Hunting Booters - A Law Enforcement Perspective

A data analytical approach to comparing DDoS police reports to real world data

J.P. Koenders

Technische Universiteit Delft

Hunting Booters - A Law Enforcement Perspective

A data analytical approach to comparing DDoS police reports to real world data

by

J.P. Koenders

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Wednesday December 7, 2016 at 3:00 PM.

Student number:	4049209	
Project duration:	March 1, 2016 – December 7, 2016	
Thesis committee:	Prof. dr. M.J.G. van Eeten,	TU Delft, Chair
	Dr. ir. W. Pieters	TU Delft, First Supervisor
	Dr. H. Asghari,	TU Delft, Second Supervisor
	E. van Veldhuizen,	Politie - THTC, Head Supervisor
	Y. Horst,	Politie - THTC, Supervisor
	P. Wagenaar,	Politie - THTC, Supervisor

This thesis is confidential and cannot be made public until December 7, 2018.

TU Delft

Summary

Throughout the four-some decades the internet has existed, it has gone through a myriad developments and evolutions at an incredible pace. Major societal change seldom comes without issues and trouble makers, the internet is no exception to this rule. Cyber- and cyber enabled crime have been on national and international news for years now, more often than not the articles speak of major outages of major infrastructure. Underlying these outages are often a simple attack methodology called Distributed Denial of Service attack or short DDoS. DDoS attacks work by overloading a certain system or service with illegitimate traffic such that legitimate users won't be able to access them. While 20 years ago DDoS attacks were still in their infancy, and only used by highly capable criminals or activists, they are now the prime example of the commoditisation of cyber crime. Attacks can be bought online via Paypal or using cryptocurrency and only require a target IP address to be executed, hence can be executed by almost everybody. The result of these attacks are large financial losses on the side of the victim, be it due to loss of potential customers or protection services to mitigate the attack. While hacktivisim or a simple joke used to be the prime reasons for attackers to execute DDoS attacks, intentions have changed and DDoS for bitcoin or in other words extortion with threat of distribution of a system or service has become a frequent reality.

A lot of research has been done on this phenomenon, in particular the technical methodologies used to execute these attacks, from botnets to amplifiers or reflectors, and even the combination of the two. Likewise, mitigation techniques such as filtering malicious requests or identifying compromised bots have been analysed many times by academia around the world. Less focus has however been given to how both the attackers and the victims experience the attack. This research does just this by taking the unique opportunity of viewing the DDoS problem from the point of view of the Dutch police and comparing that to known DDoS statistics. The Dutch police is currently undergoing a major restructuring of its forces, as part of this the national high tech crime unit is seeking to pass DDoS investigations on to the regional forces in order to clear up resources for more technically advanced forms of cyber crime. As part of this process, the Dutch police would like to get a deeper insight in their own operations as well as how these compare to the DDoS environment as a whole in the Netherlands in order to provide regional forces with as much information as possible. Currently the police has no information regarding the official reports coming in and what the characteristics of these case are. In part this lack of knowledge is due to the ageing administrative system used by the police offices around the Netherlands, which lacks provisions for easy data retrieval as well as large scale data analysis. In addition to the analysis, the police seeks a number of requirements and improvements that may be used to improve the investigative process.

To supply the Dutch police with the information needed three main research questions were developed. These questions split the research into three major sections, each valuable on its own. The first research question targets the data of the police itself, focusing on how it may be queried, analysed, what the characteristics of the cases are and how they may be classified hence leading to the following research question: *MQ1: What do DDoS attacks reported to the Dutch police look like and how can these cases be classified?* The Second research question introduces the AmpPot data-set as a point of reference. The AmpPot data-set was collected by setting up eight honeypots running the 6 most prominently used amplification protocols such as DNS and NTP, and recording the requests made to them. The AmpPot data is used as a representation of the DDoS population in the Netherlands as a whole. As such this section of the research gives a deeper insight in the DDoS population in the Netherlands and how it compares to the sample the Dutch police views or in other words, what part of the population decides to go to the police regarding their attack. The research question developed for this section reads as follows: *MQ2: How does the Dutch police data gathered through MQ1 compare to data available about the DDoS population as a whole and what insights can be gathered by combining the two?* The third and last section reiterates on the results of the previous two research questions, by combining the most prominent issues, limitations and insights to recommend a number

of improvements and suggested requirements for the police reports, the investigation itself as well as the usage of data within DDoS cases, hence yielding the following research question: *MQ3: Based on the insights gathered in MQ1 and MQ2, how could reporting, investigating and re-use of data be improved?* It is the purpose of this section to provide the police with additional information useful for the transition process of DDoS cases to the regional police forces.

To provide a concise and valid analysis of the police data-set, a methodology needs to be developed to query a sample from the data-set specific to the DDoS phenomenon. Due to the inherent nature of the system, a lists of keywords needs to be used to supplement the basic categorisation feature in order to identify the right cases. The data-set then shows that the majority of cases fall into one of three types; high school curiosity, gamers and bullying and extortion. Categorising the attacks by their victim demographic, demonstrates that almost half of the cases are focused on companies, while both the educational and home broadband victims represent about 25% each. The governmental victim demographic is a small fourth category account only for 3% of the cases. The attackers whom were interrogated (mostly high school age) described that they had no knowledge of the illegality of their actions and only two-thirds knew how to execute an attack but rather used links received from friends ”to turn of the internet”. Analysing the AmpPot data-set shows that DDoS attacks are increasing over time and that especially protocols used for basic internet services such as DNS and NTP are used to execute the attack. Furthermore, looking the attack durations yields clear peaks at certain intervals that indicate the usage of booter services. Comparing the police data to the AmpPot data indicates that the victim demographics are vastly different. This difference indicates that some victims groups such as the educational sector go to the police much more often than home broadband users do. Furthermore, the educational sector as opposed to all other victim groups has a suspect in the majority of cases. Finally, combining the damages noted in the police reports with the average price paid for a booter service, indicates that costs on the side of the victim are between 3 and 4 orders of magnitude bigger than the initial costs associated with executing the attack through a booter service. Lastly, looking at the investigation itself, a DDoS case has much larger tactical character to it rather than a previously assumed digital one, as more leads can be generated using traditional types of investigative techniques. This is an important result as the regional forces whom will investigate these cases in the future have a lot of experience in traditional tactical investigations techniques and as such may approach the cases in their usual way.

Future research may focus both on a more extensive approach of the DDoS analysis and a repeat of the analysis using other types of crime to support the police. The former may especially focus on a more international approach, one example could be how various countries differ in regard to what victims go to the police. In regard to the police, this thesis represents a first step in getting a valid overview of the actual workload caused by DDoS cases. Having an extensive system in place for multiple types of crimes would make the distribution of resources much easier for the leadership.

Acknowledgements

Writing this thesis wouldn't have been possible without a great number of people whom I want to dedicate these here lines to.

At the TU Delft, my thesis committee Michel, Wolter and Hadi whom have nudged me in the right direction and share my passion for this topic, as well as Arman who on more than one occasion has taken the time to help me out with my thesis both on a conceptual and methodological level.

Eelco, Peter and Yvonne, my supervisors at the police, whom have taken the time to help me circumnavigate the bureaucracies of a large governmental organisation and have asked the right question at the right time. As an extension to that, all those within the organisation that even at a difficult time decided to, rather than keeping me at a safe distance, make me part of the team and help me with my countless questions.

The completion of this thesis marks not only end of a 10 month academic endeavour, which taught me many more lessons than I care to admit, but also the end of my academic career for the foreseeable future. My time at the TU Delft, I dare not count the years, have been colorful and full of unexpected opportunities. For all those moments, all those headaches and all those sleepless nights; I thank you all: my family, my friends, my housemates and all those souls from far and near that I've had the pleasure of working with.

J.P. Koenders
Delft, November 2016

“Privacy is something you can sell, but you can’t buy it back.”

by Bob Dylan

Contents

1 Introduction

1.1. Background

Throughout the 40-some years the Internet has existed [30], it has grown immensely; where in 2004 just 10% of the world were considered to be Internet users, this number has increased to more than 40% ten years later. Looking at data from 2014 about western European countries such as the Netherlands, this number is as high as 93.17% of the population [59]. As such, it comes as no surprise that the importance of the Internet is ever increasing and much of today's society is run via some kind of web-interface; shopping (e.g.: Amazon.com, Bol.com, Ebay.com), socialising (e.g.:Whatsapp, Facebook, Twitter), banking (e.g.: bankieren.rabobank.nl, Paypal, Bitcoins) and some government services start switching to digital mail only such as the Dutch tax authority [7]. Looking at what the Internet or rather its users have achieved in its relative short history, is nothing short of astounding.

Like with any other new invention or major societal change however, there is not only good to be found. Internet enabled crime, Internet crime and high tech crime, in other words cyber crime, has been on the rise for years and impacts more and more businesses but also internet users personally every day [11]. Looking at major news publications from around the world, barely a week goes by without some kind of article relating to cyber crime [52][14][47]. Getting a precise grasp of the financial impact of cyber crime, or even all parts that are connected to it is difficult. Researchers however suggest that the total societal costs should not only include the already large figures related to direct damages created by theft or downtime, but also cost incurred due to defence and mitigation investments as well as costs related to enforcement [3][6][54][23] and thus becomes astronomical. The U.S. Federal Bureau of Investigation (FBI) suggests that cyber crime in 2005 alone cost the country's economy more than US$67 billion or about 0.5% of its GDP [40][23].

While cyber crime knows many different faces and is constantly evolving, one of its more prevalent [16] and damaging incarnations is that of Distributed Denial of Service (DDoS) attacks [11][32]. "Distributed denial-of-service attacks are comprised of packet streams from disparate sources. These streams converge on the victim, consuming some key resource and rendering it unavailable to legitimate clients." [33]. To further exemplify the urgency of the DDoS issue, it is important to take a look at the development over time. In Q3 of 2015 Kaspersky [27] recorded a 320 hours (13.3days) long continuous DDoS attack thereby beating the longest attack in the previous quarter by 115hours. In addition, DDoS extortion [34] (a DDoS attack is launched against an institution or organisation, quickly followed by demands for a large payment in crypto currency, such a bitcoin, to stop the attack [27]) is becoming more and more prevalent and hence a more costly issue for large companies but also public institutions such as schools. The reason why DDoS attacks are so popular is threefold: They are effective; while there are ways to mitigate the attacks they are expensive and thus only an option for large corporations [3], they are easy to execute; so called booter services openly advertise their DDoS as a service offering that only require an IP address to execute an attack [39], and they are cheap; booter services offer subscriptions to their services for between $0.19 and $14.99 [18] and some providers even offer free test attacks.

While law enforcement agencies actively pursue DDoS cases world wide, DDoS investigations are notoriously difficult as the attacker usually doesn't use their own machine but rather makes use of some kind of proxy [29], one example is that of hacked machines that are part of a botnet, which are being controlled via a command & control server [5] [12]. In addition, there is often very little evidence other than network log files from the victim (organisation). Due to the relative difficulty of analysing these log files and building a case on them, the investigation is often left with the high tech crime units within the police force. As made apparent by the previous paragraphs, these teams are in high demand however and sadly limited in their capacity. In addition, it is the purpose of these teams to use their limited resources to investigate cases on the bleeding edge of innovation and to stay up-to-date to ensure swift reaction to constantly evolving criminals.

1.2. Problem Context

As mentioned in the previous paragraphs researchers have a fairly good grasp of the DDoS landscape as a whole, various researchers world wide have classified the type of attacks executed [10][48][4], the damage both economically and socially caused by these attacks, the volume these attacks can have and the victims that experience most strain etc.. These have mostly built on data made available by DDoS mitigation providers as well as honeypots and estimates made by academics in the field. However, little has been written about the law enforcement point of view and how they may experience the DDoS plaque. The position of law enforcement is however a very interesting one, as it places itself between the victim and the attacker, by accepting official reports from the one side, and investigates to find the attacker with retribution as final goal on the other. As it happens, the Dutch national police is currently undergoing a large reorganisation. In line with this reorganisations, the police is seeking to get a deeper insight in types of crimes that are being reported as well as how this compares to known statistics of the occurrence these crimes.

Due to a combination of the inherent properties of the ageing police administrative system that require the analyst to search through entire plain-text reports rather than categorised files, and lacking quality of said reports, the police has very little such statistics at the ready. For cyber related crimes, the available information is even more limited as the system was simply never built to accommodate complex cases such as DDoS attacks. On a basic level this lack of knowledge leads to a number of troubles such as a mismatch of resources compared to the magnitude of the issue as well as a lack of specialised knowledge. On a more concrete level, the police does not know whether the reports made are a good representation of what actually happens, and whether certain types of cases are left entirely un-investigated. In either case, the leadership of the Dutch police is lacking basic managerial knowledge in order to steer the reorganisation.

1.3. Problem Definition

As of this moment, "Team High Tech Crime" (THTC) of the Dutch national police force, which will act as the problem owner within this research, handles most cases regarding DDoS attacks. This is mostly due to a lack of knowledge on other levels of the police body. Since the reorganisation however, it has been decided that regional police forces need to take the reigns in more cyber crime cases to clear up investigative resources for more complex cases. Hence it is the plan to move DDoS investigations to the regional forces. To make the transition as smooth as possible and to be able to sketch a picture of what is being transferred to the regional forces however, THTC would first like to get a better understanding of the DDoS investigative landscape and how it compares to the entire DDoS environment. In addition to a lack of descriptive information regarding the DDoS reports made, the police and that includes Team High Tech Crime have no formalised approach to DDoS attacks. New cases coming in are always handled as a entirely new phenomena, knowledge is only transferred by investigators that have been part of a previous investigation. Within the high tech crime units of the police these experienced investigators are present in sufficient numbers, at the regional level however very little experienced personnel is available. Due to this investigations are slow to start and often end up unsolved.

As mentioned in the background section of this chapter, DDoS attacks have been around for a long

period of time. During that time the DDoS phenomena has evolved, both on a technical level as well as how they are used by attackers and how it effects victims. Due to the flexible nature of DDoS attacks, there is no easy step-by-step manual that can be developed and handed to investigators. There are however a number of basic requirements to the varies entities connected to an investigation that if satisfied, make for a much easier investigative process. These may regard the questions asked whole recording the official report, secondary data-sets to consult, data to request from third parties and an indication of which team should investigate certain types of cases. These requirements will come forth from both limitations and issues discovered during this thesis as well as results from the various analysis steps. These can then in turn be used as a basis to educate regional forces on their investigative strategy.

1.3.1. Problem Statement

DDoS attacks are one of the most devastating and costly type of cyber crime today. Due to organisational changes within the Dutch police force, DDoS attack investigations will need to be facilitated with the regional police forces rather then the high tech crime teams. As of now the police has little knowledge about the reports they have received and investigations they have executed that could be used to support that change. Hence, the police requires a thorough analysis of their own data. Furthermore, this data needs to be put into perspective in regard to the entire DDoS environment.

1.3.2. Research Objective

It is the objective of this research to supply Team High Tech Crime with a thorough understanding of the current state of affairs of DDoS investigations within the Dutch police and how they compare to the entire DDoS world as it is perceived throughout the internet. Additionally, it is the objective to provide the Dutch police with a set of investigative requirements that should be implemented in order to facilitate a better investigative process in order to aid regional police forces in the execution of a DDoS investigation.

1.3.3. Knowledge Gaps and Deliverables

As stated in the previous sections, there are a number of knowledge gaps the Dutch police or more specifically THTC needs filled. First and foremost there is the internal information gap, or in other words, what information is available and where, how can case files regarding a certain type of crime be extracted from the police administrative system and how can these files be categorised and analysed. The second gap is the external knowledge gap regarding the DDoS environment and how this environment looks like especially in regard to the victim demographics. Lastly there is the data analysis gap regarding the comparison between the DDoS issue as it is perceived by the Dutch police and the DDoS environment as a whole. In accordance with these three knowledge gaps, 4 deliverables will be be developed. 1. A concise methodology to query DDoS cases from the police administrative system 2. A thorough analysis of the DDoS case files both in regard to their content as well as their quality or lack there of 3. A comparison between the police data-set and the DDoS environment as a whole 4. A list of requirements to improve the investigative process.

1.3.4. Scientific Relevance

The scientific relevance of this research is two fold. Firstly, the data analysis will give a deeper insight in properties of DDoS attacks as they are perceived from the law enforcement point of view. This is an entirely new data-set that may shed some light on the type of cases victim deem worthy of a police report. Secondly, it shapes the basis for future research in regard to allowing researchers a glimpse in the vast, albeit difficult to access pool of data law enforcement has and tangibly, adds to the knowledge about the perceived cost of DDoS attacks as the reasoning and aims behind police reports are analysed.

1.3.5. Societal Relevance

As stated in the first section of this paper, DDoS attacks are a constant and big issue for society; customers can’t access their online banking, victims can’t make a statement via the police’s website

and citizens can't submit their tax review. This is where the societal relevance comes forth, adding to the knowledge about the cases the police investigates, and what the files contain adds to the quality of future investigations and as such may improve the success rate of cases. Additionally, the comparison may yield a victim demographic that the police was previously unaware off. Lastly, this research will provide the police with a number of requirements that may be used to improve future investigations and yield improvements in both the information available on investigation as well as on the investigative process.

1.4. Research Question

Based on the previous sections, one may distill two main research goals; an analysis of the current DDoS investigative population based on police case files, and the comparison of that analysis to the DDoS population in the Netherlands as whole. The three main research questions coming forth from these two goals go as follows:

MQ1: What do DDoS attacks reported to the Dutch police look like and how can these cases be classified?

MQ2: How does the Dutch police data gathered through MQ1 compare to data available about the DDoS population as a whole and what insights can be gathered by combining the two?

MQ3: Based on the insights gathered in MQ1 and MQ2, how could reporting, investigating and re-use of data be improved?

In order to answer the above main-questions, the following set of sub-questions has been devised. These questions will lead the research throughout the various stages and ultimately lead back to answer the main questions and bring forth the previously discussed deliverables.

MQ1.

SQ1. How can cases files related to DDoS attacks be queried from the police administrative system?

SQ2. How can these case files be analysed to categorise them?

SQ3. How are the DDoS attacks reported to the Dutch police characterised on a descriptive level and what can be said about the identified attackers?

MQ2.

SQ4. What does the current DDoS population in the Netherlands as a whole look like and how does it compare to the cases recorded by the police?

SQ5. What additional information can be gathered by combining the honeypot data-set with the police dataset?

MQ3.

SQ6. Based on MQ1 and MQ2, what requirements may be identified in the police reporting system in order to improve the information basis to data querying and analysis?

SQ7. Based on MQ1 and MQ2, what lead generators may be identified for the investigative process in order to improve the investigative process both on a tactical and digital level?

SQ8. Based on MQ1 and MQ2, what requirements may be identified for the use and recycling of data across various DDoS cases in order solve cases more efficiently?

1.5. Research Approach

To ensure a logical and valid flow of the research, a methodological framework will be used as a basis to for the research approach and as such to answer the above stated research questions. The development of this framework in large takes cues of of the qualitative analysis framework used [31]. Where the framework used for this thesis (depicted below) mainly differentiates itself, is in the addition of the stakeholder analysis, methodology development and meta analysis sections. The addition of these sections originates from the unique character of this research, which is in large due to the stakeholder in question as well as the data provided. While for many governmental and private organisations there exists a certain status quo to refer to, the police is somewhat of an outlier, not least because of it tasks, but also because of the way it is governed and the value data analysis and interpretation of in it. To capture this very difference, a stakeholder analysis will be executed, allowing the researcher to identify the unique characteristics that will provide an environment to place the results of the coming chapters in. Due to the current position data analysis has within the organisation, resulting in a lack of previously developed data collection and analysis methodologies, the methodologies section will preface the empirical data sections. Lastly, the meta analysis phase will reiterate not only on the results found in the data analysis sections, but also include insights gathered during the stakeholder and literature analysis as well as the methodology section to create a set of investigative requirements.

Figure 1.1: Research Framework

Research Methodologies

Inline with the research framework proposed above, this thesis will make use of four research methodologies. These four methodologies, shown in the middle pillar of the Research Approach and Structure figure 1.2 below, in turn support the seven chapters of this paper. The first research methodology is a literature and stakeholder analysis, the purpose of this methodology is to build a basis both on the technical level as well as on a stakeholder level on which the rest of the thesis can build. The stakeholder part of the analysis is vital as it will provide a deeper insight in the unique position the police is in both due to its inherent tasks as well as the current reorganisation. Additionally, it will shed light on why data analysis or even the provision of data is as underdeveloped as it currently is. The literature analysis focusing on the technical side of DDoS attacks is to give the reader a better insight in the various developments as well as mitigation difficulties connected to the attack type. The second research methodology is the methodology development. As mentioned in previous sections and as will become more apparent in the stakeholder analysis, the police currently has no validated way to

gather empirical data about the types of police reports they receive or the type of investigations they have executed. In order to be able to execute the rest of this project, a new methodology needs to be developed in order to gather a valid DDoS data-set from the administrative system. The third research methodology is empirical data analysis. This method will be used for the largest part of this thesis and will make use of two main data-sets, the police data-set and the AmpPot data-set. The purpose of this methodology is to provide an insight into the police reports filled specifically about DDoS attacks and how they compare to the DDoS world as a whole. Lastly, the meta analysis will reiterate not only on the results found in the data analysis sections, but also include insights gathered using the other research methodologies to create a set of investigative requirements.

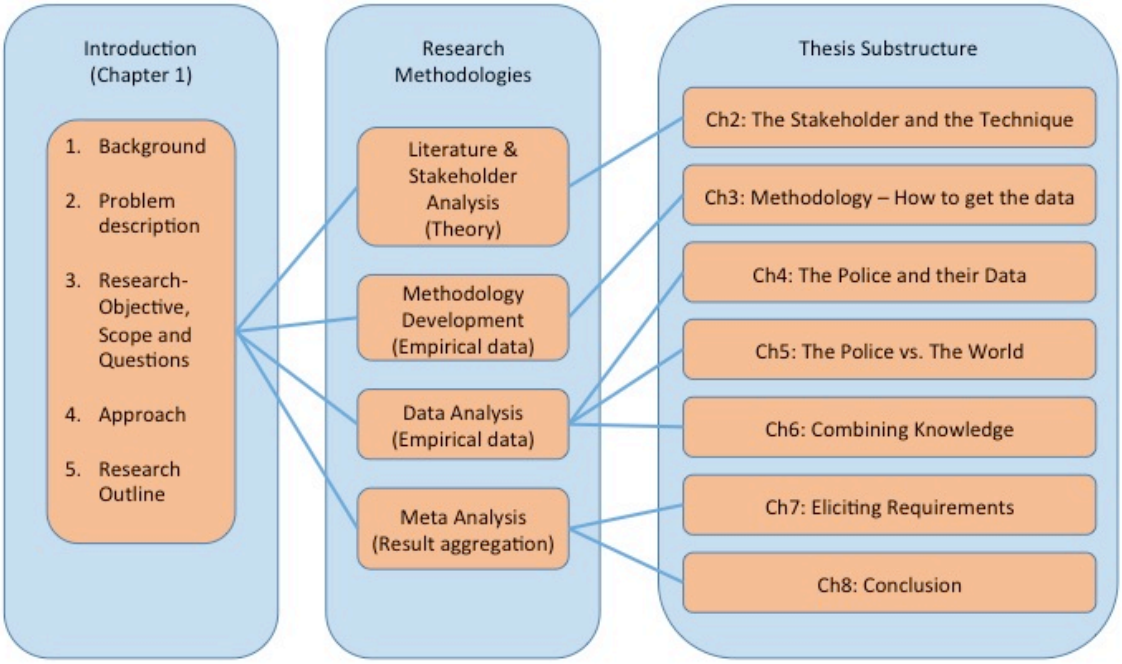

Figure 1.2: Research Approach and Structure

1.6. Project Scope

To get a clear understanding of the scope of this research, one must first look at delimiters used to create the research space. The first delimiter is the origin of the data utilised throughout the coming chapters. In essence there are two major suppliers of data, the Dutch national police and with that its internal administrative system, containing official police reports, as well as a DDoS through amplification data-set dubbed ”AmpPot” [28], which was gathered through a set of honeypots. The AmpPot data-set in the state it was used for this research was enhanced by Arman Noroozian and his team of the TU Delft for their own work [39] and as such contains additional data, such as the addition of historic DNS data in order to get an understanding of the domains connected to the victim IP at the time of attack, that was used throughout this paper. The second delimiter is that of the geographical location. While the Dutch police data-set naturally only contains attacks on victims located in the Netherlands, the AmpPot data-set is international and as such contains victims from all over the world. To ensure the validity of this research however, before kindly providing the data-set, Arman Noroozian limited the original data-set to only contain victims located in the Netherlands. While this may create limitations in terms of international validity, the Netherlands is one of the most active countries world wide in regard to cyber crime, both in terms of the occurrence of crime as well as the enforcement of laws in this area and as such lends itself perfectly as a sample case. The third and last delimiter is that of time. As the AmpPot data-set contains attacks from the years 2014 and 2015, most of this research will utilise data from this time frame. The scope of this paper will be within said boundaries, focusing on the exploration of the police data as well as placing that data into its environment in order to judge how representative it is.

1.7. Research Limitations

There are a number of limitations or potential issues that may arise from the above proposed research. First and foremost, the data-sets. The AmpPot data-set used as a comparison data-set is specialised to DDoS amplification/reflection type and as such ignores other types of DDoS cases. The police data-set on the other hand contains all kinds of DDoS types and as such the two data-sets are not a perfect match. However, to prove that amplification/reflection type attacks account for the vast majority, a third party data-set provided by the Dutch scrubbing agency will be used. The second limitation regarding the AmpPot data-set is the limited coverage in regard to victim classification. A quick analysis shows that roughly 50% of the attacks do not contain a classifier. To still be able to use the data-set as a whole, the other 50% were classified by hand and as such the data provides an excellent accuracy level. The police data-set knows a number of problems as well. First of all the data quality. The first point of contact between the police and victims of DDoS attacks are usually the local precinct (wijk teams). The officers at these precincts are trained to cope with traditional types of crime, ranging from theft to violent crime, cyber crime however is a completely new and unknown area for them and as such they lack vital knowledge to note the most important details central to any DDoS attack. The issue in fact is so big that in some cases DDoS case files may not even contain the phrase ”DDoS” much less the magnitude or targeted IP address. To ensure a valid search result and to minimise the time needed to get the right case files, the methodology chapter will focus on developing a method to collect these case files from the police administrative system. The development of this method is not without problems itself as the age of the system combined with queries needing to search through plain text files provide for a tedious querying process.

Thesis Outline

Throughout the various chapters of this thesis the previously posed sub-questions will be answered, finally leading back to the main-questions. Throughout the coming subsections the various chapters will be detailed and the sub-questions will be distributed. Figure 1.3 below is a graphical depiction of that distribution, noting what the corresponding chapter and sub-question(s) are. The three main-questions are answered throughout the conclusion chapter and as such are all noted at the bottom of the graph. It’s important to note that the graph below is lacking the first chapter as it doesn’t contain any research methodologies nor answers any research questions.

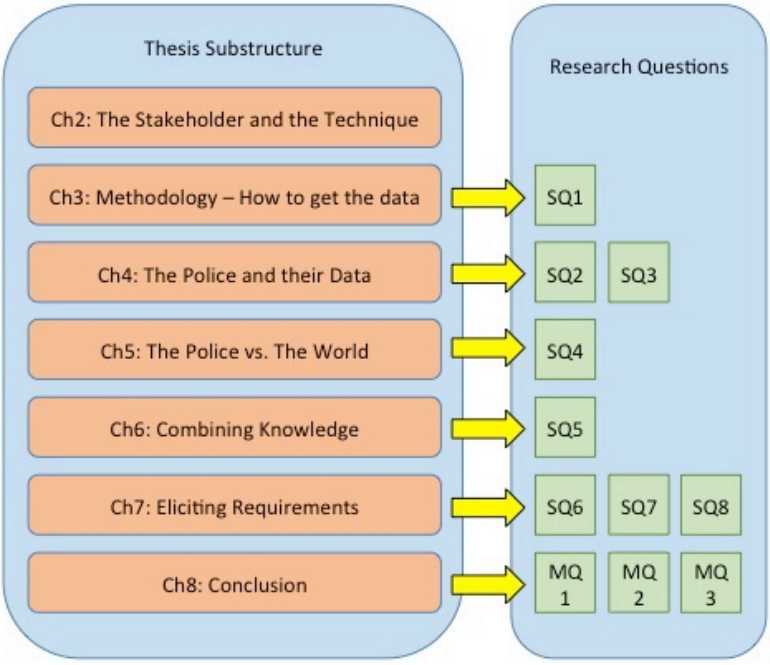

Figure 1.3: Chapter Question Outline

Chapter 1: Introduction
The first part of this thesis, the introduction will give a basic insight in the topic DDoS as a whole, identify the problem owner, give an insight in the problem at hand and note the research objectives and knowledge gaps before stating the research questions and research approach.

Chapter 2: The Stakeholder and the Technique
The second chapter consists of two major parts, a stakeholder analysis and a literature review of the DDoS attack type. The purpose of the stakeholder analysis is to zoom in on the the problem-owner and to give an insight in the special character it has. The purpose of this analysis is to lay a groundwork upon which the later chapter can be built. It will also provide the necessary knowledge to understand why certain decisions impacting the usage of data analysis were made. The literature review will provide both a technical basis as well as a descriptive analysis of the most important trends in the DDoS world. Both of which will provide a back story to the rest of this thesis and provide the reader with a better understanding of the DDoS issue.

Chapter 3: Methodology – How to get the Data
The third chapter will focus on the methodology development part of the thesis. The methodology is a vital part of this project as there is currently no way to get an overview of the recorded police reports nor the investigations that were executed in certain crime category. As such this method will not only be of importance for this project, but may also serve as a basis for future research. The development process is based on various test trials to develop valid keywords and search filters. Next to that the results of the data retrieval process as well as it's limitations will be discussed. The final result of this chapter will answer the first sub-question.

Chapter 4: The Police and their Data
Throughout the fourth chapter the second and third sub-question will be answered. It is the purpose of this section to get a better understanding of the DDoS attacks that reach the police via official police reports. As such this chapter will make use of the data gathered through the third chapter and analyse the data both on a more general descriptive level, as well as a more specific level, defining victim categories and analysing interrogations, thereby answering the second and third sub-question. To execute the analysis the Python statistical add-on Pandas will be used in addition to basic excel calculations. The importance of this chapter is rather big, as it provides a baseline for the coming chapters, but also supplies the police with a first look into their own data-set and what their actual work-out put is in regard to DDoS attacks.

Chapter 5: The Police vs. The World
The fifth chapter is much of a continuation of chapter four as it uses both the data analysis research methodology as well as the results of the previous analysis as a baseline. It is the purpose of this chapter to compare the results found in the police data-set with data from the DDoS population as a whole in the Netherlands. This will provide a unique insight in the victim demographics. The problem-owner is especially interested in whether there is a certain group of victims that categorically avoid making an official report, or whether certain types of attacks combined with an extortion scheme are more prevalent than previously expected. The forthcoming results will allow for answering the fourth sub-question.

Chapter 6: Combining Knowledge
The sixth chapter shapes the last of the data analysis chapters and continues the use of both the police and AmpPot data-sets. Where chapter five compares the two data-sets, the sixth chapter combines the two, to get a deeper insight in the attacks the police are receiving reports of. Especially interesting is the monetisation side of DDoS attacks, or in other words how much damage can be imposed on a victim for a certain amount paid to a DDoS service provider. Chapter six will answer the fifth sub-question.

Chapter 7: Eliciting Requirements
The eliciting requirements section will make use of a meta analysis like research methodology to combine information and issues gathered throughout the previous chapters in order to provide the police with a set of suggestions and requirements for DDoS investigations as well for the data analysis and usage. The purpose of this chapter is to 1.Improve the investigative process 2.Provide a basis that may

be used when transferring cases from the high tech crime units to the regional police forces. Through this, the sixth, seventh and eighth sub-question will be answered.

Chapter 8: Conclusion

The eighth and last chapter will shape the conclusion of this thesis. As such the results of the eight sub-questions will be revisited and the three main research questions will be answered by reiterating on the results found in the previous chapters. Furthermore, the contribution of this thesis both on an academical as well as on a social level will be discussed. Lastly, suggestion for future research will be posed.

2

Literature and Stakeholder Review

2.1. Stakeholder Analysis

Throughout this section the relevant stakeholders will be identified and the problem-owner will be inspected on a more detailed level. The purpose of this section is to draw a picture of the environment Distributed Denial of Service attacks happen in, and what the unique position of the police is. It's important to note that due to the inherent nature of DDoS attacks, writing a full list of all potentially involved parties is near impossible, yet the list below gives an indication of what kind of actors are impacted or directly connected to the issue.

- Booter service providers
- Anti-DDoS service providers
- Internet service providers
- Schools
- Banks
- Police
- Hosting providers
- The public prosecutors office
- Attacker
- Criminal organisation (in extortion cases)

While many of the above mentioned actors are large organisations or sometimes even umbrella organisations positioned above those organisations, most of them may roughly be categorised in the victim category. This is due to the simple architecture DDoS attacks use, which allow them to target any entity with an IP address. It's important to note that more often then not there is evidence for presence of a criminal organisation, this however excludes extortion cases who have been known to be connected to larger criminal groups such as the Armada Collective [42] and LizardSquad [56]. Most of these extortion cases call for payment via the virtual currency Bitcoin as these are hard to trace and can easily be transferred to other accounts. While usually this would entail the addition of a financial partner, the Bitcoin currency is based on a decentralised system and as such no banks are involved and hence no additional actor is introduced. The table below depicts the actors with their corresponding role, again victims are simply listed as one entity as their role is the same across the board.

Table 2.1: Actors and their roles in the DDoS environment

Actor	Role
Booter service providers	Provide DDoS attacks as a service for payment
Anti-DDoS service providers	Provide DDoS mitigation services by re-routing and filtering the traffic directed to the victim
Internet service providers	Provide an Internet connection and are the first line of defence against DDoS attacks
Victim	The party that is the recipient of the DDoS attack
Police	Investigate the attackers as well as the facilitating parties such as booters
The public prosecutors office	Uses the cases brought forth by the police to prosecute attackers a facilitating parties using Dutch criminal law
Criminal organisation	Group of criminals utilising DDoS attacks to extort organisations by threatening to take a site/service down if no payment is done
Attacker	Entity that utilises the DDoS methodology to attack a victim

Although many of these actors, take ISPs and the financial sector as an example, have their ownm initiatives to fight DDoS attacks; and others such as Akamai, Kaspersky and Cloudflare, being represented by the Anti-DDoS service provider bullet above, have built their entire business model around the issue, this thesis focuses on the party that directly targets the attacker behind the trigger of the virtual DDoS gun, the police. As such this party is unique, as all other parties focus on mitigating the symptoms of an attack i.e. the network traffic or try to make underlying methodologies such as IP spoofing impossible to use. As this thesis focuses on the police as a problem-owner and an internal problem within that organisation, the number of actors on general level shrinks down to the three that are directly in contact with the police through the DDoS investigative procedure. The figure below gives an indication of how these actors interact.

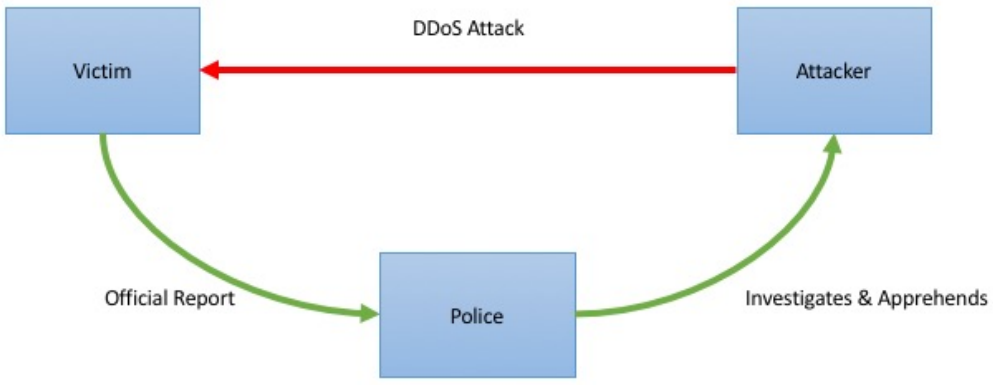

Figure 2.1: Three Central Actors

The unique position of the police is obviously due to the mandated package of tasks the government has put on it. In its core these tasks are categorised into six main tasks: Ensure safety, prevent and combat crime, ensure public order, detect offences, provide help in emergencies and carry out tasks for the department of justice [41]. These tasks are independent of the type of crime and as such also apply to cyber related crimes. The Dutch national police has had an eye on cyber crime for a number of years already and as such has produced a number of reports not only identifying the various types of cyber crime as well as the newest developments within the field, but also where it sees itself within that area [8] [49]. The report notes an interesting distinction between three levels of cyber crime and how it identifies these. The first level is cyber enabled crime, denoting traditional crime that is in some way

supported or facilitated through cyber. In this day an age one can imagine that almost any crime is part of this category, be it for the digital communication criminals use or the online market place they use to sell products. The second type is cyber crime, the police defines this category as any crime that makes use of automation for either the collection or the processing of data. Examples of cyber crime would be defacement of a website or skimming as well as cyberstalking. Lastly, there is hight tech crime which is defined as the bleeding edge of the cyber crime, using the newest developments to reach high impact.

To cater to cases that fall under the third category, the police has a special unit called Team High Tech Crime (THTC). This units uses, next to the definition of high tech crime as noted above, six pillars to identify cases that fall within its jurisdiction. These pillars are black swans, facilitators, ransomware, vital infrastructure, banking sector and botnets [49]. To be able to handle these kinds of cases, the unit has a 50/50 policy meaning half of the unit's members come from the ”traditional” or how the police calls it ”tactical” part of the police body, while the other half consists of specialised people either coming directly from higher education or from the high tech industry. As mentioned in the introductory chapter of this thesis, currently the high tech crime unit handles DDoS cases. Looking at the defining characteristics of cases this unit usually handles, DDoS cases certainly fall outside of the scope. Additionally, as the internet is growing, so is the amount of cyber crime as well as the speed of development in this area. In order to stay on the bleeding edge, THTC has to take on more and more cases. Due to this all DDoS cases are supposed to be handled by regional forces as the previously defined cyber crime cases fall under their jurisdiction.

While the police traditionally speaking sees recording of data i.e. police reports, as a central tool to fulfil their set of tasks, the analysis of said data is not. To understand the current situation the police finds itself in, one must look no further than the digital registration system used to record the cases. It is no secret that the BVH system is ageing and was built in a time where priorities and technical capabilities were different [58]. The underlying purpose of the time was to combine various technologies to get all information into one system. This was to digitalise information more than it was to allow for large scale data analysis [58]. When designing the system, cyber crime was still in its infancy and hence was not included in the design procedure. Due to this, the system lacks fundamental categorisation options vital for cyber crime cases and thus does not facilitate them very well [25]. Additionally, the system was built to coordinate single cases but not to run statistical analyses over the entire database. Due to this environment and also due to the priorities set within the investigative community no analysis of the case load has been executed.

2.2. Literature Review

Throughout the second part of this chapter the DDoS phenomena as a whole will be inspected. As such this chapter has two main goals; give the reader a better understanding of the technical methodologies underlying the DDoS attack methodology, and to note the most recent developments and most urgent issues internet users are fighting with. This section will entirely be based on the literature review research methodology, mainly focusing on papers published by academia as well as industry reports. While many of these industry reports give quite a detailed account of the various types of DDoS attacks, one may argue that they would suffice as a point of comparison for 5. The trouble with these reports however is that they are created by companies that directly profit from making these attacks look as bad as possible. While there is certainly a scientific basis in these papers, they are mainly a marketing tool and as such should be viewed with a grain of salt. Hence, while the reports provide a great basis to introduce the topic of DDoS attacks, later finings section will make use of a honeypot data-set made available by the TU Delft.

To get a better understanding of the DDoS landscape and how it functions today, one must first understand the basic properties and categories of DDoS attacks. To start lets look at the 7 layer or Open Systems Interconnection (OSI) model, (see figure 2.2) which is a ISO standard of representing the basic building blocks of an IT system as well as a framework for the definition of protocols [61]. Due to the direct relation to network protocols, it gives a great overview of potential target layers for DDoS attacks.

The various layers of the model can roughly be categorised into two basic types of layers, applica-

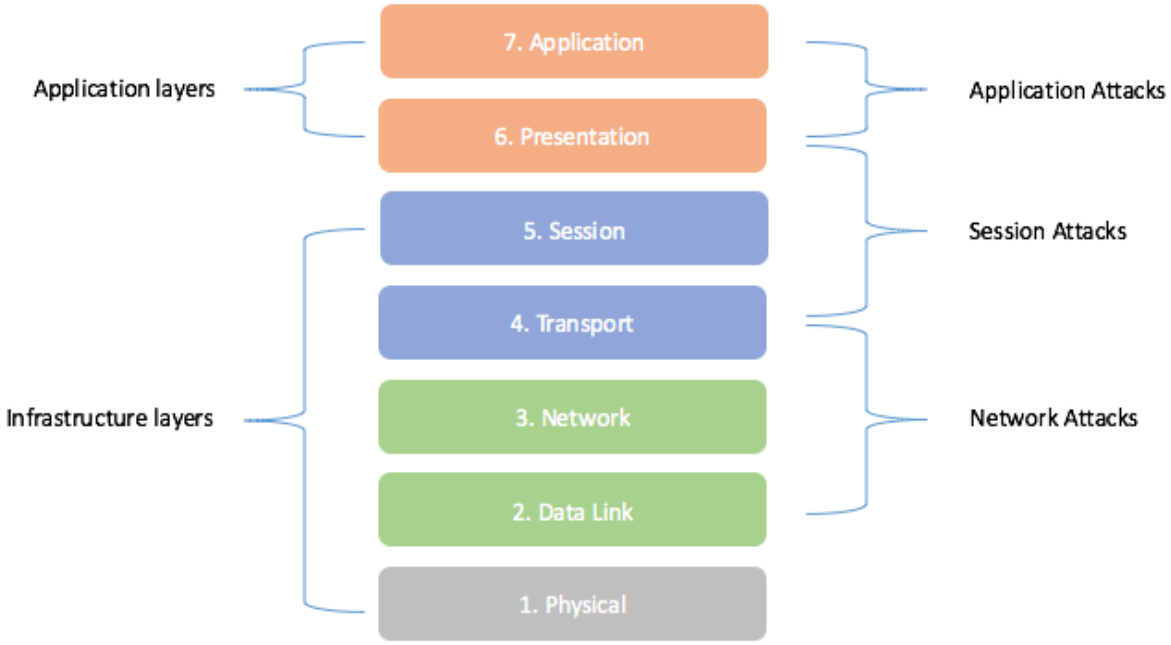

Figure 2.2: 7 Layer OSI model and Attack Layers

tion layers and infrastructure layers, as depicted on the left hand side of the figure. While the application layers carry the operating system or application in use, the infrastructure layers takes care of everything needed to support the application layers such as network package delivery and routing. Attacks on the application layers by way of targeting application weaknesses and exhausting resources[60] account for the minority of attacks, depending on the report ranging between 3% [1] and 17% [38], of all DDoS attacks measured in 2015. One possible reason for this is that these types of attacks are more advanced and are both harder to execute but also harder to detect as the overall network load may be very small yet have a very big effect [57]. Conversely, infrastructure layers are being attacked for the majority of the observed cases, again ranging between 97% [1] and 83% [38] depending on the source. Infrastructure layer attacks contain both layer 3 (Network) and layer 4 (Transport) attacks, where layer 4 attacks account for the majority [20].

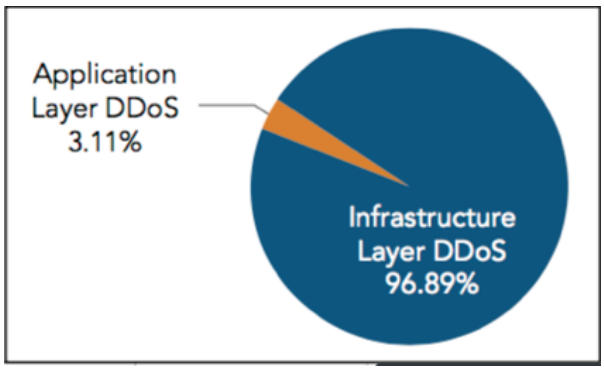

Figure 2.3: Attack layer distribution according to the Q4 2015 Akamai report [1]

Zooming in a bit more, one can split the infrastructure layer into two subcategories; session attacks and network attacks, as depicted on the right hand side of figure 2.2. Depending on the sublayer type, the attacker may make use of a variety of protocols or attack types. One of the reasons why the infrastructure layer attacks are becoming more wide spread and take up such a large percentage of the overall DDoS world, is the development of amplification attacks. Amplification attacks use one of

the attack types as noted in the network or session categories in the figure below but make use of a proxy to amplify the amount of packages that arrive at the victim [43].

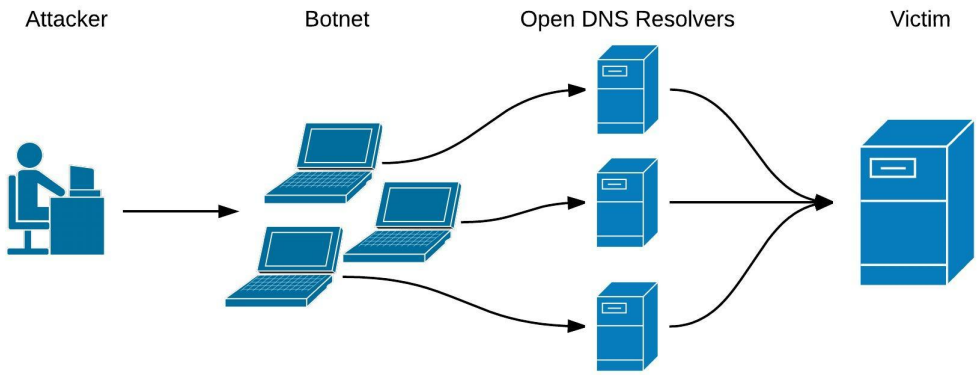

Figure 2.4: Graphical representation of a DNS amplification attack

To do so the attacker sends a certain request to a server but hides their own identity by, rather than using their own IP address as a return address show that of the victim. To get a better understanding on how these request and their corresponding responses look like see A.6. The trick is to make the answer expected from the server to be larger than the original request, hence amplifying the amount of traffic that is sent towards the victim [43]. The figure above is a graphical representation of one such attack utilising the DNS system. As an additional parameter the attacker makes use of a botnet as to create more traffic and to be able to reach more DNS servers.

2.2.1. Developments in the DDoS world

While DDoS attacks are no new phenomena and some of the most relevant research within this area dates back to the early ’00s [32][22][24], recent data indicates that the DDoS attack methodology is on the rise again [26]. Major actors in the industry such as Akamai [1] and Verisign [56] as well as Kasperski [27] see roughly a 50% increase of DDoS attacks onto their customers in both the third and fourth quarter of 2015. With that overall increase in attacks, two major developments have established themselves in the DDoS world. The first one is the commoditisation of DDoS attacks through services and the second one being extortion cases [19] in which DDoS attacks are used as a tool to pressure organisations to pay a certain ”safe fee”. As such both of these developments represent the ongoing effectiveness of the attack type as well as the limited effective mitigation that is available to protect potential victims.

Taking a better look at the commoditisation of DDoS attacks and where this development is coming from, directly leads to an interesting development in the attack properties all three companies have noted in their reports. There seems to be a disproportionate increase of attacks on the infrastructure layers (notably layer 3 and 4) where popular amplification attack types such as DNS amplification account for almost 60% of all DDoS attacks. Within the infrastructure layers, layer 4 attacks (often referred to as SYN flood attacks) seem to be especially popular [28] [13]. This increase is very much related to the overall increase of the usage of the amplification attack methodology [13][28][43]. As explained in the previous section, the methodology makes use of a spoofed IP address to forward traffic to the victim. One of the features of this attack that make it so popular is that without having a large infrastructure of their own, relatively large scale attacks can be launched and public servers such as DNS services can be abused without having to be hacked in advance to the attack. Continuing down the same path, using a simple botnet increases the attack vectors substantially. Due to this relatively efficient, cost efficient and easily scalable technique building a powerful infrastructure isn’t too difficult. Join that with a ever increasing demand for DDoS attacks especially in the gaming world [44][13][1], and a new industry is born. The DDoS as a service phenomenon is a very interesting one as it allows customers to choose from a wide variety of packages and sometimes even custom tailored attacks.

Customers usually pay for the attack type or combination of various protocols, the bandwidth and duration of the attack [21]. Next to that, customers can pay by Paypal or cryptocurrencies such as bitcoin, the latter providing a certain sense of privacy and protection for both the service provider as well as the customer. As such it shows clears signs of a mature market, which handles its customer's requests professionally and is able to quickly respond to new demands. Hutchings and Clayton in their paper "Exploring the Provision of Online Booter Services" get in contact with ten of these service providers to get a better understanding of their views on the matter, and while they find the main motivation to be rather simple "easy money", they almost unanimously believe to provide a valid service for entities that may want to stress test their own infrastructure. Another interesting note is that the majority of the actors believe their services to be legal and thus don't expect any legal repercussions [18].

While the increase in amplification attacks and with it the increase in booter services portray the advance in DDoS tooling, DDoS extortion cases portray the advance in utilisation of the attack methodology. Historically DDoS attacks were used with the goal of service disruption or plain vandalism in mind. DDoS extortion uses the fear of said effects to extort victims [19]. The basic notion entails that the attacker asks the victim to pay a certain amount, usually by means of a virtual currencies such as bitcoin, or else a attack will occur [27]. Again, this development shows a clear evolution of the DDoS attack methodology, were the damage done by an attack is not so much the final goal as the tool used for financial gain. For many large enterprises especially those in volatile industries that are built on customer trust, such as the financial industry, and may be inclined to pay the ransom instead of experiencing a possible decline in customer trust. While economic model behind such an attack may be even easier to execute on a smaller level. Small and medium enterprises often can't afford the large financial burden of corporate DDoS protection that are used to protect financial institutions and large multinationals. For them an extortion case may end up being a simple calculation, what is more expensive, getting DDoS protection or simply paying off the attacker [34]. On the side of the attacker the calculation of may be of equal ease, attack a set of companies all at the same time and some are bound to pay. All that would be needed would be one or multiple booter subscriptions that would be financially recovered by the payed ransom.

3

Methodology

As mentioned in the previous chapter, there are a number of challenges in using the Dutch police data in the raw. This chapter will focus on the methodology used to retrieve and analyse the data-set provided by the Dutch national police as well as which tools were used and what limitations one must account for when viewing the results.

3.1. The Dutch national police as a data provider

Within the Dutch police force a variety of software suits are used to capture and archive reports. As with many large organisations, different software choices were made at the various levels of the institutions. So while the national police uses a relatively new system called Summ-IT that allows for a complete work-flow from the report down to sending files to the public prosecutor [36], the other units within the Dutch police use the ageing BVH system that according to official evaluative reports fights with many issues due to it's complexity and interconnection with many third party systems [58]. For this research specifically however, the hardship lies both in the categorisation and formulation of the reports made. Due to the original intend of the software as well as the environment it was developed in, much of the system is focused on traditional types of crime. As such there is a plethora of categories too choose from for the various types of vandalism, theft and physical harm that may be reported. For cases relating cyber crime however, very little in this regard is available and most officers will simply categorise anything remotely related to technology under one single large category called ”Computercriminaliteit” (computer criminality). Since this means that there is no way of easily filtering DDoS cases from all these other cases, a text search had to be executed that would go through the entire official report set up by the police officer that recorded the report. As these reports are traditionally written in plain text, a secondary internal police tool called ”Cognos” can query through them looking for specifically assigned key words. As one may imagine, due to the age of the database, these queries can take rather long. To ensure a swift execution of the data as well as a relevant and up to date date-set, which would be usable in an analysis, the search was limited to the last year.

In addition to the categorisation issue, the official repots differ widely in quality of the information noted as well as how concise they are. Furthermore, there seems to be a relatively loose interpretation of of the spelling of cyber crime related jargon. This is mostly due to the underlying training the police officers have gone through which, much like the the development of the administrative system, focuses on traditional types of crime. As such cyber crime related knowledge is often lacking leading to the comparatively lower quality reports in these cases [25]. One very interesting observation made while reading the reports was that officers would use one of three ways of indicating a DDoS attack. 1. Using the official name ”DDoS” or ”Distributed Denial of Service” 2. Naming the tool/service used ”Booter” or ”Stresser” 3. Describing the symptoms of an attack ”Network down”, ”Server down” or ”no internet”. Due to these possible variations, various search queries were executed to get the final data-set utilised in this project. These queries were built upon the various ways an officer could spell the abbreviation ”DDoS” such as ”D-DoS” and ”D-D-o-S” or simply ”DoS”.

3.2. Data Retrieval

Throughout this section the process of data retrieval with the Dutch national police will be detailed. The purpose of this extensive procedure is to ensure the validity of the data-set and as such to counteract the issues noted in the subsection 3.4. It is important to note that due to both legal and privacy concerns, no information from the reports will be mentioned verbatim in this or any other section of this paper. Any and all data snippets that will be shown contain fictional data and are used purely for demonstrative purposes and as such cannot be traced back to real cases or the people linked to them. The case files used for this research originate from the BVH system that is used by the regional police forces to store and archive any and all official reports in [58]. The BVH system is renowned for its many issues, one being its limited search functionality, especially in regard to nationwide searches. To do just that, search reports from the entire country utilising one query, collecting the results in a structured overview, and making it all easily exportable, the police makes use of the Cognos system. Cognos is a search-machine like tool that allows the user to create extensive queries for searching through police report files. As such it has the capabilities to not only find reports that are marked with a certain category or other defining characteristic, but can also search directly through the text of the report itself, seeking pre-specified keywords. This very feature was used to gather the data for this research.

The search procedure itself was based around a set of filters that would limit the overall size of the Dutch police database by first filtering all results for the general and broad computer criminality category and second, for the time-frame of 1st of January 2014 until the very day the query was executed. The purpose of gathering data over such a long period of time was to allow for an extensive descriptive analysis containing as many cases as possible, while also allowing for easy restricting of the data to be used in a comparison with the AmpPot data-set which contains data from the the beginning of 2014 through the end of 2015. To ensure that the final export wouldn't be too large for processing and to execute the queries as quickly as possible, the amount of characteristics listed in the export were limited to the essential information that would allow classification as well as identification of the cases such as the unique identifier, the plaintext report itself as well as a short description.

The keywords used for the query were developed in two ways, first a visit was payed to a regional police force to discuss various DDoS cases and how they were investigated. During this visit various ways of writing a DDoS report were discussed and some of the Dutch terms were identified. Next, various types of popular synonyms such as ”DDoS” and ”Denial of Service” (it's important to note that the system is case-insensitive and thus no additional keywords had to be developed in order to catch the different kinds of capitalisation) were collected. Utilising these two sources in addition to keywords identified in a number of test searches, which mostly yielded additional synonyms, such as Dutch translations of the effects of an attack. These seem to be especially popular as the police officers recording the report have an easier time naming the effects, as opposed to the name of the attack type itself. The final collection of keywords used to get the data-set analysed throughout this paper reads as follows: dosaanvallen, DDoS, dos, dos, Denial of service, flooding, flood, booted, booter, stressed, stresser, amplification, platleggen, netwerkaanval. The final result being an overview of cases that matched the various criteria. Naturally many of these cases were false positives due to one of the keywords being mentioned in a non-DDoS related case. A big offender in that regard was ”DoS” as it yielded cases that would mention a ”DOS prompt” referring to the Microsoft command prompt. To filter the mentioned cases and to delete them from the final set as well as to categorise the cases in order to use them in the analysis in chapter 4, all case files were read and sorted in excel. For more information on the analysis that was executed while reading through the files and categorising them, refer to the data analysis section below. After all data cleaning steps were executed, 209 cases related to DDoS attacks were found. These 209 cases build the basis used for the descriptive findings section of chapter 4.

For the categorisation section of the same chapter, which will be used as a basis for comparison with the AmpPot dataset, more filtering needs to be done in order to make the police data-set a valid counterpart. The first step is to delete all cases in which organisations or people were threatened or extorted with the possibility of a DDoS attack, but weren't followed up by an actual attack. While especially the extortion cases are an interesting category of cyber crime in and of itself, they may only be used for the descriptive section as the cases would skew the statistics without actually being a real attack. This filtered the cases down to a number of 209 cases. Lastly, the cases were filtered by time of

occurrence and limited to between 1/1/2014 to 31/12/2015 in order to ensure they would overlap with the attacks recorded by the AmpPot data-set. This final filter left 144 cases for the categorisation.

3.3. Data Analysis

The analysis of the police reports has a number of issues connected to it, one being the layout it is presented in, or rather the lack there of. The reports themselves are saved as plain text files and rather than following a certain question and answer like structure that would be susceptible for querying or any other automated read-out procedure, are a simple storyline of what the victim has to say. As such, save for IP addresses, all 209 cases had to be read in full, in order to gather information utilised in the analysis. Furthermore, the greatly varying quality of the reports meant that while some cases provided everything from which IP address was attacked to all kinds of attack vectors such as the number of attacking IP addresses, a copy of the network logs and the packet type used for the attack; others only mentioned that the internet was down due to a DDoS attack. Likewise, some of the cases contain many so called mutations, which are additional pieces of information such as interrogations or the analysis of networks logs, hence providing much more data to work with. Due to these difficulties, all files were read manually and notes were made in an excel file that would later be translated into an extensive database noting the particulars of the case such as the victim type (on two levels, one board general one such as educational or company and a second more specific one denoting the type of educational institution e.g. high school or university), whether there is a suspect or not, particulars of the case such as did the reporter know what kind of attack it was and whether there was bullying or extortion involved. One may note the absence of whether somebody was found guilty by a judge or not. This is due the police system not being connected to the system of the Public Prosecution Service (”Openbaar Ministerie”) and hence only includes what the police collects in terms of evidence. For more information on how the data was analysed and what kind of scripts were used to automate part of the analysis to speed up the process, refer to the section A.1 of the Appendix.

To execute the above stated analytical steps, the basic police report framework is used. Since police files are an official document with a set of requirements and rules, many pieces of information need to be present by law and as such are available in all reports made. In the case of Dutch police reports, most of these rules are set in the criminal law (Wetboek van Strafvordering) and more specifically in article 163. While there are many nuances to the requirements, which are best read in the law itself, for this research there are essentially two major points of importance: 1. The victim definition 2. The article that makes the reported action a criminal act. The victim definition is usually stated in the first paragraph of a report. In the definition three major pieces of information are brought forward; the declarant - which is a description of the person making the report, the victim - a description of the victim which may or may not be the same as reporter, a declaration of the declarant stating that they are legally allowed to report the incident in the name of the victim. There is no fixed location to quote the criminal law article that makes the actions reported illegal, nor does the whole article need to be quoted. It is however the task of the police officer to write the report in such a way that the various properties noted in the particular article of the criminal law come forth in the report [2]. For DDoS cases both article 138b and 350a of the Dutch book of criminal law (Wetboek van Strafrecht) are of relevance. As such it is the task of the police officer to note as many details proving the relation between the case brought forth and both articles. Finally, much like the police report interrogations use a set of basic rules that prove valuable for their use in this thesis. In essence all interviews follow the same procedure containing two main parts: the social interview; containing basic personal information such as the name, age and current address as well the current emotional state of the person, the case specific interview; containing questions specific to the offence such as how the DDoS attack was executed and how the attacker got the idea etc. [2]. If the interviewee is a minor, they may have a guardian present during the interview, at this point the guardian may give some additional insight on the suspect in order to put their answers into context. The layout of the interviews follows a basic (Q)uestion and (A)nswer process [2].

3.4. Data Limitations

As mentioned in the previous sections, there are a number of issues regarding the police data-set. It is important to understand the limitations originating from these issues as well as the data retrieval steps, as they shape the glasses through which this research can be read and its results valued. Limitations

can roughly be categorised into three major categories: Data limitations, data retrieval limitations and data analysis limitations. Data limitations are issues as mentioned previous sections that directly impact the amount of information that can be gathered from the police case files. The main issue as mentioned is the differing content of the files as well as the detail and differing ways of writing. The major limitation arising from this issue is the level of interpretation on the side of the researcher that needs to be executed in order sort information written out in the case files. One may argue that this step is subjective and as such limits the validity of the data-set. While this may be true, and many of the reports were written in a different style, the major characteristics of a report are written in a constant way and as such the subjectiveness was limited as much as possible. Data retrieval limitations are issues associated with querying data from the system, or in other words distilling the DDoS cases from all other case files. The major issue here is that, as there are no DDoS classifiers other than searching through the entire plain text, cases may be lacking from the overall data-set if they use an obscure or unknown way of describing a DDoS attack. Due to this issue, it is possible that the data-set gathered via the retrieval procedure detailed in section 3.2 is not a complete representation of all cases that were filed in the period. To counteract this limitation as much as possible an extensive procedure was set-up including various tests. This procedure is available for scrutiny in section 3.2. Lastly, data analysis limitations are about issues limiting the analysis of the data itself. The major limitation arising from this type is that of differing amount of detail given by the reports, and as such not all analysis steps could be executed for the entire data-set. The downside of having such a large amount of missing values is the potential of miss-representing the data-set. To limit this as much as possible, all analysis steps that do not contain the entire data-set mention so and as such warn the reader of this limitation.

3.5. Conclusion

As noted in the introductory chapter of this research, due to the inherent properties of the police administrational system, a data-set containing all DDoS cases investigated by the Dutch police isn't readily available. To ensure the validity of the data-set and by that, that of the research the first sub-question posed reads as follows:

SQ1: How can cases files related to DDoS attacks be queried from the police administrative system?

Throughout the previous paragraphs this question was answered using a number of methodologies, providing the following results. The BVH system as deployed by the Dutch police was developed in a day-and-age where cyber crime didn't play major role yet. As such the categorical possibilities are limited and currently only one cyber crime category exists. As some regional offices have tried to utilise other categories, not originally meant for cyber crime, it's vital to keep these in mind in the querying process. This type of behaviour was also present for DDoS cases resulting in a large amount of cases being filed under the fraud category. To further limited the search query the keyword search option is utilised. This option allows the analyst to search straight through the plain text case files. In order to get a valid matches a list of keywords was developed by contacting investigators at regional offices as well by trial and error searches, yielding the following list: dosaanvallen, DDoS, dos, dos, Denial of service, flooding, flood, booted, booter, stressed, stresser, amplification, platleggen, netwerkaanval. Since the files themselves contain no inherent structure or layout, most analysis had to be done by hand by reading the entire case file. The data is limited by a number of factors; firstly, the data retrieval methodology isn't an absolute way of finding cases, as due to the way the keyword development was done more keywords may be needed. Due to the inherent structure of the system however, there this was the most efficient and practical way to solve the issue within the bounds of this thesis. Secondly, the information noted throughout the files are limited and vary widely in terms of quantity and quality.

4

The Police and their Data

As this thesis consists of three separate findings sections, this chapter constitutes as the first of the three, focusing entirely on the police data-set. As such the following sections will detail the DDoS attacks as they are experienced by the Dutch police through reports made by victims. The purpose of this section is twofold, the first is to give the police a taxonomy of their own workload as well, the second is to build a basis to which other data-sets can be compared. Due to the inherent properties of the data and how it is stored and extensive data retrieval and data analysis procedure had to be developed and executed. For more information on how this process looked like and what the limitations are, please refer to chapter 3.

4.1. Descriptives

As mentioned in chapter 3, 209 cases were queried from the police report system of which 144 fall into the period of 2014 to 2015. Of those, 58 cases regarded attacks occurring in 2014 while 86 regarded attacks executed in 2015. This indicates an increase of 48% over a one year period. The magnitude of the increase may be due to a variety of reasons such as an overall increase of DDoS attacks, a reason for this increase may be the general trend of DDoS attacks becoming a commodity and thus are executed by more and more people [46]. Additionally, the increased media attention due to large scale attacks such as the Ziggo DDoS of august 2015 [55] add to the common knowledge and hence make people more aware of a DDoS happening. Looking at the official reports themselves, it becomes apparent that the wealth of information differs significantly between the various cases files. Like mentioned in chapter 3, this is in part due to the lack of specialised knowledge on the side of the police officer making the report, but also signifies the relatively low knowledge level of the victims themselves. One example of the knowledge level is in the details victims mention during their first contact with the police. As such, in only 10% of the cases did the victims note detailed information of the attack such as what kind of DDoS attack it was or what the magnitude of the attack, either in terms of traffic or the number of attacking IP addresses, was. 34 cases mention the IP address of the victim, while 21 cases mention a suspect and perform a interrogation based on this. The latter is limited to cases in the educational sector.

Table 4.1: Summary of cases gathered from police system

Queried for analysis	209
2 year period 2014-2015	144
2014	58
2015	86
Victim IP is mentioned	34
Suspect is mentioned	27
Attack vectors mentioned	21

In his paper Jose Nazario [37] noted 5 types of- or motivations for attacks: Home user attacks to nag

or anger somebody, retaliation attacks on anti-spam/anti DDoS instances, extortion attacks, attacks on internet infrastructure such as DNS servers, and politically motivated attacks. In his eyes these types can be deduced by looking at the victims that were attacked as well as the strength of the attack. Utilising the inherent structure of the police reports as described in chapter 3, one can execute the very same methodology as used by Nazario to check whether these types are present in the Dutch police data-set as well. In addition to the information used by Nazario, the police case files yield both the personal insight of the victim as well as 21 interviews with the attackers themselves, hence allowing for a more detailed approach. Analysing the case files in said way, shows that two of Nazario's categories don't show up in the police data at all, namely the attacks on anti-spam/anti DDoS instances and attacks on internet infrastructure such as DNS servers. Politically motivated attacks do in fact occur, however they are very limited in their quantity and only represent relatively ineffective attacks. Both the home user attacks to nag or anger somebody as well as the extortion attacks are very much present in the attacks. Generalising the attacks much like Nazario, the categories would take the following shape:

High School curiosity

Many of the victims are related to high schools, either by the victim literally being a school, or the victim being a student in high school age. This falls in-line with comments made by both victims and suspects alike, whom mention that using booter services to ”turn off” somebody's internet connection for a short period of time is completely normal between friends, as a way of nagging and to show somebody's ability. In other cases, especially regarding schools, it's often curiosity that takes over and leads the young attackers to execute an attack without knowing what the eventual repercussions would be. Import to note is the intent underlying the attack, which in many cases either doesn't exist or in the worst case pertains to the social status of the attack themselves.

Gamers and Bullying

Like many DDoS reports all over the internet, from credible academic sources to personal blog posts mention, DDoS attacks are a tool that is often used in online gaming circles to gain an advantage over another user. This may be by taking up just enough bandwidth to slow the opponent down, or by simply disrupting their internet connection all together. Reading through the case files it becomes apparent that quite a few of these cases aren't as clear cut as gamer 1 versus gamer 2, but often come coupled with a certain bullying nature. This becomes especially apparent when the attacker and victim are still minors, which in most cases they are. This is not to say that bullying cannot occur on its own, in fact seven cases were identified in which minors in their mid teens were subject to bullying via Skype conversations which were followed up by DDoS attacks.

Extortion cases

Throughout the data-set 6 extortion cases were identified, only one of them was actually executed. Extortion cases have a very different character than most other cases mentioned here, simply because of their inherent financially driven intentions. Generally speaking DDoS extortion cases follow one of three schemes: 1. Pay or else; which consists of a simple email stating the intention of the attacker to DDoS the victim unless they pay a certain ”safe fee”, 2. Attack-Pay-attack; In this scheme the attacker shows their seriousness by executing a sample attack, then sending the victim the extortion message noting that they will attack again unless the victim pays the ”safe fee”, 3. Attack until payed; in this scheme the attacker starts their DDoS and doesn't stop until the victim pays the ”safe fee”. It's important to note that only one of those six cases led to an attack, not because the victims payed but simply because many of these are trying to use the name of notorious groups such as Armada Collective [42] whom have gotten extensive media coverage [9] to make victims pay.

4.2. Attack Classification

While the previous section focused on analysing the cases on a descriptive level, this section focuses on the classification of the cases. As the results of this section will be used as a basis for comparison to secondary external sources, which have been recorded over the two year period 2015-2016. This section will do likewise by only selecting the sub-set of cases that note a date of the crime between the 1/1/2015 and 31/12/2016 hence resulting in the aforementioned 144 case sub-set.

To categorise the cases by the type of victim attacked, two sets of categorises were created. The first level is a fairly broad distribution of the four main types of organisations found in the data-set: Company, Educational Institutions, Home broadband and Governmental institution. The second level goes one step further and makes use of the detail provided by the police reports to further categorise the victims. For the category company, the same sub-categories as defined by Verisign in their 2015 industry report [56] will be used with the addition of a other category for cases that fall outside of the prescribed categories which read as follows: IT Services/Cloud/SaaS, Media & Entertainment/Content, Financial, Public Sector, Telecom, E-Commerce/ Online Advertising. The Educational sub-categories will be based on the various levels of the Dutch education system as noted by the Dutch Ministry of Education, Culture and Science [51] and are defined as: PO = Primary and special education, VO = Secondary education, MBO = Secondary vocational education, BOL = Vocational training pathway in MBO, BVE = Adult and vocational education, HBO = Higher professional education, WO = Universities. There are no official classifications for the Home Broadband category. As such the categories defined were developed after reading all of the cases and deducing major groups that are also indicated in literature such as teen age youth [45] and gamers [18]. All other attacks on home connections that do not resemble any of the two will be under the other category. Cases in this category are all average home users, or people that work from home. Lastly, looking at the cases categorised under the main governmental category proves that all victims are ministries and as such cannot be sub-categorised.

In order to apply any of the categories stated above to the police reports themselves, one of the inherent properties present in any police report is used, the victim definition. This definition is present in every police report and is usually stated in the first paragraph of a report. In the definition three major pieces of information are brought forward; the declarant - which is a description of the person making the report, the victim - a description of the victim which may or may not be the same as reporter, a declaration of the declarant stating that they are legally allowed to report the incident in the name of the victim (article 162 - Wetboek van Strafvordering). Using this paragraph, one can clearly identify the various types of victims and place them into the categories. For a more extensive explanation of the police reports and how these are created and what if information is noted in them, please refer to chapter 3.

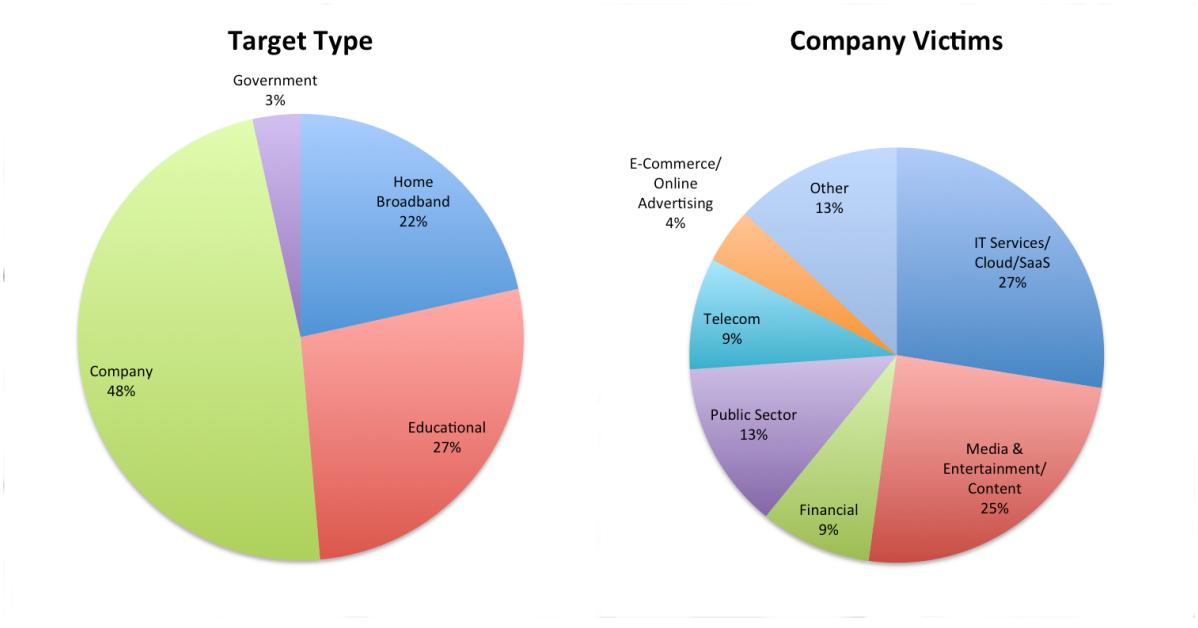

Figure 4.1: Victim demographics of police reports & Company victims by type of business

The first level (as depicted in figure 4.1) splits the cases into the previously defined major groups: companies, governments, educational institutions and private broadband connections. Companies account for nearly half of all police reports, followed by 27% educational institutions and 22% broadband home connections, trailed by a relatively small amount of governmental targets. Zooming in on the

relatively large company group, accounting for 67 reports in absolute terms, shows that the IT services and Media & Entertainment categories account for half of the attacks. It is important to note that the Media & Entertainment category, with its 25% also includes gaming related companies, which are so often connected to the DDoS phenomena. Additionally, it's notable that both the financial industry as well as the telecom businesses (which include ISPs) are victim to the same amount of attacks. Furthermore, the public sector e.g. health care and infrastructural services are the target of 13% of the reports, hence more than both telecom and the financial industry. Lastly, the other category accounts for 13% of cases. The "other" category contains a diverse set of business, everything from travel agencies to private contractors, thus showing that all kinds of organisations may become the victim of a DDoS attack.

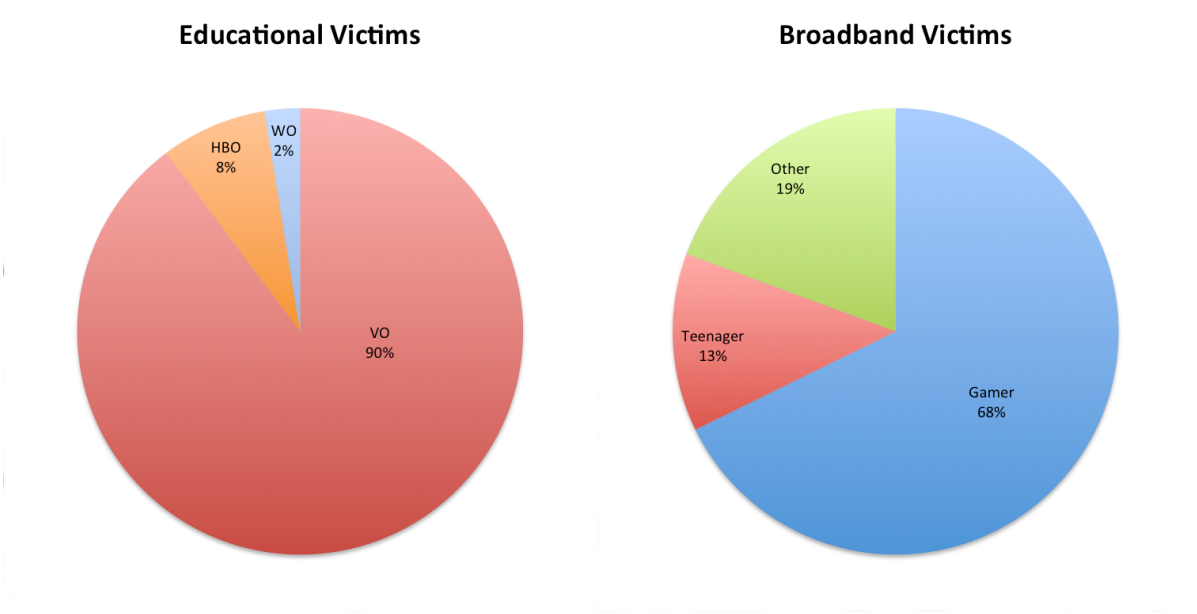

Figure 4.2: Victim demographics Educational sector & Broadband Victims

Zooming in on the second largest group, the educational victims, the distribution is much more clear cut. As visible in figure 4.3most of the categories defined in the earlier paragraphs don't show up as no reports were filed in their name. High schools account for 90% of all cases hence representing the vast majority. Other levels of educational institutions jointly account for 10%. Lastly, moving the spotlight to the third largest category, the broadband victims exhibited on the right in figure 4.3, a clear majority of 68% in the gaming category may be observed. The teenager category, which accounts for 13% of the broadband victims, exemplifies teenagers being attacked as a way of bullying or extensive nagging. 19% of the cases fall under the other category, which are either people whom work from home or are so-called "youtubers". Lastly, the government main-category accounting for 3% of all reports, or 5 reports in absolute terms, consists of various ministries as well as one report by the police themselves.

Suspects supplied

Information, or in law-enforcement terms 'investigative indicators', are central to any case the police investigates. These indicators more often then not must originate from the original police report, or in other words the first point of contact between the police and the victim. Analysing the DDoS case files proves this all the more as very little cases make any progress beyond taking in the initial report. The educational victims category however, falls outside of this general trend. Not only do educational victims supply a plethora of information in their reports, they actually go as far as supplying a suspect as well as a evidence backing-up their suspicion. As depicted in figure 4.3 two-thirds of all reports made by educational victims contain the name of a suspect as well as the aforementioned evidence.

Zooming-in even further, the high school sub-group further increases this value to 80% of all cases. There are multiple reasons for this such as the way the network is set-up and the unique situation they

Table 4.2: Summary of cases wit suspects supplied

Type of cases	Number of cases
Suspect is mentioned	27
EDU victims with a suspect	27
High School victims with a suspect	26
Suspect interrogations	21

are in. For more information on this please refer back to sub-section 4.1. It goes without saying that these cases also yield the most follow-up actions such as interrogations and finally provide the most solved cases.

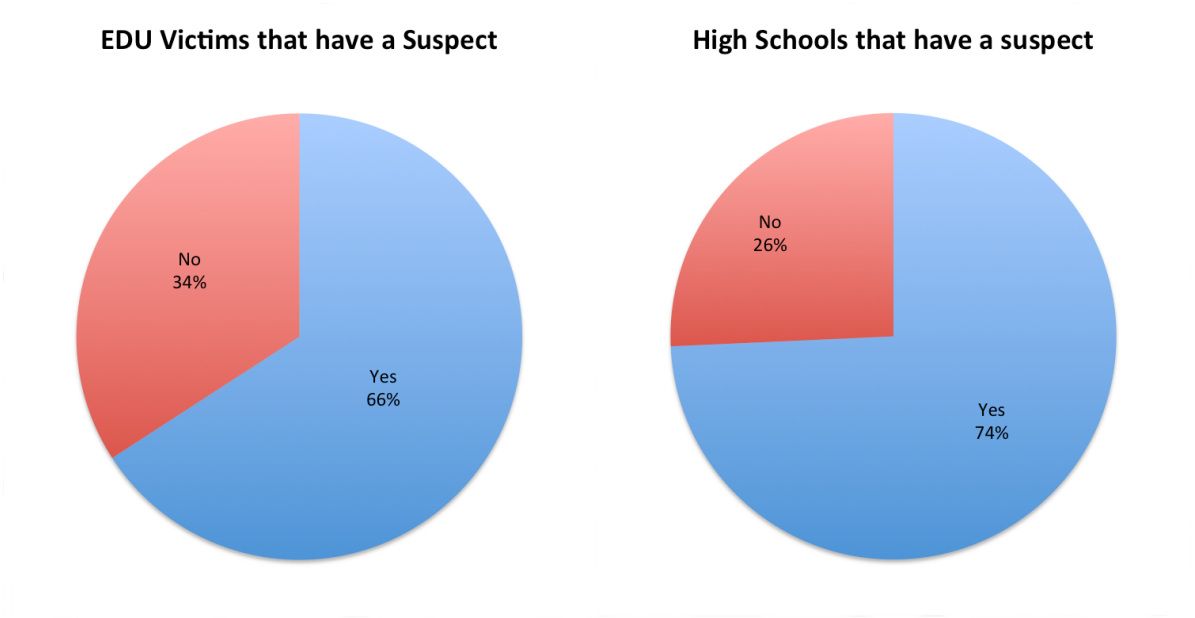

Figure 4.3: EDU victims that had a suspect when making a report

4.3. Interview analysis

The police reports provide a unique opportunity to not only view the victim side of an attack but, if a suspect was identified, also provide the personal view of the attacker in the form of the interrogations that were executed. As such this section will focus on 21 interrogations executed during various DDoS investigations executed by the Dutch police. It's important to note that these interrogations are excerpts from cases files as presented by the police and as such contain no court ruling on whether the suspect is guilty or not. In other words, the analysis below is not on 21 convicted attackers but suspects that may still try to talk their way out of potential punishment. To execute this analysis, all cases that contained a suspect and therefore held an interrogation with said person were selected. This search yielded 21 cases, all part of the educational victim category. Of these 21 cases, 20 were targeting high schools while 1 was targeting a HBO (Higher professional education). To analyse the data two methods were used: 1. A three count of questions were used to give a deeper insight in the underlying knowledge attackers have 2. A descriptive analysis was executed to gather more general information on the attackers and how they got the idea to execute a DDoS attack. Since the interviews were already held, new questions could not be introduced. However, since police interrogations consist of the same steps and most questions are the same, the same amount of information can be extracted from all of them. In essence all interviews follow the same procedure containing two main parts: the social interview; containing basic personal information such as the name, age and current address as well the current emotional state of the person, the case specific interview; containing questions specific to the offence such as how the DDoS attack was executed and how the attacker got the idea etc. [2]. While all interviewees have the right to a lawyer and to have the lawyer present during the interviews, only 2 had

contact with a lawyer before the interview and none had one present at the time. If the interviewee is a minor, they may have a guardian present during the interview, at this point the guardian may give some additional insight on the suspect in order to put their answers into context. The layout of the interviews follows a basic (Q)uestion and (A)nswer process.

Cyber criminal or script kiddie

While cyber crime is often referenced as one of the most complex and difficult to understand types of criminality, it is the purpose of this section to get a better understanding of the technical capability of the attackers. Are they aware of the underlying actions that are being executed when attacking a victim and as such, do they also have the capability to estimate what the damage may be when executing an attack. Inline with the ability to estimate the damage comes the question on whether the attackers know whether their actions are illegal in the eye of the Dutch law. As such the answers to the following three questions were identified in the interviews: 1. Does the suspect know what a DDoS attack is? 2. Does the suspect know that a DDoS attack is against the law? 3. Does the attacker know how to execute a DDoS attack? All three of these questions were asked in all of the interviews and as such the answers could be directly used for this analysis. Figure 4.4 depicted below indicates the answers of the mentioned questions on a percentage scale.

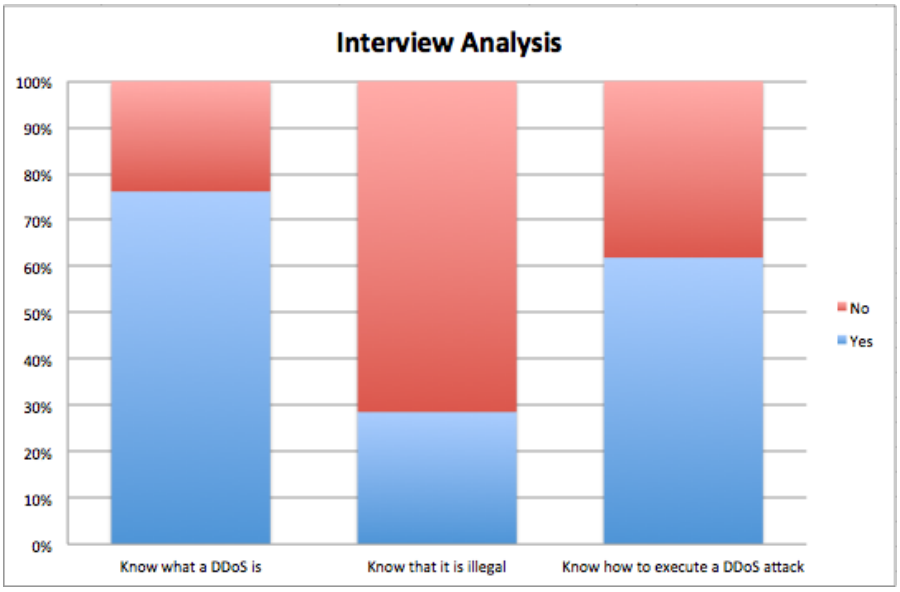

Figure 4.4: Interview Analysis

As indicated 80% of the attackers understand what a DDoS attack entails while only 29% realise that their actions are against the law. The last question may be the most interesting one, only 71% know how to execute an attack. One may argue, how can one execute a DDoS without knowing how to execute one let alone not knowing what a DDoS attack actually entails. This is where the interviews show their additional value, the attackers mentioned the ease with which the attack, often spontaneously, are executed. They mention links of websites that ”turn off the schools internet” for a certain period of time. In more specific words, the attackers are usually young males between the age of 14 and 16 who receive the link to booter services, often via some kind of message client such as skype or whatsapp or from web-forums they visit. The booter services are simplified in such a way that all the user has to do is to press a button on a website and the attack is executed on the IP address of the requester. Conversely, this means that all the attacker needs to do is connect to the wifi of a certain organisation, go to the website and click on attack. Hence, the attacker doesn’t even need to have the knowledge to find the public IP address of their target.

Descriptive

In addition to the answers to the three questions that were posed above, a number of descriptive pieces of information were found while analysing the interrogations. As such, all of the interviewees

Table 4.3: Attack motivation

Motivation	# of cases
Home situation	2
No reason, just for fun	8
Testing	2
Pressured into it	2
Curiosity	2

were males between 14 and 16 years of age studying at either the ”preparatory middle-level vocational education” (VMBO) or ”preparatory scholarly education” (VWO). Furthermore, the attack methodology often originates from how-to guides on websites or links that were sent via group chats. Many of the attackers mentioned that DDoS attacks are common place between friends as a way of nagging each other and that ”everybody knows how to do it” referring to the usage of a booter service. A few also mention the fact of being outsiders at school played into their decision to execute an attack, and that they tried to use their knowledge about DDoS attacks to better their social circumstances. Figure 4.3 depicts an overview of the various motivations noted in the interrogations, to the left are the number of time each motivation was mentioned. Most prominent is the ”Just for fun/No reason” motivation while bad home situations as well as testing and simple curiosity all account for 2 cases. One of the more worrying outcomes is that two further cases noted that the attacker was being pressured into the attack. Curiosity seems often to be sparked by forum threads or big chatrooms mentioning a ”website that can turn of the internet” leading the relatively young actors to press a button without understanding what would happen. Repercussions and effects seem to be completely unknown as almost all of the interviewees note that they had no idea what the possible effect of such an attack could be. Cases in which the attack believes that they are attacking one single computer within a school but rather are taking down the network in an entire school community.

In chapter 1 and 2 the phenomenon of the commoditisation of cyber crime was introduced. The idea being that executing a certain type of cyber crime is so easy that almost everybody can do it. Much of this phenomenon is supported by the servitisation of the various types of cyber attacks. DDoS attacks may be a prime example, with the vast numbers of booter services available on the internet, operating in plain sight for everybody to see on the clear web. Within the police case files there was also one interview, in addition to the 16 mentioned above, which indicates that the DDoS attack type takes the commoditisation one step further, the commoditisation of facilitator role. The person in question, still a minor in high school, noted that he had been running his booter services for 5 months until he shut it down. He argued that, due to the large amount of booter offerings online, he only earned approximately 200 euro over the entire period. In regard to the commoditisation of the facilitator role, he explained that all he had to do to run his own booter site was to download the source files of the site from a forum he found online, change a few variables in order to direct payments to his own Paypal account and he was ready to go. Therefore, all that is needed to set up a booter service is some basic IT knowledge, making not only executing a attack, but also profiting of attacks available to almost everybody.

4.4. Conclusion

Throughout this chapter the DDoS phenomenon from the point of view of the Dutch national police was analysed. To do so the methodology developed in chapter 3 was executed and the resulting cases were analysed. The goal of this analysis was to answer both the second and third sub-question as developed in the first chapter.

SQ2: How can these case files be analysed to categorise them?

Due to the limited amount of consistently available information, it is difficult to find multiple characterising attributes that could used for categorisation. One that is however both interesting and present across all cases is that of the target type. The data shows that the vast majority of attacks target companies, what kind of company specifically varies per case and while IT related categories add up to about 37% of the company related cases there aren’t any significant results to be found. The second and third largest group are the educational sector and private broadband connections respectively. Within

the educational sector, the mammoth share is held by high schools while the broadband connections are mostly made up of gaming related targets.

SQ3: How are the DDoS attacks reported to the Dutch police characterised on a descriptive level and what can be said about the identified attackers?

On a descriptive level the case files yield interesting insight the type of attack executed. Generally speaking three major types could be identified; High school curiosity, which are attacks executed by students who do not have malicious intend but are rather interested in the possible effects of testing the network or are simply bored and were playing around with a booter service. Gamers and bullying are another returning theme, where games DDoS each other to get a competitive advantage or sometimes simply to nag each other. A more serious derivative of this type are bullying cases, in which often high school age teenagers, are bullied and sometimes even pressured into passing on passwords or other information. Lastly there are the extortion cases which are in terms of numbers, especially in regard to executed attacks, extremely limited. Many of the threats are made in the name of large criminal groups yet don't back their threats up with actual attacks. The malicious intend of these threats is rather clear and certainly represent one of the darkest sides of the DDoS phenomenon. While not all cases fall into these classifications, they are interesting nonetheless and show the most prominent overall themes. Taking a better look at the attackers whom executed the attacks and were interviewed by police afterwards, one finds a group of exclusively high school age male teenagers. The interviews proofed that only three-fourths of the attackers knew the meaning of the term DDoS while even less seem to understand how to actually execute one. One may conclude that while many of them are playing around with booter services they don't understand what the purpose of tool nor what the entailing repercussions would be. When asked whether they knew that utilising these tools is illegal, only a small minority of the attackers answered with yes. Aggregating all this information clarifies just how normal it is, for especially young teenagers, to utilise booter services. They often don't know what they are doing other than that it turns of the internet for a short period of time. It may also be concluded that many of them simply follow quick tutorials on "how to turn of a friends internet" found in online chatrooms or forums and as such have understanding of what they are doing. An interrogation with a teenager hosting his own booter service indicates that within the DDoS world, the commoditisation goes one step further, making hosting a booter service easy enough almost everybody can do it.

5

The Police vs. The World

5.1. Intro

Throughout the fifth chapter the DDoS population in Netherlands as a whole will be inspected. This is to give a deeper insight in what actually happens and what the attack vectors actually are. This will provide a counterpoint to which data found in the police reports can be compared to, to give the Dutch police a better idea of how their view of the DDoS issue differs from what really happens. To do so, two new data-sets will be introduced. The first one originates from the Dutch scrubbing agency who provides anti-DDoS services to its Dutch member organisations. Due to its inherent properties, it only contains data of attacks whose victims are located in the Netherlands and as such allows us to characterise the state of affairs specifically for the Netherlands. The second step is to compare the police data to the DDoS population as a whole. For this step the AmpPot [28] data will be utilised as well as previously established results from [39]. The third and last step will zoom-in even further by combining the AmpPot data with the police data, hence finding the attacks noted in the police reports, in the AmpPot data to get a better understanding of how the reports translate to the real attack vectors.

5.2. Amplifiers in the Netherlands

As indicated by the literature review executed in chapter 2.2, there are many different types of DDoS attacks. Each of these attack types makes use of a different methodology to eventually reach the final goal, deny service to legitimate users. In the past, botnets were often used for these kinds of attacks. The methodology is rather simple, the attacker makes connection to their command and control server, which in turn connects to the bots whom attack the victim directly [12]. All the attacker has to do, give the command to let the bots connect to a certain website. In recent years however, this and many other traditional type of attacks have proven inefficient and not powerful enough. To remedy this issue, amplification and reflection was introduced. In a nutshell the attack stays the same but rather than letting the bots directly attack the victim, the bots make connection to another service on the internet, sending a simple query which results in a much bigger answer, hence amplifying the amount pf packets sent. This amplified attack is then directed to the victim [28]. DDoS mitigation companies such as Akamai argue that the vast majority of cases world wide are of this type [1][27]. However, this thesis is very much focused on the Netherlands, and as none of the aforementioned companies have a data-set specific to the Netherlands a additional data-set needs to be introduced. As such the Dutch scrubbing centre, tasked with the mitigation of DDoS attacks specific to the Netherlands, has a much better view on the environment. The data provided by the scrubber shows that almost 99% of all attacks observed make use of amplification or reflection methodologies. Thus, one may conclude that amplification and reflection attacks account for the vast majority of all DDoS attacks targeting Dutch victims.

5.3. AmpPot and the Netherlands

To get a better understanding of how the DDoS environment in the Netherlands looks like, this section will take a closer look at the AmpPot data-set as recorded over the period form 2014 to 2015. Before delving into the results it's important to note that the data-set is a record of all attacks the honeypots

were supposed to be a part of. Due to this, the data-set is not a record of all attacks but rather of targeted IP addresses that were sent to various amplifiers. Hence, multiple records may actually be part of the same attack. To consolidate the data-set to represent attacks, the following definition was handled for all analysis steps in this papers: all unique combinations of the target IP, attack protocol e.g. DNS, NTP, etc. and start date and hour of an attack. Overall this approach yields 53055 unique attacks over the two year period. The figure below is a graphical representation of the attacks over the period.

Table 5.1: Summary of cases gathered from the AmpPot data-set

Type of cases	Number of cases
Number of attacks	53055
Unique IPs attacked	22580

To create the various graphs in the upcoming paragraphs, the attack definition as described in the previous paragraph was utilised. The attacks were then plotted according to the date of attack. To graph the protocol distribution, the same methodology was used, however the attacks were first split by attack protocol. To graph the attack duration, a sort like procedure was used, through the attacks were sorted by attack duration instead. To create the graphs python in combination with the pandas add on was used. For more information regarding the code used to create the graphs see appendix A.5.

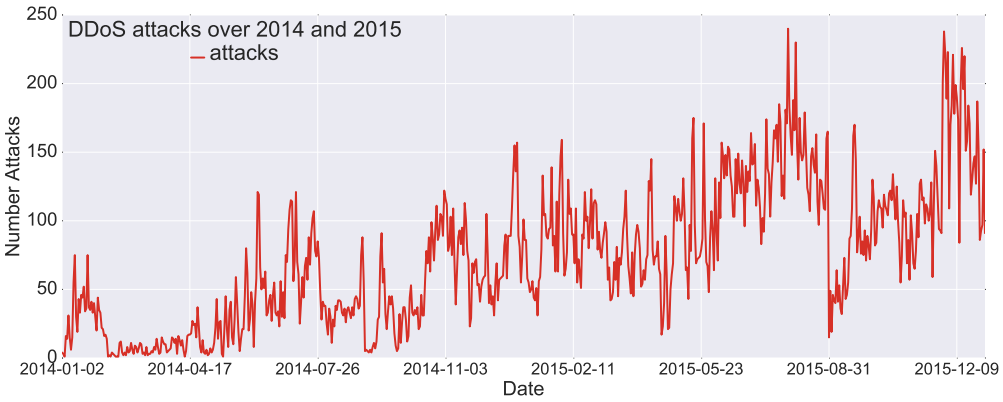

Figure 5.1: DDoS attack volume over the years 2014 & 2015 according to the AmpPot data-set

While the curve shows a clear upward trend, it's important to note that the effect cannot in its entirety be assigned to a higher overall volume of DDoS attacks. Since the data-set uses a set of eight honeypots that were activated in the beginning of 2014 [28] a certain discovery effect, in other words over time more and more web-crawlers used by booter services find the the honeypots, is to be expected.

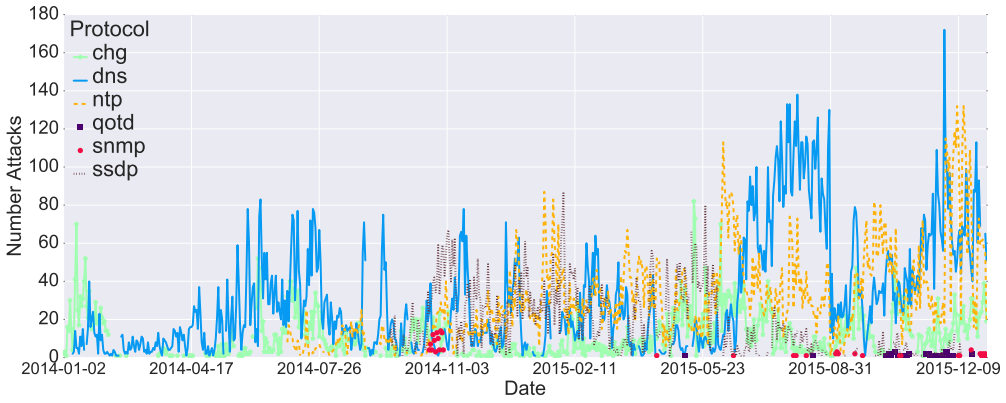

Figure 5.2: Protocol distribution of attacks on Dutch victims over 2014 & 2015

Since the honeypots offer a variety of protocols for amplification it's interesting to see what the preferred protocols are. Figure 5.2 depicts this distribution as recorded over time. On an an aggregate level the attacks were distributed: DNS = 40.33%, NTP = 30.89%, CharGen = 14.32%, SSDP = 14.12%, SNMP = 0,24% and QoTD = 0.07%. These results are comparable with other studies [39], the relative popularity of DNS and NTP is mostly likely due to them being essential open internet services that cannot simply be updated or turned off in order to limit the amount of usage. Additionally there are simply a lot more DNS servers than thee are QoTD servers. A further aspect the AmpPot data-set can shed some light on is the actual number of victims behind a certain DDoS attack. While the attack definition as described above identifies unique attacks on certain IP addresses, they do not identify what is behind that IP address. While in many cases these will simply be a home connection or one organisation, in the hosting environment these may very well be the IP address of a shared hosting server. Shared hosting is a cost efficient way of hosting, where rather than assigning an entire server to one customer, a server and with that the IP address is shared by multiple customers. In the event of a DDoS attack on one of those customers on the sever, all fellow users will feel the effects as well. On the level of the whole population, 22% of all attacks have more than 1 domain, meaning more than one victim to an attack. Looking at the traced attacks this is almost exactly the same with 21%. This is an important indicator for the police as due to the unique character of shared hosting, a police report may be filed by an entity that wasn't actually the original target. While there is northing inherently wrong with another party filing the report, investigative indicators only present at the intended target (e.g. threats, ransom, etc) may be lost.

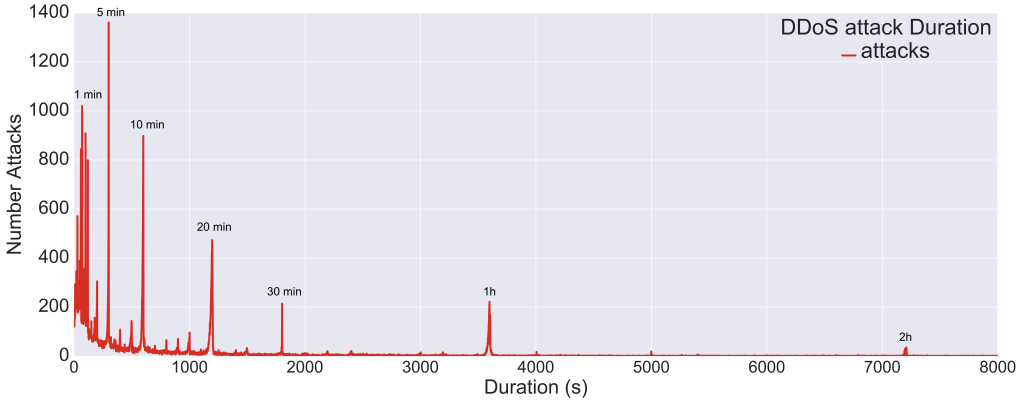

Figure 5.3: Attack durations

Lastly, the data-set gives a great insight in the duration of the various attacks. The duration is an important indicator of the amount of damage an attack causes since the longer an attack takes, the longer a service is unavailable or the longer an anti-DDoS solution needs to be utilised. As the graph

above indicates, there are very clear peaks in the duration at 1min, 5min, 10min, 20min, 30min, 1h and 2h. Since booters offer their services in the form of packages that usually use attack duration as a differentiator, these peaks make sense. Especially the cheaper 1, 5, and 10 min packages seem popular. To put this in perspective, the longest attack executed was 51,24h or just more than two days.

5.4. Police vs. AmpPot

Results presented in section 5.2 indicate that the vast majority of all DDoS attacks on Dutch victims are of the amplification variety. In the following two subsections the two data-sets, the police reports as well as the AmpPot data will be combined to get a better insight in the attacks themselves. The underlying question that will be answered through this section is how the police reports compare to the DDoS population as a whole and whether the police reports are a good representation of the overall population. Before going on to analyse the results, it's important to note the format in which the two datasets were supplied. The police data-set consists of reports, usually these reports contain one attack but may also contain multiple if they were executed in close succession of one-another and hence may be noted as one occurrence. The AmpPot data-set however, notes attack queries sent to various amplification servers utilising multiple protocols. One such example would be 5 different DNS servers being connected to at the same time in order to attack one victim. Due to this set-up as well as the distributed nature of DDoS attacks, the raw data contains multiple data point per attack. To make the data compatible, a attack was defined as all unique occurrences of an IP address being attacked via a certain protocol and that starts on a certain date and hour. For more information on the definition and how this filter was put on the data see appendix section A.4.

Furthermore, one of the main goals of the coming paragraphs is to compare the victim demographics as brought forth by the two data-sets. While the police data, as noted in the chapter 4, gives detailed information about the victim, the AmpPot data does not and a proxy needs to be used. The proxy used to categorise the attacks are the Autonomous Systems the IP addresses belong to. Autonomous Systems (AS) are "a set of routers under a single technical administration" [17] or domain that works under a single routing policy. In practice this means that large organisations such as universities or large multi-nationals as well as hosting parties have their own AS. Home broadband users are usually under one of the ASes managed by their ISP. Due to these properties, AS are a good tool to approximate the nature of the underlying IP addresses. It's important to note that Autonomous Systems are not bound by geographical boarders and as such an AS may be partially located in the Netherlands and partially in one or more other countries. This thesis only analyses data specific to the Netherlands, and as such only looks at the Dutch part of an AS. Lastly, as the AS classification was in large done by various actors at the TU Delft in order to create papers such as that of Noroozian et al., the classifications don't match classifiers as defined in chapter 4. More specifically the ASes are categorised in the following victim categories: Hosting, Edu, ISP-broadband, ISP-mobile, ISP-other and Non-intermediary. EDU and ISP broadband are the same as defined before, Hosting is part of the company category in the previous definition and non-intermediary proofs upon manual inspection to be all companies as well. ISP-mobile describes ASes containing mobile internet users and ISP-other are ASes belonging to major ISPs but that could not be categorised further. To make the comparison to the police data easier, the non-intermediary category is renamed to company and isp-mobile and isp-other are joined under the a additional other category. Since hosting is such a strong category it retains its own category.

Compare results

While law enforcement has a variety of goals, one of the most prominent ones is the identification and assistance of victims to the various kinds of crime. As such it's of great importance to know whether the victims identified by the police represent the actual victim population and if not which groups of victims are left out. To find out just that figure 5.4 depicts the victim demographics on the basis of both data-sets. As visible the two figures differ significantly, most notably the educational category only accounts for 1% of all DDoS attacks according to the AmpPot data while it accounts for 27% of all police reports. Equally large is the difference between the amount of targeted broadband connections, the AmpPot data-set argues that these account for 44% of all attacks while they only account for 22% of the police reports. The discrepancy between the two data-sets in regard to attacks on companies is rather small however, in the AmpPot data-set these are represented by not only the company category

but also the Hosting category, hence accounting for 51% of all attacks while they account for 47% the police reports. Note, that hosting companies are the largest sub-set of these companies, accounting for 47% of the attacks according to the AmpPot data-set yet only 6% of all police reports.

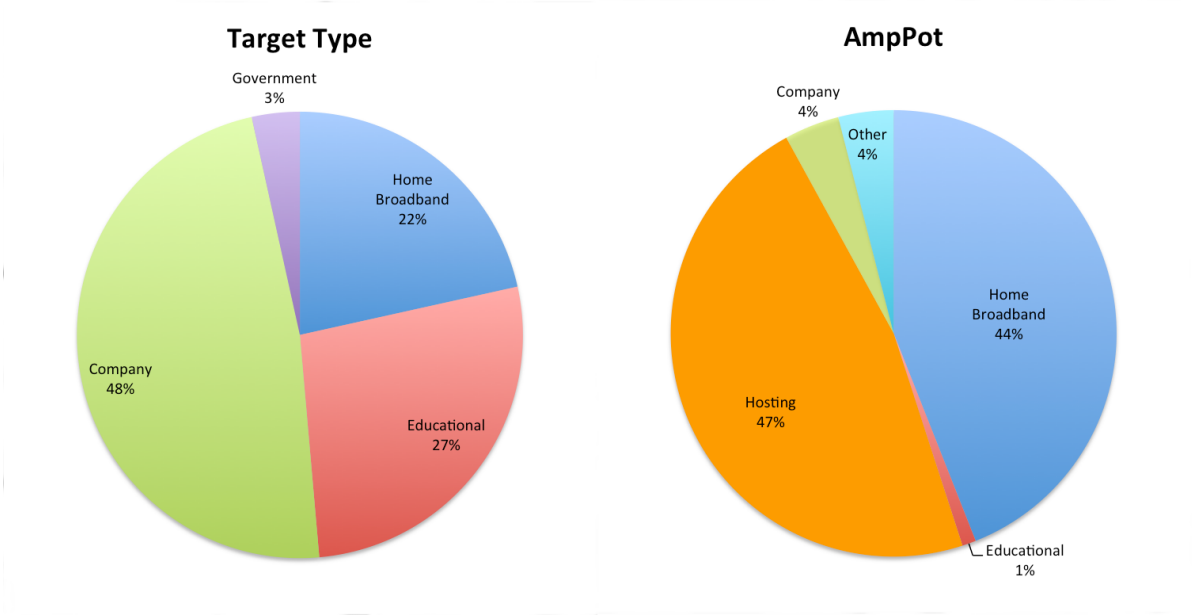

Figure 5.4: Victim breakdown Police vs. AmpPot

One may ask why the differences are so big and while a complete answer to this question may be a whole research topic on its own, the police reports and interrogations do note a number of explanations. Regarding the broadband victims, attackers note in their interrogations that DDoS attacks are part of their daily life, they argue that in high schools DDoS attacks are used for nagging and to be cool. Knowing this, it comes as no surprise that kids nagging each other won't lead to police reports, especially considering the relatively short and quick attacks that would be used for these kinds of attacks. Additionally, most of these attacks don't have tangible damages other than a short outage of a personal internet connection, not all that different from a normal service interruption. The educational reports are an interesting sub-group as opposed to other victims, they usually have a suspect already, as such they may be more prone to go to the police as they can supply sufficient information for an investigation. This directly leads to the company sub-group which accounts for the biggest part of the police reports. Looking at the reports these victims are the ones with the clearest idea of what the tangible damages are and as such may be more inclined to contact the police.

5.5. Conclusion

While the previous chapter focused on the data-set provided by the Dutch national police, this chapter moves on to first describe the Dutch DDoS environment as a whole, then introduces a data-set to be used as a sample of the population, and finally compares results found in the previous chapter with those found in the newly introduced data-set. Jointly these steps were used to answer the following research question:

SQ4: What does the current DDoS population in the Netherlands as a whole look like and how does it compare to the cases recorded by the police?

Throughout the literature study executed in chapter 2.2 DDoS through the amplification/reflection methodology has been identified as a rising and ever popular way of executing DDoS attacks. As proven by the data provided by the Dutch scrubbing agency, the Netherlands falls inline with these developments, in fact one may conclude that attacks in the Netherlands are almost exclusively of the amplification variety. Hence, one may conclude that the AmpPot data-set, recorded by eight amplification honeypot servers, is a valid representation of the Dutch DDoS environment. Utilising this data-set,

it can be concluded that there is an upward trend in the overall amount of attacks. In terms of protocols used for the amplification, both DNS and NTP are the most popular, jointly accounting for roughly 70% of all attacks. Looking at the attack duration, shows very clear peaks at 1min, 5min, 10min, 20min, 30min, 1h and 2h; indicating the usage of booter services whom use attack duration as a differentiator between the various subscriptions they offer. Comparing the victim demographics of the police system to that of the DDoS population as a whole shows that they differ significantly. Hosting seems to be targeted especially often, yet the number of police reports from this group is comparatively low. A possible explanation for this is the inherent nature of the hosting business and thus the presence of a good mitigation system that makes many of the attacks unsuccessful. Additionally, there may be the simple issue of having so many attacks that's impossible to go to the police to make a report for every single one. The second big difference is the educational sector, which are only target to roughly 1% of all attacks, yet represent more than one-fourth of all police reports. As found in the previous chapter, this difference may in large be due to the educational sector often knowing who the attacker is as opposed to most other victims. Lastly, the home-broadband connections represent almost half of the attacks, while only being 22% of the police reports. As the majority of these attacks are executed by teenagers playing video games and the like, it makes sense that very little would go to the police and make a formal report, especially since the monetary damages would be rather low too.

6

Tracing and analysing attacks

Through the previous two chapters, findings discovered in both the police data-set as well as the AmpPot data were discussed and victim demographics were compared. This chapter goes one step further by combining the two data-sets to get a full picture of an attack. Combining the two data-set allows the researcher to trace attacks from the booter server giving the order to attack, down to the victim with the damages incurred. Additionally, it provides an opportunity to use information extracted from one data-set to be directly compared to the other and to compare the traced attacks to the entire population.

6.1. Traced attacks

This subsection will zoom in one step further and analyse attacks that occur in both data-sets. To do so, the police case files will be analysed for IP addresses, the resulting addresses will in turn be used to search through the AmpPot data to match the corresponding attack. The last step is to check the date of the attack in both the police report and the AmpPot to ensure they are the same. The result of cross-referencing the two data-sets is a list of 32 cases containing 138 unique attacks that are present in both data-sets. Throughout this process, the biggest issue was the data quality as many of the police reports simply didn't note the victim's IP address. As such the value of 32 matches could be much higher and more valid statistics could be executed. Due to this limitation results presented in this section must be taken with a grain of salt. Nonetheless, they provide a first full picture of an attack from the amplifier down to the victim. Additionally, it allows the researcher to place the police report into context and providing information the victim usually wasn't able to.

Table 6.1: Summary of traced attacks data

Type of case	Number of cases
Victim IPs available	34
Traced cases	32
Unique attacks	138
Median attack length	5min

While, like mentioned previously, the dataset is rather small and executing extensive statistical analysis will simply over analyse the data-set, it is sill of interest to understand what type of victims are present in both data-sets if only, to get an understanding what types of victims give extensive information in their reports e.g. IP addresses. On the left side of figure 6.1 the previously defined victim demographics of the police data-set as a whole is defined, the right side represents the demographics as found in the traced attacks. The traced attacks target home broadband connections for 66% hence being a lot more prominent than they are in the overall data-set. Companies which are the sum of both the hosting and companies category account for just 22% of the traced attacks while they are 48% of all cases. For the educational category the results are much of the same, 27% in the overall data while just 12% were traced. One may conclude that the home broadband users that take the effort to go to the police supply ample information for the investigation. A possible explanation for this may be

that home users knowledgeable enough to identify a DDoS attack also have the knowledge to supply additional technical information. It does however not explain why the various companies do not supply this information, since most organisations have technical administrators, knowledge shouldn't be the limiting factor. Looking at the duration of the attacks, the traced attacks fall inline with the conclusions made in the previous chapter where there were peaks at various major minute marks such as 1min, 5min, 20min, 30min and 1h 20min. The median was 300s or 5min while the longest attack lasted 2h 11min.

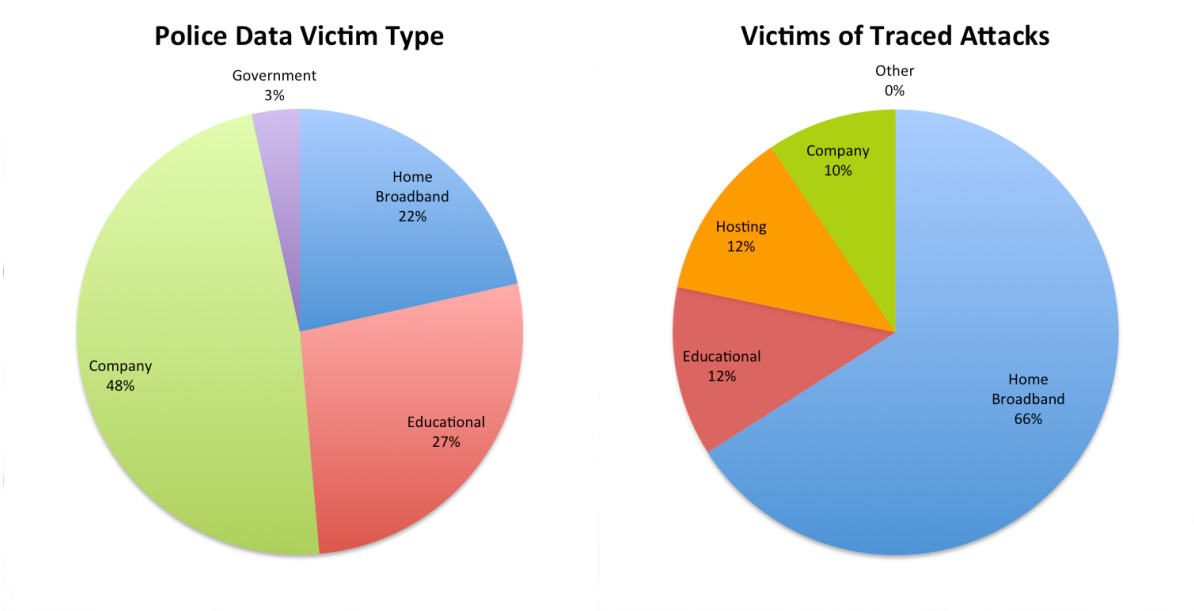

Figure 6.1: Victim distribution Police data vs. traced attacks

Figure 6.2 depicts two more attack properties, specific to the traced attacks. To the right, the distribution of protocols used to execute the amplification attacks is depicted. Again, much like in the previous chapter, DNS represents the majority of cases with almost three-fourths of the attacks, while NTP, CharGen and SSDP account for between 7% and 12%. SNMP and QoTD weren't detected in the traced attacks whatsoever. The diagram to the left of the attack protocols depicts another interesting statistic, the distribution of the number of connected domains per victim IP. Domains are not directly equal to IP addresses and various cases multiple domains may listen to the same IP address. One prominent example of this is shared hosting. Due to the inherent costs associated with renting an entire server from a hosting company, most entities in need of web-hosting make use of so-called shared hosting. In essence shared hosting means that multiple customers (in some cases more than 1000) make use of the same server/IP-address. This cost efficient solution has the downside that if an attacker is targeting one of those, all other users will be victim of the attack as well. Next to the large collateral damage, this phenomena poses an interesting question; who was the target of the attack? While this research won't be able to answer that question, the data does allow the researcher to see how big the impact of the attacks brought forward really were. In the case of the traced attacks, 80% of the cases have zero or one domain connected to them, 12% have between 2 and 12 domains while, except for two outliers at 702 and 515 domains representing just 1% of the data, the remaining 7% are between 13 and 120 domains. Pulling up the police reports these cases indicates that while some of these cases have such an extremely large group of victims, only one ended up going to the police, indicating that while many should have felt the effect only one decided to take the step and come forth to the police.

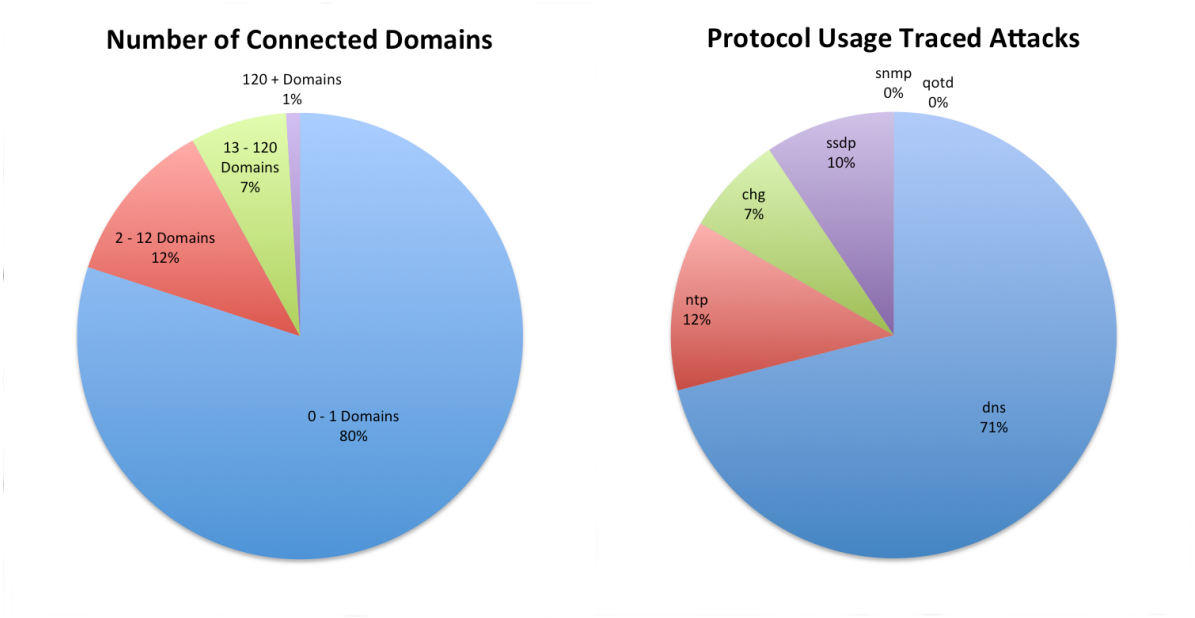

Figure 6.2: Protocol/Domain count distribution of traced attacks

6.2. Cost of attack vs. cost of damages

As discussed in the introduction to this paper, the DDoS as a service industry has grown immensely in its size and popularity, a quick google search suffices to see just how big and diverse the offerings are. From a pricing point of view, Hutchings and Clayton has done extensive research on 63 sites noting that prices of monthly subscriptions range between $0.19 and $14.99 with a median of $4.00 [18]. In some cases, the service providers even allow prospective buyers to test their service for free [26]. Seen as the majority of DDoS attacks are based on the amplification methodology executed via booter services, the booter costs structure is a great way of representing the costs associated with executing a DDoS attack on the side of the attacker. Knowing the burden of investment from an attackers point of view poses an interesting question; how much damage does one Euro worth of booter services yield on the side of the victim? To get an answer to just that question, the police report files were scanned for damage reports, most desirable were those that listed specific monetary damages and costs. Like mentioned in previous chapters, many of the cases lacked this kind of detail or simply noted that they couldn't properly describe the damages incurred. This kind of lack of detail makes sense for home internet connections and thus comes as no surprise. Bigger organisations such as schools and companies however, were able to provide a higher level of detail, resulting in a large count of qualitative descriptions such as "30h x 1000 students = 30 000 hours that couldn't be spent using the online education environment" as well as a 16 count of cases with monetary descriptions used for this analysis.

Table 6.2: Cost Data

Number of cases noting costs	16
Average victim costs	€15,352
Minimum victim costs	€1,740
Maximum victim costs	€70,000
Median cost DDoS service	$4.00

Since, much like previous sections, police case files are used as a source of information, the same rules as defined in chapter 3 can be used to deduce information. As mentioned in said chapter, it is the purpose of the report to proof the applicability of a certain criminal law, in the case of DDoS attacks article 138b and 350a of the Dutch book of criminal law (Wetboek van Strafrecht) are of relevance. Part of being able to proof the applicability of these articles is to state the damage that was incurred through the attack. As such many of the officers try to incorporate a section noting what the extent of

the damage is. Most cases denote the damages on a descriptive level without making these damages explicit with financial values. These descriptions range from ”I couldn’t use the internet” for most of the home broadband users, while the company related organisations suggested that they lost customers due to their website being offline, yet couldn’t calculate the exact losses. Educational victims suggested a loss of ”educational opportunities” such as the usage of online learning environments as well as content management systems such as Blackboard and Moodle. Lastly, the governmental institutions noted a loss of access to their website.

Table 6.3: Victim Cost factors

Internal Costs	External Costs
Unavailable sales website	DDoS mitigation services
Inactive employees	Third Party Consultancy services
Labour costs to repair/bring up the system	Third Party Investigative services
”Loss of face”	

Moving on to the 16 cases that noted specific financial damages, the first point of interest is the type of damages that are actually incurred by the victim. The reports mention a number of factors (noted in table A.1) that either internally or externally create costs. On the internal side, there is a loss of income incurred by a website or service being offline, hence no new customers can get in contact with the victim organisation or with other words loss of business. Secondly, employees of a system may not be able to work while the attack is ongoing, hence the cost of the employees is part of the damages. Thirdly, additional costs due to additional labour in order to repair- and bring up the system again. Lastly, there is a certain monetary value attached to the ”loss of face” of an organisation that was brought down by a DDoS attack. On the external side, there are cost incurred through the mitigation of a DDoS attack as well as costs incurred by third party investigative institutions or consultancy companies. Not all 16 cases mention every single one of the listed factors and as such make it hard to give a clear conclusion on the average costs are on a per factor basis. Generally speaking, the two external costs proved to be relatively high. A community of schools that noted the use of an external company to execute an investigation spent €70,000 and mentioned that the gross of the amount was spent on the investigation. Furthermore, 3 organisations mentioned the repair of their network separately at €1,740, €2,400 and €3,000. All others only state aggregate amounts. Of the 16 cases that mention monetary values 5 were schools and 11 companies. For a list of the values with corresponding target category see section A.2 in the Appendix.

$$\frac{\text{Ave. damages}}{\text{Med. cost booter service}} = \text{Ave. damages p. euro spent on a booter service} \tag{6.1}$$

Looking at the list of costs one may identify a number interesting insights. Firstly, in some cases the detail in the values given is very large while others are approximations. Some cases note the values down to the cent while other note general approximations in the order of thousands or ten-thousand. The police files support this further as some victims even note whether the value quoted includes taxes or not. Secondly, the range is very big, with a minimum of €1,740 and a maximum of €70,000, hence averaging €15,352.21 for the 16 cases. Having an understanding of both the costs on the side of the attacker as well as the victim allows us to circle back to the original question posed, how much damage does one Euro worth of booter services yield? To calculate just that the average damages incurred by the victims will be divided by the most frequently purchased booter subscription (see equation 6.1). Since the range of damages is so large, the same calculation will be repeated for the maximum and minimum amount of damages found in the police reports. It’s important to note that due to the limited data-set as well as the differences in detail, the results need to be taken with a grain of salt. Additionally, since all victims noted are of the company or school category, they are actors that are naturally more inclined to spend large sums in order to mitigate all issues resulting from the DDoS attack as opposed to a home broadband user, whom may not have the financial means nor need to do this. Lastly, since the paper by Hutchings and Clayton notes booter prices in US dollars rather than euro, a conversion step was made utilising the conversion rate as of the time of writing ($1 = €0.89). Combining all factors leads to the following calculations

$$\frac{1,740}{3.56} = 488,76 \quad (6.2) \qquad \frac{70,000}{3.56} = 19,662.92 \quad (6.3)$$

$$\frac{15,352.21}{3.56} = 4,312.42 \quad (6.4)$$

Dividing the damages by the DDoS costs yields an average about €4,312 with a minimum of €489 and a maximum of €19,663 damages worth per euro spent on a DDoS booter. One attack that could be traced in its entirety that also indicated a monetary value for damages incurred, recorded €3,000 worth of damages for an NTP based attack with a duration of about 20min and a load of about 4,000 packets per amplifier. One may conclude the damages incurred by one euro worth of damages is between 3 and 4 orders of magnitude bigger than the initial costs associated with executing the attack through a booter service.

6.3. Conclusion

It was the purpose of this chapter to continue the analysis of both the police and AmpPot data-set. While the previous two chapters focused on analysing the data-sets separately as well as comparing and contrasting them, this chapter will combine the two to get a deeper insight in the attacks that are being reported. In order to be able to combine the two, IP addresses from all police reports were compared to victim IPs noted in the AmpPot data-set. To ensure the attacks were actually the ones mentioned in the report, dates were used as a second criteria. Utilising this technique 32 cases were found, containing 138 unique attacks. Utilising this sub-set, the following sub-question may be answered:

SQ5: What additional information can be gathered by combining the honeypot data-set with the police dataset?

Due to the lack of detailed information in many of the case files, the sub-set used for this chapter is rather small and as such has limited validity however, it still provides a unique insight in the cases the police investigates and gives a good indication of what additional information secondary data sources may supply. Overall the data falls inline with that of the AmpPot data, where the vast majority of the cases target home broadband connections, trailed by the hosting and edu. In terms of protocol usage, the attacks noted in the police reports mostly make use of the DNS protocol, which falls in line with conclusions made in the previous chapter regarding the selection of DNS in regard to it being a needed basic service of the internet. Attack durations have the same peaks as mentioned previously, hence indicating the booter subscription models. Additionally, 80% of the attacks targeted IPs with none or one domain attached, while the remaining 20% represent cases in which an attack had multiple victims going up to as many as 700 domains attached to a single IP. Lastly, one of the properties the police tries to capture in a police report are the damages incurred by a victim. While most cases don't quote tangible much less monetary damages, a 16 count cases give a direct insight in the financial repercussions of an attack. Utilising this information and combining it with data regarding the cost structure of an attack shows that on average the costs created on the side of the victim are between three and four orders of magnitude higher than those on the side of the attacker. Hence proofing that even though DDoS as a service is widely available, it is still a very dangerous tool with a high impact on whoever is the victim.

7

Eliciting Requirements

Throughout the previous chapters a methodology was developed, police reports in regard to DDoS attacks were analysed and compared to a secondary honeypot data and the two data-sets were combined in order to gain additional information. Throughout these chapters a number of shortcomings and limitations were identified, that if handled correctly, could improve the investigative process as well as the data analytical possibilities. This chapter will focus on presenting these issues as well as eliciting a number of requirements to provide a better investigative environment. These requirements focus on three main sections of the law enforcement process specifically, starting from the first official contact of the police with the victim, and finishing with the reuse of left over information after an investigation has finished. More specifically, the three parts are defined as firstly, the police report as it shapes the basis of every investigation, the second area is the investigative process following the report and lastly, the aftermath of an investigation or rather the results and leftover data after a specific investigation has finished.

7.1. Police Report

Following the inherent order of the various steps within the law enforcement process, the first part to take a better look at is the police report. Police reports are created during initial contact of the police with the victim. The purpose of the report as defined in chapter 3.3 is to identify the criminal activity, how that activity was executed, who the victim is, who the declarant is and what the damages are. Additionally, it's the basis on which the investigation is judged for probability of success as well as providing investigators with key indicators to start the investigation. As indicated in chapter 4 the amount of detail about a DDoS attack varies significantly between the various reports and as such investigators have little to go on. A number of must have pieces of information specific to DDoS attacks, besides those that are required by law, are: the attacked IP address, the duration of the attack, the attack type and the specifications of the victim system. The reasoning behind this list is that they characterise the attack, give a better indication of the actual magnitude as well as resulting damages and may give identifying properties of the origin. Additionally, both the IP address and a specification of the IT infrastructure of the victim may be used to cross reference the case with booter databases as well as to identify further victims as detailed in section 7.2. In chapter 6 the cost analysis showed that in only 16 of the 209 cases analysed for this thesis, victims were able to quote detailed information about the costs incurred through the DDoS attacks. Both the attack duration and protocol used as well as a specification of the victim system may give the investigators a better understanding of the damage, if not on a financial level they at least may indicate the resulting effects. Furthermore, to make the information supplied more usable and cohesive as well as to support the police officer making the report, a standard template should be introduced. Templates are nothing new to the police administrative system, according to officers at the regional offices for many of the more common types of crime these templates already exist in order to ensure that all vital information is recorded. The template should be created in such a way that the aforementioned data is recorded in a logical and coherent way, meaning that the various questions supply enough information to the officer to create the report and to the victim to allow make them understand what information is needed. Furthermore, The template should be

created in such a way that the cognos system used to execute search queries in the police system as well as data analysis tooling can easily read through the files in an automated fashion to allow for extensive case analysis.

Table 7.1: Vital additional pieces of Information for police reports

#	Piece of information
1	Victim IPs
2	Attack Duration
3	Protocol used to
4	Specification of victim system

7.2. Investigation

Throughout the case analysis phase of this research a number of investigative steps were identified that, given that sufficient information was gathered in the police report, may allow the investigator to get a deeper insight in the case at hand. First and foremost it's important to make a differentiation between what the police categorises as digital and tactical investigative work, where the former indicates complex technical task requiring advanced knowledge of various IT systems, while the latter indicates more classical police work that relies on traditional investigative techniques. This distinction is rather important as many of the regional police forces lack educated digital investigative resources and as such rely on tactical investigations to gather the needed evidence. As indicated by multiple police reports, one of the most important leader generators within a DDoS investigation is the inspection of recent occurrences such as threats via email or phone as well as chatter from people directly connected to the victim. This has been especially prominent in cases regarding attacks on large companies as well as schools. Disgruntled employees or students threaten their target and later execute the attack via some kind of booter service. An even clearer example of this are the extortion cases mentioned in chapter 2.2 and 4. These kind of leads are vital as attackers often rely on the anonymity of booter services, yet leave identifying evidence in their threats.

The second important generator of leads are logging systems of wifi networks. This technique has proven very helpful in cases related to the educational sector. The underlying theory relies on the methodology used by the attackers. These often log onto the wifi system utilising their smart phone using their account provided by their educational organisation. The attack itself is then executed by surfing to a booter service, which in turn is tasked to attack the phone. At this point the phone has the public IP address of the organisation and as such the entire network becomes victim to the attack. The investigative clue lies in the logging system of the school which, if focused on the last few minutes before the attack was executed, will yield which account browsed to a booter service. The third and last tactical leads generator connects back to the shared hosting or multiple domain per IP address phenomenon introduced in chapter 2.2 and investigated in chapter 5 and 6. Shared hosting in the DDoS environment directly relates to multiple victim per attack or in other words, for every IP address attacked multiple domains feel the effects of the attack. This off course poses an interesting dilemma, if a certain party (victim to an attack) makes a official report with their local police office, were they the actual targeted victim. From an investigative point of view, having an overview of all victims is vital as any of them could be the original intended target and as such any additional investigative indicators may be found there.

Moving on to the digital investigation, there are a number of important pieces of information that need to be collected as well. First, the network logs of a victim may give a deeper insight in the attack on a technical level. The type of DDoS attack such as amplifier or not, protocol used, strength of the attack and length of the attack. All of these traces are of use to specific the attack and categorise it. One possible use for this may be the identification of a certain signature in the attack that may be traceable to other cases. The second, while most DDoS attacks make use of spoofed traffic to hide the origin of the attack, many attackers have made contact with their victim prior to the actual DDoS attack. This connection may be as simple as a visit to a website, but often is also in the form of a port-scan or comparable methodology without any legitimate use. Methodologies such as port scans are often executed from the home-network of the attacker and thus may be identifiable in the network

Table 7.2: Lead generators

Tactical	
#	Lead generator
1	Recent Incidents - Investigation of recent incidents such as threats via email or phone as well as chatter from people directly connected to the victim, which may indicate somebody's ill will towards the victim.
2	Logging systems - Logfiles of wifi networks, especially in organisations that require a username and password to log-on may identify the user that accessed a booter site.
3	Domains connected to victim IP - To get an overview of all victims of an attack. Allows the investigators to find intended victim, which may supply evidence towards a suspect.
Digital	
#	Lead generator
4	Network logs - Network logs of the victim allow the investigators to get a better insight in the properties of the attack, which helps to categorise the attack and potentially identify a known attack signature.
5	Suspicious IPs - IP addresses connecting to the victim in order to, for example execute port scans are a great piece of evidence as the attacker may use their home IP to execute the scan.
Stratigical	
#	Lead generator
6	Investigations regarding the facilitators behind the DDoS phenomenon should be moved to the high tech crime units, whom are responsible for facilitating parties and have capability to pursue these suspects.

logs. In addition to the digital and tactical sides of an investigation, there is also a strategical one. As discussed in actor analysis presented in chapter 2.2, the high tech crime unit of the Dutch police has selected facilitators as one of their central pillars. In relation to DDoS attacks the facilitator role is a rather large one represented by booter service providers. As such it's important to make that distinction when approaching a DDoS case, keeping a single attack or even multiple attacks with the regional forces, yet when the investigation yields significant insights in the service provider used to attack the victim, the high tech crime units should be informed. As such it also becomes the task of this unit to make data centrally available.

7.3. Use of data

Gathering information to eventually reach information superiority is one of the core goals of the police. The ability to have information on tap, ready to be used in new cases is invaluable. As with many cyber crime related cases, Distributed Denial of Service attacks generate large amounts of data, be it in terms of networks logs, attack signatures and victim IPs. Cases that directly relate to a booter service may in-fact yield entire attack databases as executed by a certain booter service. Saving this information and making it available is a great tool for fighting DDoS cases as cross matching the information and combining cases may yield a much better possibility of solving the case. A prime example would off course be a collection of victim IP addresses that has been collected over time that is combined with a database of a booter service. Having this wealth of information however, is not enough. As mentioned at multiple points throughout this research, and made specific in chapter 3, the data is almost impossible to query in a straight forward way. This is due to a number of issues: first, there is only one single cyber crime category rather than a more detailed structure. Due to the limited categorisation options as well as a limited description of what the categories entail, officers place DDoS attacks in other categories such as fraud. Additionally, since the system requires the user to query keywords right from the plain text police reports, it is important that there is a consistent usage of vocabulary. One of the main pitfalls that had to be circumvented through the development of the methodology in presented in chapter 3, was the collection of all synonyms used for DDoS attacks. While some of the synonyms make sense from the point of view of an untrained non-digital literate person, such as descriptive terms

of the actual result of an attack, they are very difficult to use for querying cases. As of now, an analyst needs first develop an overview of all types of these keywords, which also lead to a large number of false positives, in order to get the whole dataset. As such it's important to use the industry standard vocabulary. If it isn't possible to do this in the report itself, a keyword line at the bottom of the report should be introduced.

The use of data however does not need to be limited to internal data such as police case files or left over data from previous cases. Much like this research additional external data-sets may be introduced to extend the information one has about a certain DDoS attack. Throughout the previous chapters the AmpPot data-set was used, which contains a combination of data gathered via a set of amplification honeypots joined by historic DNS data, in order to supplement and compare and contrast the case files. This begs the question, should the police run their own set of honeypots to improve their DDoS investigations. Throughout this thesis and more specifically in chapters 5 and 6 the two data-sets were combined, as a result the attacks that could be traced through both data-sets became clearer and more information regarding their attack vectors could be added. However, most of these pieces of information are trivial to the investigation since the attack vectors are often the same across many different booters, combined with the inherently spoofed nature of the IP addresses these pieces of information won't yield attacker identifying information. As mentioned in section 7.2 as lead generator 3 however, there is a use case for historic DNS data as it may be used to identify all victims of a certain attack and through that the actual intended target. Historic DNS data can best be described as a historical ledger of what domains were attached to a certain IP address. Knowing the data and time of an attack will thus allow the investigator to get an overview of these domains. Historic DNS data is partially available for free online, while commercial packages offer a more detailed and wider view. As such one may conclude that running a entire honeypot data-set is not needed, however supplementing police investigations through historic DNS data can in-fact improve the investigative process.

Table 7.3: Uses of data

Data usage		
#	Type of data	Use
1	Victim IPs	Victim IP addresses as noted in the police report should be saved in a database. Whenever a attacker is caught or a booter service is brought down, the IPs can be cross-referenced to solve the case.
2	Booter databases	Whenever the facilitators behind the DDoS attacks are caught or their database become otherwise available, containing the targets they attacked as well as information regarding the attacker, they should be saved so they can be matched with victim IPs of past case files.
3	Historic DNS data	To identify all victims of a certain attack and to find the actual intended target. See table 7.2 #3.

7.4. Victim Demographics

Throughout chapter 5 the victim demographics of the Dutch police reports as well as the AmpPot data-set, representing the Dutch DDoS environment as a whole, were compared. Throughout the analysis it became clear that there is a significant difference between the two and that the police data-set portrays a different picture than the AmpPot dataset. As such there are victim groups that report much less frequently than others. Home broadband users and to an even bigger extent hosting companies are attack much more often than they make known. Analysing the willingness to report a crime is a research topic on its own, which has been heavily researched in recent years [53] [50] [15] [35]. Throughout this research however many of the companies contacted, both in the hosting and banking sector, argued that the biggest issue is the amount of of time and work connected to making a report. The amount of DDoS attacks, small and large, that reach these companies on a daily basis is so big that reporting them all would simply take too much time. Home broadband users also find making a official report a bit too serious, especially if the attacker is somebody they may know. As such it stands to reason that a simpler and more hassle free way of noting a DDoS attack should be introduced in order to allow these underrepresented groups to come forth. Within the police this issue has been heard as well, and as

such various incarnations of a ”hotline” or ”DDoS Desk” have been offered. Sadly none of these ideas have been executed. Implementing a first version of a point of contact with hosting parties may be a good start. Hosting companies represent a large percentage of the victims and as such can supply a large part of the unreported attacks. Eventually, a sort like procedure for all victim groups would be needed to ensure all victim groups would be represented.

7.5. Conclusion

While the majority of this thesis focused on the analysis of data, as well as comparing and contrasting various data-sets, it is the purpose of this chapter to take a step back and collect all the lessons learned about the Dutch police system, as well as the current investigative process and to dilute a number of requirements for both. These requirements are to ensure an improved process and to give the police an indication where the major short-comings in their system are located. Three sub-questions were devised to provide the Dutch police with a set of requirements starting with the following.

SQ6: Based on MQ1 and MQ2, what requirements may be identified in the police reporting system in order to improve the information basis to data querying and analysis?

The main issues in regard to the actual content of the police report are three fold: 1. the quantity of the data is lacking as many of the cases don’t contain vital properties of both the attack and the victim 2. the quality of the data varies widely between the cases, both in terms of usage of descriptive terminology as well as detail of attack properties 3. consistency is lacking as every officer writes their report in a completely different manner, negating any standard layout that would aid in the analysis of large amounts of cases. As such the following requirements should be set for police reports; A basic template needs to be devised that includes both detailed technical information on the attack as well as the victim, including the attacked IP address, the duration of the attack, the protocol used , as well as a description of the victims infrastructure such as whether they use shared hosting or host their own servers. Furthermore, the usage of standard terminology e.g. DDoS, needs to taught in order to make the cases easier to find for analysts later on.

SQ7: Based on MQ1 and MQ2, what lead generators may be identified for the investigative process in order to improve the investigative process both on a tactical and digital level?

The units investigating cyber crime cases, including DDoS cases, often execute what they call both tactical and digital investigations. The former focuses on more classical police work, while the latter focuses on complex technical tasks. DDoS cases yield a number possible investigative steps in both on digital but also the tactical level which is especially important for the regional police forces whom will take on more and more of these cases. On a tactical level there are a number of tactical leaders generators that must be checked when investigating a case. First, all occurrences of threats in every shape or form in the recent past need to be investigated, as these are often the basis for future actions. Second, in many instances the attacker makes use of the wireless internet connection of the victim which may yield identifying information, especially if the wifi connection requires an account. Lastly, many companies make use of shared hosting or other infrastructures that cause multiple domains to be attached to a single IP address. When a victim goes to the police, yet has no idea why they may be attacked, looking into other companies hosted on the same IP may lead to investigative cues. On a digital level, the network logs may yield a number of pieces of information to help identify which booter service was used to execute the attack. Network logs may also be used in order to look back in time to see whether the attack may have executed a port-scan to test the waters before executing the DDoD attack itself. Lastly, it’s important to make a distinction between the investigation of a DDoS attack and the facilitating party behind it. The booter service itself may yield a historical account of DDoS attacks on all kinds of victims. Furthermore, investigating these facilitator falls right into the area of expertise of the high tech crime unit of the Dutch national police, which could function as a central point to disperse data to regional forces.

SQ8: Based on MQ1 and MQ2, what requirements may be identified for the use and recycling of data across various DDoS cases in order solve cases more efficiently?

Collecting and recycling of data is vital to the DDoS phenomenon as even if cases cannot be solved at the time they are being reported, they may be solved down the road when a booter service is taken down and the database is secured. As such it is vital to make information easily available and easily searchable. To make this process as smooth as possible, the following requirements need to be implemented. First, the number of categories responsible for cyber crime is extremely limited and needs to be extended to allow for more specific searching. Second, the officers making the reports need to be educated on which categories to use as currently other categories such as fraud are used. Third, a keyword line should be added to the case file so officers can note various synonyms for a case, making searching through the plane text files easier for analysts.

8

Conclusion

Distributed Denial of Service attacks, are one of the most damaging and frequent type of cyber crime today. Their application ranges over a wide variety of victims from broadband home connections to schools all the way to governmental organisations. The growth in the usage of the attack type, especially via the amplification and reflection attack methodology has not gone unnoticed. As such the DDoS phenomenon has characterised itself as a prime example of the commoditisation of cyber crime, where the amount of service providers and the ease of use have increased, while costs have decreased to trivial amounts. Due to this the application has become a commodity to be used by everybody with internet access and basic technical knowledge. While DDoS attacks in general have been studied in a multitude of ways, most these studies focus on the zombie machines or bots used to execute the attack, or the analysis of ISP data to get a general insight in the attack. This thesis however, uses the police reporting system as a database to query a DDoS data-set from the point of view of the Dutch national police. This unique point of view allows the researcher to get a better understanding of what kind of DDoS cases make it the police and how these are characterised as well as how they compare to the DDoS population as a whole. Lastly, this thesis uses the data gathered throughout the analysis phase to set a number of requirements or improvements to be used by the Dutch police in order to improve the investigative process. To reach these goals and to provide a logical order to research, three main research questions are spelled out. In order to answer these questions cohesively eight-subquestions help to shed light on important aspects of the main-questions. This chapter is to aggregate the results found to these eight questions to answer the three main questions. Additionally, suggestions for future research will be proposed, in order to allow future researchers to build on this work, as well as to give an indication on what sort-like research could improve on.

8.1. Dutch Police Data

To shape a basis for the two following research questions, the first research question posed focuses on data available with the Dutch police and reads as follows:

MQ1: What do DDoS attacks reported to the Dutch police look like and how can these cases be classified?

The DDoS cases recorded by the Dutch police show three major types of attacks in their data-set; High school curiosity, which are attacks executed by students who do not have malicious intend but are rather interested in the possible effects of testing the network or are simply bored and were playing around with a booter service. Gamers and bullying are another returning theme, where games DDoS each other to get a competitive advantage over another player or sometimes simply to nag each other. A more serious derivative of this type are bullying cases, in which often high school age teenagers, are bullied and sometimes even pressured into passing on passwords or other information. Lastly there are the extortion cases, these cases have the clear intent of forcing a victim into paying a certain sum in a crypto currency or otherwise endure a DDoS attack. The analysis however shows, that while there are a few reports of threats coming in, only one of them was executed. While many of the threats are made in the name of large criminal groups in order to give the threat credibility, most of these don't

seem to originate from these groups. While not all cases fall into these three classifications, they are interesting nonetheless and show the most prominent overall themes.

Furthermore, the data shows that the vast majority of attacks target companies, the type of company specifically varies per case and while IT related categories add up to about 37% of the company related cases there aren't any significant results to be found. The second and third largest group are the educational sector and private broadband connections respectively. Within the educational sector, the mammoth share is held by high schools while the broadband connections are mostly made up of gaming related targets. Taking a better look at the attackers whom actually executed the attack and were interviewed by police afterwards and whom are exclusively high school age male teenagers, supplies additional insights. As such only three-fourths of the attackers who the actual meaning of the term DDoS while even less seem to understand how to actually execute one. One may conclude that while many of them are playing around with booter services they don't understand what the purpose of tool, nor what the entailing repercussions would be. When asked whether they would know that utilising these tools was illegal, only a small minority of the attackers answers with yes. Aggregating all this information clarifies just how normal it is, for especially young teenagers, to utilise booter services. They often don't know what they are doing exactly other than that it turns of the internet for a short period of time. It may also be concluded that many of them simply follow quick tutorials on "how to turn of a friends internet" found in online chatrooms or forums and as such have understanding of what they are doing.

8.2. The Police vs. the world

MQ2: How does the Dutch police data gathered through MQ1 compare to data available about the DDoS population as a whole and what insights can be gathered by combining the two?

Much like the world wide DDoS environment, attacks on Dutch victims are predominantly of the amplification/reflection variety. Comparing the victim demographics of the police system to that of the DDoS population as a whole shows that they differ significantly. Hosting seems to be a especially often targeted victim, yet the number of police reports is comparatively low. A possible explanation for this is the inherent nature of the hosting business and thus the presence of a good mitigation system that makes many of the attacks unsuccessful. Additionally, there may be the simple issue of having so many attacks that's impossible to go to the police to make a report for every single one. The second big difference is the educational sector, which are only target to roughly 1% of all attacks, yet represent more than one-fourth of all police reports. This difference may in large be due to the educational sector often knowing who the attacker is as opposed to most other victims. Lastly, the home-broadband connections represent almost half of the attacks, while only being 22% of the police reports. As the majority of these attacks are executed by teenagers playing video games and the like, it makes sense that very little would go to the police and make a formal report, especially since the monetary damages would be rather low too. While most cases don't quote tangible much less monetary damages, a 16 count cases give a direct insight in the financial repercussions of a attack. Utilising this information and combining it with data regarding the cost structure of an attack shows that on average €1 of investment into a booter service yields €4,312 worth of damages on the victim side. This tally runs up quickly if consultancy companies are called-in to provide investigative services.

8.3. Requirements

MQ3: Based on the insights gathered in MQ1 and MQ2, how could reporting, investigating and re-use of data be improved?

The main issues in regard to the actual content of the police report are three fold: 1. the quantity of the data is lacking as many of the cases don't contain vital properties of both the attack and the victim 2. the quality of the data varies widely between the cases, both in terms of usage of descriptive terminology as well as detail of attack properties 3. consistency is lacking ass every officer writes their report in a completely different manner, negating any standard layout that would aid in the analysis of large amounts of cases. As such the following requirements should be set for police reports; A basic template needs to be devised that includes both detailed technical information on the attack as well as

the victim, including the attacked IP address, the duration of the attack, the attack type, the IP address of the victim as well as a description of their hosting infrastructure such as whether they use shared hosting or host their own servers. Furthermore, the usage of standard terminology e.g. DDoS, needs to taught in order to make the cases easier to find for analysts later on. The units investigating cyber crime cases, including DDoS cases, often execute what they call both tactical and digital investigations. The former focuses on more classical police work, while the latter focuses on complex technical tasks. DDoS cases yield a number possible investigative steps in both on digital but also the tactical level which is especially important for the regional police forces whom will take on more and more of these cases. On a tactical level there are a number of tactical leaders generators that must be checked when investigating a case. First, all occurrences of threats in every shape or form in the recent passed need to be investigated, as these are often the basis for future actions. Second, in many instances the attacker makes use of the wireless internet connection of the victim which may yield identifying information, especially if the wifi connection requires an account. Lastly, many companies make use of shared hosting or other infrastructures that cause multiple domains to be attached to a single IP address. When a victim goes to the police, yet has no idea why they may be attacked, looking into other companies hosted on the same IP may lead to investigative cues.

On a digital level, the network logs may yield a number of pieces of information to help identify which booter service was used to execute the attack. Network logs may also be used in order to look back in time to see whether the attack may have executed a port-scan to test the waters before executing the DDoD attack itself. Lastly, it's important to make a distinction between the investigation of a DDoS attack and the facilitating party behind it. The booter service itself may yield a historical account of DDoS attacks on all kinds of victims. furthermore, investigating these facilitator falls right into the area of expertise of the high tech crime unit of the Dutch national police, which could function as a central point to disperse data to regional forces. Collecting and recycling of data is vital to the DDoS phenomenon as even if cases cannot be solved at the time they are being reported, they may be solved down the road when a booter service is taken down and the database is secured. As such it is vital to make information easily available and easily searchable. To make this process as smooth as possible, the following requirements need to be implemented. First, the number of categories responsible for cyber crime is extremely limited and needs to be extended to allow for more specific searching. Second, the officers making the reports need to be educated on which categories to use as currently other categories such as fraud are used. Third, a keyword line should be added to the case file so officers can note various synonyms for a case, making searching through the plane text files easier for analysts.

8.4. Contributions

While the above stated answers to the three main research questions represent the most important and immediate contributions of this paper, there are a number of secondary contributions that are are important for both academia and the police alike. These secondary contributions are often results required to the reach the final answers but are interesting on their own none the less. The first contribution that falls into this category is the development of the a methodology to create a data-set for analysis on the basis of the administrative system of the Dutch Police. While the development process was targeting DDoS attacks specifically, the method is easily manipulated to work for other cases as well. The second contribution is the list of manually categorised AS from the AmpPot data-set that was resubmitted to the TU Delft for future research. Third contribution is the script to quickly analyse the the police report data set for any victim IP addresses. Fourth, a descriptive analysis of the DDoS cases in the Dutch police administrative system. Additionally, a taxonomy of the victim demographics for both the Netherlands as a whole and the police data-set were created. Furthermore, a attacker analysis based on the interrogation transcripts was executing giving an insight into the attackers knowledge level. Lastly, an overview of the Dutch DDoS environment was created and compared to the data provided by the police displaying the differences between the two.

8.5. Suggestions for Future Research

To derive future research ideas from this thesis one must look at two different type of derivatives; research on DDoS attacks and research on police data. In regard to the former; due to its prevalence the DDoS phenomenon is a heavily researched topic by academia world wide. Much research has gone

into the analysis of the technology used by the attackers and how attacks can be mitigated or filtered form legitimate traffic. There is however still a lot to be learned about the effect of the attacks and how society is dealing with them. As such a more international approach to the what victims decide to go to the police, what their damages are and how they decide to do so would be in order. One example could be how various countries differ in regard to what victims go to the police and what the underlying reasons are for this behaviour. Furthermore, if the recommendations in regard to what data to record in the official reports would be implemented, reanalysing the police data utilising IP matches to other honeypot data, may yield a more valid view as the number of cases with matches are still relatively slow in this thesis. From the police side of things, this thesis gives an insight in the actual work done by Dutch police in regard to DDoS cases. Ideally this type of research would be executed many more types to put the results of thesis into perspective but also to allow the leadership of the police to distribute resources accordingly. It would also be of interest to have a continuous business analysis system developed, that would give the police an insight in change sin crime statistics so problems could be resolved earlier.

Appendix A

A.1. Police Data Analysis

Throughout this section an overview of the activities executed to analyse the police report dataset will be given. As mentioned in section 3.3, the data was exported form the Cognos system to a large Excel file containing a number of characteristics but mainly consisting of plain text entries representing the original reports as well as so called mutations which contain additional data gathered throughout the investigation. To connect the report to the various mutations, every case contains a unique so-called BVH number that is attached to every entry in database. The BVH numbers look as follows ”PLXXXX_BVH_YYYYXXXXXX” and can roughly be split into two parts. The first part ”PLXXXX” denotes the regional police force whom recorded the case where ”XXXX” is a number that identifies the regional unit. The second part ”BVH_YYYYXXXXXX” is the unique identifying BVH code of the regional BVH system. The four ”Y” identify the year in which the case was recorded while the six ”X” are a increasing number starting from 0.

```
#! /usr/bin/python
import csv

txt_file = r"./Find_EDU.csv"
csv_file = r"./EDU_Only_Test1.csv"

BVH_numb = 'PL2400_BVH_2014014455'

out_csv = csv.writer(open(csv_file, 'w'))

reader = csv.reader(open(txt_file, "rU", encoding='utf8', errors='ignore'),delimiter=';')
for row in reader:
^^Iif BVH_numb in row: out_csv.writerows([row])
```

A Python script that prints all mutations connected to a certain case-number.

```
cat alt_DDoS.csv | strings | grep -E -o
^^I"([0-9]{1,3}[\.]){3}[0-9]{1,3}" | sort | uniq -c | sort
```

A quick Bash script that goes through the entire Cognos export looking for IP addresses using a regular expression.

A.2. Police Data - Victim Costs

The table below is a list of all 16 police reports that noted specific financial damages due to a DDoS attack. Next to the costs, the organisation type is noted.

Table A.1: Victim Costs

#	Victim Category	Costs (€)
1	school	14,007.50
2	school	17,329.44
3	school	70,000.00
4	school	20,000.00
5	school	6,581.45
6	company	3,000.00
7	company	3,701.70
8	company	1,740.00
9	company	20,000.00
10	company	17,000.00
11	company	6,000.00
12	company	20,000.00
13	company	10,000.00
14	company	3,875.00
15	company	2,400.00
16	company	30,000.00

A.3. AmpPot Data

The AmpPot dataset was collected by a group of researchers to get a deeper insight in the inner-workings of the popular amplification and reflection DDoS techniques. As such the data was collected by first developing a number of honeypots that would server up some of the most most prominent amplification protocols, namely ”QotD (17/udp), CharGen (19/udp), DNS (53/udp), NTP (123/udp), SNMP (161/udp) and SSDP (1900/udp).” [39]. While the data set as a whole makes use of 21 honeypots and two additional protocols namely, MSSQL (1434) and SIP (5060/5061) [28], this research will focus on just 8 of these honeypots, all located at major Japanese ISPs. The benefit of using just these 8 rather than the full 21 is that these were all consistently run in the same period of time (2014-2015) and as such portray a valid and fairly current overview of DDoS attacks in that period of time. It must be noted that several steps of data processing were executed by Arman Noroozian, such as adding historic DNS data in order to get an insight in the domains connected to the IP addresses at the time of attack, for his very own research before kindly offering the data set to be used in this paper. Examples of processing steps taken are the binding of attacks that were run over several days and thus over several log files, as well as identifying the domain count underlying the various attacked IPs. For more information on these calculations please reference [39]. To get a better understanding of what the actual data looks like, the table below shows an overview of the most important fields for this research and by using three imaginary data points. For the original data-set get in contact with Arman Noroozian at the TU Delft. For a extensive definition of the dataset and how the various honeypots were created please refer to the paper ”AmpPot: Monitoring and Defending Against Amplification DDoS Attacks” Krämer et al..

Table A.2: AmpPot Sample Data

#	target_ip	sensor_id	service	duration	attack_time	packets	as_type	raw_country	raw_as	dc
0	123.456.789.123	sensor001	DNS	578.0	2014-12-24 02	434	hosting	Netherlands	AS12345 Big Network	1000
1	234.567.890.123	sensor002	NTP	576.0	2014-12-24 04	431	isp-broadband	Netherlands	AS23456 Medium Network	50
2	345.678.901.234	sensor007	SSDP	591.0	2014-12-24 02	440	edu	Netherlands	AS4567 Small Networks	1
...	...	...	...	...	...	...	...	...	...	...

A.4. AmpPot Data Preperation

Central to this thesis is the discussion about DDoS victims and how these can be categorised. As such this thesis knows two chapters discussing and comparing the victims of DDoS attacks from the point of view of the police as well as the AmpPot data-set. While the categorisation of police case files is discussed in section A.1, this section will discuss the data preparation steps taken to make the AmpPot data ready for analysis. Just like the data analysis section, python in combination with the pandas plug-in as well as several secondary libraries were used for the data preparation. Before going into the data manipulation part of this section, one must note the inherent structure of the AmpPot data-set, which rather than the police data, doesn't contain one data-point per attack but rather notes connections made to IP addresses from 8 honeypots via 6 protocols within a certain period of time. In other words, to be able to compare the AmpPot data with the police data, the data-points need to be consolidated to attacks. While there are many ways to define an attack, multiple were tested, for the purposes of this paper the following definition will be used: all data-points containing the same target IP address utilising the same attack protocol and having the same start date and hour will be considered one attack. The last of the three, the starting date + hour is a column that needs to be created as the starting date in the data-set also includes minutes and seconds, which during testing seemed too precise for the purposes of the Dutch data. The creation of the so-called "attack_time" column was executed as follows:

```
# This script parses all DDoS start times
PSDT = [dp.parse(x) for x in TUD['start_time']]
#Selects the year, month, day and hour
ST = [x.strftime("%Y-%m-%d %H") for x in PSDT]
#Writes those to a new column in the data set
#called attack_time
TUD['attack_time'] = ST
```

Hence allowing the following function to be used to select all unique attacks:

```
def count_unique_attacks(df):
    attacks = df[['target_ip', 'service', 'attack_time',]]
    attacks = attacks.drop_duplicates()
    return len(attacks)
```

Executing this definition yields a number of 53055 for the total amount of attacks recorded in the AmpPot data-set for dutch victims within the timeframe of 2014 and 2015:

```
# Execute the unique attacks function and count
total_attacks = count_unique_attacks(TUD)
print(total_attacks)
```

It's important to note that the AmpPot data-set knowns multiple ways to categorise the victims yet none, beside the CAIDA classification, is 100% complete. For an extensive discussion on the up- and down-sides of these classification types please see Noroozian et al.. In short, the downside of the CAIDA classification lies partially in it's rough graining (it only knows three types of classifiers: Transit/Access, Content and Enterprise) as well as its relatively high false classification rate. As such this thesis uses the partially implemented AS type classifier as implemented by Arman Noroozian. As this classifier was implemented on an international level with size as the major deciding factor, it comes as no surprise that many of the Dutch ASes lack a tag. As such all DDoS attacks without an AS tag were selected, grouped by their raw "Autonomous System" number and name and sorted by size with the following python pandas command:

```
#Select all unique DDoS attacks
x=TUD.drop_duplicates(subset=['target_ip', 'service', 'attack_time'])

#Find all cases in which the as_type field is empty
l=x[x['as_type'].isnull()]
```

```
#Count those cases
len(l)
```

Of the aforementioned 53055 cases overall, 26648 cases or roughly half did not contain an AS tag. To select these cases, sort them and export them, the following command was executed:

```
#Select all unique DDoS attacks
x=TUD.drop_duplicates(subset=['target_ip', 'service', 'date'])

#Find all cases in which the as_type field is empty
l=x[x['as_type'].isnull()]

#Group the resulting cases by their AS number/name and sort
#them by size
v=l.groupby('raw_as').size().sort_values(ascending=False)

#Export the resulting list to csv using ';' as a seperator
v.to_csv('AS_lookup.csv', sep=';')
```

The resulting list was manually filled by looking up each AS number specifically and finally imported back into the data-set. It's important to note that Autonomous Systems are not bound by geographical boarders and as such an AS may be partially located in the Netherlands and partially in one or more other countries. This thesis only analyses the part that is specifically located in the Netherlands.

A.5. AmpPot Data Analysis

To analyse the fairly large and extensive honeypot data set from Japan, the data analysis toolkit Pandas was used. Pandas is based on python and as opposed to commercial products such as SPSS is open source and rather than providing the user with a graphical interface to interact with, can be controlled utilising the python programming language as well as some additional Pandas syntax. The benefit of using Pandas as opposed to other products is that the code can easily be shared for other researchers to use and also provides reviewers a convenient way of validating the results. To make use of this very property, the various code snippets used to analysis the data set will be quoted here. In addition to the code itself, comments will give the unpracticed reader an idea of what the various commands are supposed to do.

```
#Import the various toolkits and libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import dateutil.parser as dp
import datetime as dt

#Import the AmpPot data-set
TUD=pd.read_csv('ddos_NL.csv',delimiter='|',
parse_dates=True,infer_datetime_format=True)
```

To get a better understanding of what kind of victims were targeted by the attacks, the attacks were classified. Since there is no personal information available on the various IP-addresses the autonomous system (AS) of the IP-address is used as a proxy identity. Categorising the attacks works as follows:

```
#Select unique attacks
x=TUD.drop_duplicates(subset=['target_ip', 'service', 'attack_time'])
#Categorise by AS type
x.groupby("as_type").size().sort_values(ascending=False)
```

Continuing with the victim analysis, another interesting characteristic is the domain count underlying the attacked IP address. In other words, how many factual victims are there for one target IP address.

```
x=TUD.drop_duplicates(subset=['target_ip', 'service', 'attack_time'])
x['dc'].describe()
```

To get a better understanding of the attack the attack type is analysed. On a general level this was done as follows:

```
x=TUD.drop_duplicates(subset=['target_ip', 'service', 'attack_time'])
x.groupby("service").size().sort_values(ascending=False)
```

Looking at the same data over time, in the same way Arman Noroozian (The below quoted code was developed by him as well) has in his own paper.

```
#Create a grouping by date and attack type
protocol_gr = TUD.groupby(['date', 'service'])

#Create a time-series of the attacks utilising the unique
# attack definition
timeseries_attacks = pd.DataFrame(protocol_gr.apply(count_unique_attacks), columns=["
timeseries_attacks = timeseries_attacks.unstack("service")
timeseries_attacks.columns = timeseries_attacks.columns.droplevel()


#Utilising a pre-defined color pallet
with some_contract_colors:
#Create a graph with multiple lines (per attack type)
    fig, ax = plt.subplots()
    ax = timeseries_attacks.plot(ax=ax, style=['-p', '-',
    '--', 's', 'o', ':'], markersize=8, linewidth=3)

    for tick in ax.yaxis.get_major_ticks():
         tick.label.set_fontsize(28)

    for tick in ax.xaxis.get_major_ticks():
         tick.label.set_fontsize(25)
#Add labels to the graph
    ax.xaxis.set_label_text("Date", fontsize=30)
    ax.yaxis.set_label_text("Number Attacks", fontsize=30)

#Create a legend for better understanding
    plt.xticks(rotation=0)
    plt.gca().get_legend().set_title('Protocol')
    plt.gca().get_legend().get_title().set_fontsize('33')
    plt.setp(plt.gca().get_legend().get_texts(), fontsize='33')

    ax.figure.set_size_inches(25,9)
#Save the figure to pdf
    plt.savefig("time_series_attacks.pdf")
```

A.6. Sample requests

The pieces of code below represent various types of requests and their results that frequently used for DDoS attacks. The purpose of this section is to give the reader an understanding how little the request is, yet how big the answer is and as such how amplification fundamentally works.

DNS or Domain Name System protocol is a hierarchical way of naming and storing names within a network.

```
jank$ dig @8.8.8.8 www.google.com
```

```
; <<>> DiG 9.8.3-P1 <<>> @8.8.8.8 www.google.com
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 19884
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;www.google.com.       ^^I^^I^^IIN     ^^IA

;; ANSWER SECTION:
www.google.com.^^I^^I299    ^^IIN     ^^IA      ^^I216.58.212.164

;; Query time: 33 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Fri Sep 16 10:01:23 2016
;; MSG SIZE  rcvd: 48
```

NTP or Network Time Protocol which is used for time synchronisation within a network.

```
jank$ ntpdate -q pool.ntp.org
server 83.98.201.134, stratum 2, offset 0.005728, delay 0.03731
server 188.166.95.178, stratum 2, offset 0.003305, delay 0.03818
server 31.3.104.60, stratum 2, offset 0.004409, delay 0.03865
server 89.188.26.129, stratum 2, offset 0.004099, delay 0.04057
16 Sep 10:06:19 ntpdate[49477]: adjust time server 83.98.201.134
offset 0.005728 sec
```

QoTD or Quote of The Day is a an protocol used to share quotes.

```
$ telnet> telnet nicksosinski.com 17
Trying 68.117.3.73...
Connected to nicksosinski.com.
Escape character is '^]'.
I'm a private person too, and we don't ever film anything in
our home because it's off limits. It's like letting people
see your messy house.
Connection closed by foreign host.
```

CHARGEN or Character Generator Protocol is a protocol used for testing and measuring.

```
$ telnet localhost chargen
Trying 127.0.0.1...
Connected to localhost.
Escape character is '^]'.
!"#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abc
"#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcd
#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcde
$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdef
%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefg
&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefgh
'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghi
()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghij
)*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijk
*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijkl
+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklm
,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmn
-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmno
./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmnop
```

```
/0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmnopq
^]
telnet> quit
```

Bibliography

[1] Akamai. Akamai's state of the internet/security Q4 2015 report. Technical report, 2015. URL https://www.mendeley.com/viewer/?fileId=1aca831e-e0b1-6439-7278-1f1bde73a14a{&}documentId=e4832292-4709-31fd-8e32-37d5e9a431b6.

[2] Adri Amelsvoort, Imke Rispens, and Henny Grolman. *Handleiding verhoor*. Stapel & De Koning, 6th edition, 2015. ISBN 9035244532.

[3] Ross Anderson, Chris Barton, Rainer Böhme, Richard Clayton, Michel J. G. van Eeten, Michael Levi, Tyler Moore, and Stefan Savage. *The Economics of Information Security and Privacy*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-39497-3. doi: 10.1007/978-3-642-39498-0. URL http://link.springer.com/10.1007/978-3-642-39498-0.

[4] Abbass Asosheh and Naghmeh Ramezani. A comprehensive taxonomy of DDOS attacks and defense mechanism applying in a smart classification. *WSEAS Trans. Comput.*, 7(7):281–290, 2008.

[5] David Barroso. Botnets-the silent threat. *Eur. Netw. Inf. Secur. Agency*, 15:171, 2007.

[6] Johannes M Bauer and Michel Van Eeten. Introduction to the economics of cybersecurity. 2011.

[7] Belastingsdienst. Vaarwel blauwe envelop - welkom Berichtenbox!, 2015. URL http://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/campagnes/landingspaginas/prive/digitale{_}post/.

[8] Frank Bernaards, Eileen Monsma, and Peter Zinn. High Tech Crime - Criminaliteitsbeeldanalyse 2012. Technical report, KLPD, Wierde, 2012.

[9] Russell Brandon. The DDoS attack that cried wolf - The Verge, 2016. URL http://www.theverge.com/2016/4/26/11512032/ddos-ransom-armada-collective-denial-of-service-threat.

[10] Kim-Kwang Raymond Choo. The cyber threat landscape: Challenges and future research directions. *Comput. Secur.*, 30(8):719–731, 2011. ISSN 0167-4048.

[11] Even Cooke, Farnam Jahanian, and Danny McPherson. USENIX SRUTI '05 Technical Paper. *USENIX*, SRUTI, 2005. URL https://www.usenix.org/legacy/publications/library/proceedings/sruti05/tech/full{_}papers/cooke/cooke{_}html/index.html.

[12] C Czosseck, G Klein, and F Leder. On the arms race around botnets - Setting up and taking down botnets. In *Cyber Confl. (ICCC), 2011 3rd Int. Conf.*, pages 1–14, 2011.

[13] Jakub Czyz, Michael Kallitsis, Manaf Gharaibeh, Christos Papadopoulos, Michael Bailey, and Manish Karir. Taming the 800 pound gorilla: The rise and decline of ntp ddos attacks. In *Proc. 2014 Conf. Internet Meas. Conf.*, pages 435–448. ACM, 2014. ISBN 1450332137.

[14] Rob Davies. UK businesses battling huge rise in cybercrime, report says, 2016. URL http://www.theguardian.com/technology/2016/feb/25/cybercrime-uk-businesses-battling-huge-rise-silver-fraudsters.

[15] Robert C Davis and Nicole J Henderson. Willingness to report crimes: The role of ethnic group membership and community efficacy. *Crime Delinq.*, 49(4):564–580, 2003. ISSN 0011-1287.

[16] Felix Freiling, Thorsten Holz, and Georg Wicherski. Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks. *Comput. Secur. 2005*, pages 319–335, 2005.

[17] John Hawkinson and Tony Bates. Guidelines for creation, selection, and registration of an Autonomous System (AS). 1996. URL https://tools.ietf.org/html/rfc1930{#}section-3.

[18] Alice Hutchings and Richard Clayton. Exploring the Provision of Online Booter Services. *Deviant Behav.*, 2016.

[19] Nicholas Ianelli and Aaron Hackworth. Botnets as a vehicle for online crime. *FORENSIC Comput. Sci. IJoFCS*, page 19, 2005.

[20] Imperva. The Top 10 DDoS Attack Trends. Technical report, 2015. URL https://www.imperva.com/docs/DS{_}Incapsula{_}The{_}Top{_}10{_}DDoS{_}Attack{_}Trends{_}ebook.pdf.

[21] IMPERVA Incapsula. Global DDoS Threat Landscape Q2 2015. Technical report, IMPERVA, 2015. URL http://lp.incapsula.com/rs/804-TEY-921/images/DDoSReportQ22015.pdf.

[22] John Ioannidis and Steven Michael Bellovin. Implementing pushback: Router-based defense against DDoS attacks. 2002.

[23] M J G van Eeten J. M. Bauer. ITU Study on the Financial Aspects of Network Security: Malware and Spam. *ITU Final Rep.*, July 2008, 2008. URL http://www.itu.int/ITU-D/cyb/cybersecurity/docs/itu-study-financial-aspects-of-malware-and-spam.pdf.

[24] Cheng Jin, Haining Wang, and Kang G Shin. Hop-count filtering: an effective defense against spoofed DDoS traffic. In *Proc. 10th ACM Conf. Comput. Commun. Secur.*, pages 30–41. ACM, 2003. ISBN 1581137389.

[25] Marianne Junger, Lorena Montoya, Pieter Hartel, and Margo Karemaker. MODUS OPERANDI ONDERZOEK NAAR DOOR INFORMATIE EN COMMUNICATIE TECHNOLOGIE (ICT) GEFACILITEERDE CRIMINALITEIT. Technical report, Twente, 2013. URL http://www.websitevoordepolitie.nl/archief/politiewerk-na-de-digitale-revolutie-706.html.

[26] Mohammad Karami, Youngsam Park, and Damon McCoy. Stress Testing the Booters: Understanding and Undermining the Business of DDoS Services. *arXiv Prepr. arXiv1508.03410*, 2015.

[27] Labs. Kaspersky. DDOS Intelligence Report Q3 2015. Technical report, Kaspersky, 2015. URL https://cdn.securelist.com/files/2015/11/Q3{_}DDoS{_}report{_}final{_}EN.pdf.

[28] Lukas Krämer, Johannes Krupp, Daisuke Makita, Tomomi Nishizoe, Takashi Koide, Katsunari Yoshioka, and Christian Rossow. AmpPot: Monitoring and Defending Against Amplification DDoS Attacks. In *Res. Attacks, Intrusions, Defenses*, pages 615–636. Springer, 2015. ISBN 3319263617.

[29] T K T Law, J C S Lui, and D K Y Yau. You can run, but you can't hide: an effective statistical methodology to trace back DDoS attackers, 2005.

[30] Barry M Leiner, Vinton G Cerf, David D Clark, Robert E Kahn, Leonard Kleinrock, Daniel C Lynch, Jon Postel, Larry G Roberts, and Stephen Wolff. A brief history of the Internet. *ACM SIGCOMM Comput. Commun. Rev.*, 39(5):22–31, 2009.

[31] Joseph A Maxwell. *Qualitative research design: An interactive approach: An interactive approach*. Sage, 2012. ISBN 1412981190.

[32] Jelena Mirkovic and Peter Reiher. A taxonomy of DDoS attack and DDoS defense mechanisms. *ACM SIGCOMM Comput. Commun. Rev.*, 34(2):39–53, 2004.

[33] Jelena Mirković, Gregory Prier, and Peter Reiher. Attacking DDoS at the source. In *Netw. Protoc. 2002. Proceedings. 10th IEEE Int. Conf.*, pages 312–321. IEEE, 2002. ISBN 0769518567.

[34] Tyler Moore, David Pym, and Christos Ioannidis. *Economics of Information Security and Privacy*. Springer US, Boston, MA, 2010. ISBN 978-1-4419-6966-8. doi: 10.1007/978-1-4419-6967-5. URL http://www.springerlink.com/index/10.1007/978-1-4419-6967-5.

[35] Kristina Murphy and Julie Barkworth. Victim willingness to report crime to police: Does procedural justice or outcome matter most? *Vict. Offender.*, 9(2):178–204, 2014. ISSN 1556-4886.

[36] Chris Nap. Politie gaat Summ-IT gebruiken voor opsporing, 2012. URL http://www.automatiseringgids.nl/nieuws/2012/18/politie-gaat-summ-it-gebruiken-voor-opsporing?smgca.

[37] Jose Nazario. DDoS attack evolution. *Netw. Secur.*, 2008(7):7–10, 2008. ISSN 13534858. doi: 10.1016/S1353-4858(08)70086-2.

[38] Arbor Networks. Worldwide Infrastructure Security Report About Arbor Networks. Technical report, 2015.

[39] Arman Noroozian, Maciej Korczynski, Carlos Hernandez Ganan, Daisuke Makita, Katsunari Yoshioka, and Michel Van Eeten. Who Gets the Boot? Analyzing Victimization by DDoS-as-a-Service. 2016.

[40] United States Government Accountability Office. Public and Private Entities Face Challenges in Addressing Cyber Threats. Technical report, http://www.gao.gov/new.items/d07705.pdf, 2007. URL http://www.gao.gov/new.items/d07705.pdf.

[41] Politie. Politietaken | politie.nl, 2016. URL https://www.politie.nl/themas/politietaken.html{#}alinea-title-wat-zijn-de-kerntaken-van-de-politie.

[42] Matthew Prince. Empty DDoS Threats: Meet the Armada Collective, 2016. URL https://blog.cloudflare.com/empty-ddos-threats-meet-the-armada-collective/.

[43] Christian Rossow. Amplification Hell: Revisiting Network Protocols for DDoS Abuse. In *NDSS*, 2014.

[44] José Jair Santanna and Anna Sperotto. Characterizing and Mitigating the DDoS-as-a-Service Phenomenon. pages 74–78. Springer Berlin Heidelberg, 2014. doi: 10.1007/978-3-662-43862-6_10. URL http://link.springer.com/10.1007/978-3-662-43862-6{_}10.

[45] Jose Jair Santanna, Roland van Rijswijk-Deij, Rick Hofstede, Anna Sperotto, Mark Wierbosch, Lisandro Zambenedetti Granville, and Aiko Pras. Booters — An analysis of DDoS-as-a-service attacks. In *2015 IFIP/IEEE Int. Symp. Integr. Netw. Manag.*, pages 243–251. IEEE, may 2015. ISBN 978-1-4799-8241-7. doi: 10.1109/INM.2015.7140298. URL http://ieeexplore.ieee.org/document/7140298/.

[46] Aditya K Sood and Richard J Enbody. Crimeware-as-a-service—a survey of commoditized crimeware in the underground market. *Int. J. Crit. Infrastruct. Prot.*, 6(1):28–38, 2013. ISSN 1874-5482.

[47] Spiegel. Computerangriffe auf Streitkräfte: Bundeswehr zählte 71 Millionen Cyberattacken 2015, 2016. URL http://www.spiegel.de/netzwelt/netzpolitik/bundeswehr-71-millionen-cyberattacken-in-2015-a-1082536.html.

[48] Usman Tariq, ManPyo Hong, and Kyung-suk Lhee. A comprehensive categorization of DDoS attack and DDoS defense techniques. In *Adv. Data Min. Appl.*, pages 1025–1036. Springer, 2006. ISBN 3540370250.

[49] THTC. Intro to THTC, 2016.

[50] Jochem Tolsma. Aangiftebereidheid: Welke overwegingen spelen een rol bij de beslissing om wel of niet aangifte te doen? *Proces-verbaal, aangifte en Forens. Onderz. Cah. Politiestud.*, 21: 11–32, 2011.

[51] Dutch Eurydice Unit. The Education System in the Netherlands 2003. In *Forty-seventh Sess. Int. Conf. Educ.*, The Hague, 2004. Netherlands Ministry of Education, Culture and Science.

[52] Peter van Ammelrooy. Hackers gijzelen ook Nederlandse computers voor losgeld, 2016. URL http://www.volkskrant.nl/tech/hackers-gijzelen-ook-nederlandse-computers-voor-losgeld{~}a4264574/.

[53] S. van der Weijer and W. Bernasco. WODC | 2674 - Aangiftebereidheid. Technical report, NSCR, Amsterdam, 2016. URL https://www.wodc.nl/onderzoeksdatabase/2674-aangiftebereidheid.aspx?nav=ra{&}l=veiligheid{_}en{_}preventie{&}l=sociale{_}veiligheid.

[54] Michel J G Van Eeten and Johannes M Bauer. Economics of malware: Security decisions, incentives and externalities. Technical report, 2008.

[55] Marissa van Loon. Vijf jongeren opgepakt voor DDoS-aanvallen op Ziggo en KPN - NRC, 2015. URL https://www.nrc.nl/nieuws/2015/10/07/vijf-jongeren-opgepakt-voor-ddos-aanvallen-op-ziggo-en-kpn-a1412575.

[56] Verisign. VERISIGN DISTRIBUTED DENIAL OF SERVICE TRENDS REPORT VOLUME 2, ISSUE 3 – 3RD QUARTER 2015. Technical report, Verisign, 2015. URL http://www.ddos.it/wp-content/uploads/2015/11/Verisign-ddos-trends-reports-Q32015.pdf.

[57] Verisign. Types of DDoS Attacks. Technical report, Verisign, 2015. URL https://www.verisign.com/en{_}US/security-services/ddos-protection/types-of-ddos-attacks/index.xhtml.

[58] Joost Visser and Pieter Jan 't Hoen. BVH Software Risk Assessment Rapport t.b.v. vts Politie Nederland. 2008.

[59] World-bank. Internet users as percentage of population, 2015. URL http://www.google.com/publicdata/explore?ds=d5bncppjof8f9{_}{&}ctype=l{&}strail=false{&}bcs=d{&}nselm=h{&}met{_}y=it{_}net{_}user{_}p2{&}scale{_}y=lin{&}ind{_}y=false{&}rdim=region{&}ifdim=region{&}tdim=true{&}tstart=972511200000{&}tend=1414274400000{&}ind=false.

[60] Y. Xie and S. Z. Yu. Monitoring the Application-Layer DDoS Attacks for Popular Websites. *IEEE/ACM Trans. Netw.*, 17(1):15–25, feb 2009. ISSN 1063-6692. doi: 10.1109/TNET.2008.925628. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4548145.

[61] H. Zimmermann. OSI Reference Model–The ISO Model of Architecture for Open Systems Interconnection. *IEEE Trans. Commun.*, 28(4):425–432, apr 1980. ISSN 0096-2244. doi: 10.1109/TCOM.1980.1094702. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1094702.