Machine learning for network-based malware detection

Stevanovic, Matija

DOI (link to publication from Publisher):
10.5278/vbn.phd.engsci.00088

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

MACHINE LEARNING FOR NETWORK-BASED MALWARE DETECTION

BY
MATIJA STEVANOVIC

DISSERTATION SUBMITTED 2016

AALBORG UNIVERSITY
DENMARK

Machine learning for network-based malware detection

Ph.D. Thesis

Matija Stevanovic

Thesis submitted on January 29, 2016

Dissertation submitted: January 29, 2016

PhD supervisor: Assoc. Prof. Jens Myrup Pedersen, Aalborg University

PhD committee: Associate Professor Reza Tadayoni (chairman)
Department of Electronic Systems
Aalborg University

Reader Kevin Curran
Computer Science Research Institute
University of Ulster

Director Cyril Onwubiko
Cyber Security and Information Assurance (IA)
At Research Series Limited
DWP

PhD Series: Faculty of Engineering and Science, Aalborg University

ISSN (online): 2246-1248
ISBN (online): 978-87-7112-490-3

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2016

Abstract

Malware has evolved over the past decades adding novel propagation vectors, robust resilience techniques as well as diverse and increasingly advanced attack strategies. The latest incarnation of malware is the notorious bot malware that provide the attacker with the ability to remotely control compromised machines thus making them a part of networks of compromised machines also known as botnets. Bot malware rely on the Internet for propagation, communicating with the remote attacker and implementing diverse malicious activities. As network traffic activity is one of the main traits of malware and botnet operation, traffic analysis is often seen as one of the key means of identifying compromised machines within the network.

This thesis explores how can network traffic analysis be used for accurate and efficient detection of malware network activities. The thesis focuses on botnet detection by exploring the possibilities of developing a novel collaborative approach to botnet protection that would utilize insights from various detection sensors. Furthermore, we focus on network-based detection aspects of the collaborative framework by devising novel detection approaches that are aimed at identifying malware network activity at different points in the network and based on different, mutually complementary, principles of traffic analysis. The detection approaches proposed by the thesis rely on machine learning algorithms (MLAs) for identifying malicious traffic as a set of algorithms capable of identifying patterns of malicious network traffic in automated and resource-efficient manner. The proposed approaches are developed in order to cover different aspects of malware network activity and thus be suitable candidates for a future collaborative botnet protection system. We evaluated the proposed detection methods through extensive set of experiments in order to assess the capabilities of different traffic analysis scenarios and machine learning algorithms to facilitate accurate and time-efficient detection. The experimental evaluation was performed using malicious and benign traffic traces originating from honeypots and malware testing environments as well as traffic traces from large-scale ISP networks. Based on the evaluation, the proposed traffic analysis methods promise accurate and efficient identification of malicious network traffic, thus being

promising candidates for future operational deployment. Furthermore, in addition to novel machine learning-based detection approaches the thesis provides an overview of some of the biggest challenges of using MLAs for identifying malicious network activities. The challenge specially addressed by the thesis is the "ground truth" problem, where we proposed a novel labeling approach for obtaining the ground truth on agile DNS traffic. The novel labeling approach has proved to provide reliable and time-efficient labeling by discovering much wider set of malicious domain names in comparison to conventional labeling solutions. Finally, the thesis outlines the opportunities for future work on realizing more robust and effective detection solutions.

Resumé

Malware har udviklet sig gennem de sidste årtier med nye spredningsvektorer, robuste teknikker til at modstå bekæmpelse, samt alsidige og stadigt mere avancerede angrebsstrategier. De sidste generationer af malware er de notoriske bot malware, der giver angriberen mulighed for at fjernstyre angrebne maskiner, og således gøre dem til en del af et netværk af inficerede maskiner, såkaldte botnet. Bot malware bruger Internettet til spredning, kommunikation med angriberen, og endeligt til at implementere diverse ondartede aktiviteter. Den netværkstrafik der genereres i forbindelse med disse aktiviteter udgør et væsentligt træk, og bliver af mange set som et af de vigtigste redskaber til at identificere inficerede maskiner på et netværk.

Denne afhandling undersøger hvordan netværkstrafikanalyse kan bruges til præcis og effektiv detektion af ondsindede netværksaktiviteter. Afhandlingen fokuserer på detektion af botnets ved at udforske mulighederne for at udvikle en ny kollaborativ tilgang der gør brug af informationer fra forskellige typer af sensorer. Derudover fokuseres der på de netværksbaserede aspekter ved at udvikle nye metoder til detektion med henblik på at identificere ondsindet netværksaktivitet I forskellige punkter på netværket. Disse er baseret på forskellige tilgange til trafikanalyse, der gensidigt supplerer hinanden. Metoderne til detektion der foreslås i afhandlingen baserer sig på maskinlæringsalgoritmer (MLA) til at identificere ondsindet trafik, og implementeres ved hjælpe af en række algoritmer der er i stand til at identificere mønstre af ondsindet trafik på en automatiseret og ressource-effektiv måde. De foreslåede tilgange er udviklet med henblik på at afdække forskellige aspekter af ondsindet netværksaktivitet, og dermed være egnede kandidater til at indgå i et fremtidigt kollaborativt system til beskyttelse mod botnets. Vi har analyseret og evalueret de foreslåede detektionsmetoder gennem omfattende eksperimenter med henblik på at undersøge hvordan de maskinlæringsalgoritmer og forskellige scenarier til trafikanalyse kan understøtte præcis og hurtig detektion. Den eksperimentelle evaluering blev udført ved hjælp af både reel og ondsindet netværkstrafik, opsamlet ved hjælp af såvel honeypots og testmiljøer for malware som større ISP-netværk. Ud fra evalueringerne blev det konkluderet at de foreslåede metoder til trafikanalyse

er lovende i forhold til at kunne bruges til præcis og effektiv identifikation af ondsindet netværkstrafik, og dermed også lovende i forhold til at kunne anvendes i operationelle miljøer I fremtiden. Udover nye maskinlæringsbaserede tilgange til detektion giver afhandlingen et overblik over nogle af de største udfordringer ved at bruge MLA til at identificere ondsindet netværksaktivitet. Især behandles "ground truth" udfordringen, og i forbindelse hermed foreslås en ny fremgangsmåde til at finde og mærke trafik baseret på agil DNS-trafik. Det viser sig at denne nye tilgang giver både pålidelig og tidseffektiv mærkning idet den opdager langt flere ondsindede domænenavne end konventionelle mærkningsmetoder. Afslutningsvis kommer afhandlingen med et overblik over muligheder for fremtidigt arbejde på vej mod mere robuste og effektive detektionsløsninger.

Acknowledgments

The work presented in this thesis was made possible by many people. First of all, I would like to thank my supervisor Jens Myrup Pedersen for giving me the opportunity to do the PhD project at Aalborg University and for providing me with guidance and valuable feedback throughout my PhD studies.

Furthermore, I would like to thank my colleagues from Wireless Communication Networks Section and former Networking and Security Section for support and great scientific inputs over the years. A special thanks goes to Dorthe Sparre for assisting me with many practical and organizational tasks throughout the PhD process.

I would also like to thank FTW (Forschungszentrum Telekommunikation Wien), Vienna for having me as a visiting researcher during my PhD stay abroad. FTW in an excellent research environment that significantly contributed to the knowledge needed for creating this thesis. The special thanks goes to Alessandro D'Alconzo, Stefan Ruehrup and Andreas Berger from FTW. I have learned a lot from them and through their advices and guidance I have become a better researcher in many regards.

Special thanks goes to Bredbånd Nord for providing DNS traffic data sets used for the development and the evaluation of the proposed detection methods. This thesis would not be possible without the data sets they so kindly shared with us. I would also like to thank Dan Sandberg and Peter Isager for assisting in obtaining the data sets and contributing to discussions on the use of the proposed detection methods in operational networks.

Finally, I also would like to thank my wife Nevena for motivating me to go on this journey and for giving me invaluable support and encouragement along the way. Thank you for believing in me. Also, I would like to thank my parents and my family for their support during my education.

Aalborg, January 29, 2016
Matija Stevanovic

Contents

Thesis Details

Thesis Title: Machine learning for network-based malware detection
PhD Student: Matija Stevanovic
Supervisor: Jens Myrup Pedersen, Associate Professor, Aalborg University, Denmark

Thesis organization

The thesis is realized following the collection of papers thesis model, thus consisting of an introductory overview and a number of appended publications. The thesis is organized as follows. Part I of the thesis presents the problem addressed by the thesis and the research questions. This part also summarizes the contributions of the appended papers and the thesis as a whole. Part II attaches the publications that carry the main contributions of the thesis.

List of Appended Papers

This thesis is based on the work presented in the following 6 papers:

Paper I Matija Stevanovic, Kasper Revsbech, Jens Myrup Pedersen, Sharp Robin and Christian Damsgaard Jensen. "A collaborative approach to botnet protection." In the proceedings of the International Cross-Domain Conference and Workshop on Availability, Reliability, and Security, CD-ARES 2012, August 2012. Lecture Notes in Computer Science Vol. 7465, Springer, 2012. p. 624-638. DOI: 10.1007/978-3-642-32498-7_47.

Paper II Matija Stevanovic and Jens Myrup Pedersen. "On the Use of Machine Learning for Identifying Botnet Network Traffic." The paper will appear in a special issue of the Journal of Cyber Security and

Mobility as the proceedings of the 8th International CMI Conference on Cyber Security, Cyber Crime, Privacy and Trust, November 2015.

Paper III Matija Stevanovic, Jens Myrup Pedersen, Alessandro D'Alconzo, Stefan Ruehrup and Andreas Berger. "On the ground truth problem of malicious DNS traffic analysis." Computers & Security, Vol. 55, 2015, p. 142-158. DOI: 10.1016/j.cose.2015.09.004

Paper IV Matija Stevanovic and Jens Myrup Pedersen. "An efficient flow-based botnet detection using supervised machine learning." In the proceedings of the International Conference on Computing, Networking and Communications (ICNC), February 2014. IEEE Press, 2014. p. 797-801. DOI: 10.1109/ICCNC.2014.6785439.

Paper V Matija Stevanovic and Jens Myrup Pedersen. "An analysis of network traffic classification for botnet detection." In the proceedings of the International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), June 2015. IEEE, 2015. DOI: 10.1109/CyberSA.2015.7361120.

Paper VI Matija Stevanovic, Jens Myrup Pedersen, Alessandro D'Alconzo and Stefan Ruehrup. "A method for identifying compromised clients based on DNS traffic analysis." The paper is submitted to the International Journal of Information Security by Springer, December 2015.

Comments on My Participation

I am responsible for the most of the written material, the implementation of the proposed methods and for carrying out all the experiments with the exception of the cases described below. My supervisor and collaborators contributed by participation in discussions about the scope of the papers, methods used in the papers, and by providing comments on the papers throughout the writing process.

- Paper I was realized in collaboration with Kasper Revsbech, Jens Myrup Pedersen, Robin Sharp and Christian Damsgaard Jensen. I am responsible for the most of the written material and for defining the concepts of the presented methodology. Kasper Revsbech has contributed to defining the network monitoring part of the botnet protection approach. Robin Sharp and Christian Damsgaard Jensen have contributed to the presented approach with discussions on the possibilities of correlating findings from diverse information sources considering their trust and

reliability. Finally, Jens Myrup Pedersen has contributed through discussions regarding network traffic analysis.

- Paper II, Paper IV and Paper V were realized in collaboration with Jens Myrup Pedersen. I am responsible for the most of the written material, the implementation of the methods and for carrying out all the experiments. Jens Myrup Pedersen contributed by participating in discussions about the scope of the papers, methods used in the papers, and by providing comments on the papers throughout the writing process.

- Paper III was realized in collaboration with Jens Myrup Pedersen, Alessandro D'Alconzo, Stefan Ruehrup and Andreas Berger. I am responsible for the most of the written material, the implementation of the proposed method and for carrying out all experiments. The work done in this paper was built on top of Andreas Berger previous work on characterizing the agility of DNS traffic. Andreas Berger has contributed by participating in discussions regarding the proposed DNS labeling methodology and software solution that was used as the base for the presented work. Alessandro D'Alconzo, Stefan Ruehrup and Jens Myrup Pedersen have contributed through discussions on the proposed method, and by providing comments on the paper throughout the writing process.

- Paper VI was realized in collaboration with Jens Myrup Pedersen, Alessandro D'Alconzo and Stefan Ruehrup. I am responsible for the most of the written material, the implementation of the proposed method and for carrying out all experiments. The co-authors contributed through discussions about the scope of the paper, the method presented in the paper, and by providing comments on the paper throughout the writing process.

Other Papers

Apart from the papers included in this thesis, I am the first author of the following technical report:

- Matija Stevanovic and Jens Myrup Pedersen. "Machine learning for identifying botnet network traffic." Technical report, Department of Electronic Systems, Aalborg University, pages 1–28, April 2013. Accessible: `http://vbn.aau.dk/files/75720938/paper.pdf`.

The report was not included due to its excessive length. However, it should be noted that the Introduction part of the thesis is based on findings and

conclusions presented in this paper.

Furthermore, during my PhD studies I have co-authored following publications in regards to malware analysis and detection:

- Jens Myrup Pedersen and Matija Stevanovic. "AAU-Star and AAU Honeyjar: Malware Analysis Platforms Developed by Students." In the 7th International Conference on Image Processing and Communications (IP&C 2015), Image Processing and Communications Challenges 7, Springer, 2015. p. 281-287 (Advances in Intelligent Systems and Computing, Vol. 389). DOI:10.1007/978-3-319-23814-2_32.
- Radu-Stefan Pirscoveanu, Steven Strandlund Hansen, Thor Mark Tampus Larsen, Matija Stevanovic, Jens Myrup Pedersen and Alexandre Czech. "Analysis of Malware behavior: Type classification using machine learning." In the proceedings of the International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), June 2015. IEEE, 2015. DOI: 10.1109/CyberSA.2015.7166115.
- Steven Strandlund Hansen, Thor Mark Tampus Larsen, Matija Stevanovic and Jens Myrup Pedersen. "An Approach for Detection and Family Classification of Malware Based on Behavioral Analysis." The paper is will appear in the proceedings of the International Conference on Computing, Networking and Communications (ICNC), February 2016.

The first paper presents two malware analysis platforms developed through a series of student projects at Aalborg University. The student projects were supervised by Jens Myrup Pedersen and me, and we were actively involved in their design and implementation. The second and the third paper address the identification of malware at client machines and malware classification to types and families based on the behavioral analysis.

Declaration

This thesis has been submitted for assessment in partial fulfillment of the PhD degree. The thesis is based on the submitted or published scientific papers which are listed above. Parts of the papers are used directly or indirectly in the Introduction part of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

Part I

Introduction

Introduction

The growing reliance on the Internet, the advances in computing technology and the proliferation of affordable computing units have contributed to a new "connected" era of human civilization. However, the new connected world introduces numerous challenges to the protection of the privacy and the security of users and user's data.

During the last two decades, the use of the Internet and Internet-based applications has experienced a tremendous expansion to the point at which they have become an integral part of our lives, supporting a wide range of services, such as banking, commerce, healthcare, public administration and education. The number of the Internet users worldwide have surpassed 3 billion in 2015 corresponding to the penetration rate of over 40% [1]. Furthermore, the technology advances have led to the proliferation of affordable computing units in forms of either conventional personal computers or handheld devices such as smartphones and tablets. Figures for 2014 show over 2.6 billion smartphone subscriptions globally with a steady growth trends [2]. Finally, Internet of Things (IoT) together with initiatives such as Smart Grids and Smart Cities have contributed to networking of even wider set of household appliances equipping them with often capable computing units and networking ability via multiple communication technologies. The latest reports claim that the number of IoT devices in 2015 was 13.4 billion corresponding to over 2 Internet connected units on average per person in the world [3].

Although offering a number of advantages, the new connected world represents an attractive playing field for cyber criminals. Criminals rely on Internet for implementing various illegal activities in anonymous and hardly traceable manner. As over 40% of world's population uses the Internet the reach of potential attacks is immense. Furthermore, a large number of connected computational units represents a great asset in terms of both the cumulative computational power and the available network bandwidth. Attackers often try to compromise these machines and use them in diverse malicious contexts ranging from mining of digital currencies to launching powerful Distributed Denial of Service (DDoS) attacks. Cyber criminals rely on malicious software also known as *malware* for misusing the Internet connected

computers. Modern malware relies on Internet for implementing its malicious agenda and facilitating communication infrastructure through which attackers can control compromised computers.

This thesis tackles the malware detection problem from the perspective of network traffic analysis. The work presented in this thesis proposes novel methods that aim at providing efficient and accurate malware detection based on network traffic analysis. The thesis focuses on *botnets* as networks of computers compromised with malware. We have devised several traffic analysis strategies aimed at identifying botnets at different points in the network and based on different, mutually complementary, principles of traffic analysis. The proposed approaches are developed in order to cover different aspects of malware network activity and thus be suitable candidates for a future collaborative botnet protection system. For the realization of the traffic analysis we rely on Machine Learning Algorithms (MLAs) as a set of algorithms capable of identifying the patterns of malicious network traffic in automated and resource-efficient manner. Furthermore, the thesis brings an overview of both capabilities and some of the biggest challenges of using MLAs for identifying botnets, such as the "ground truth" problem. The proposed methods have been evaluated using traffic traces captured by honeypots and malware testing environments as well as traces from ISP networks. As a result, the proposed detection methods promise accurate and efficient identification of malicious network traffic, thus being good candidates for the use in a future collaborative botnet protection systems.

This chapter has a goal of outlining the work done during the PhD project and summarizing its contributions. The chapter is based on findings and conclusions of our technical report on the use of machine learning for botnet detection [4]. This chapter is organized as follows. Section 1 presents malware threat in more details by elaborating on the malware phenomena, current trends and characteristics of modern malware. The section focuses on botnets as the latest malware incarnation. Section 2 presents the main motivation for network-based detection of malware and overall concepts behind it. This section emphasizes machine learning-based approaches as one of the most promising classes of detection methods. Section 3 defines the problem statement addressed by the thesis and four research questions covered by the work done. The four research questions cover some of the most prominent topics in the field of network-based malware detection. Section 4 presents the state of the art on network-based malware detection focusing on machine-learning based approaches. This section also outlines opportunities for future work out of which several have been the focus of the work presented in the attached papers. Section 5 presents the main contributions of the thesis and appended paper. Finally, Section 6 summarizes the conclusions of the thesis. This section also discusses our findings and outlines the opportunities for future work.

1 Malware Threat

In this section we present the treat of malware by presenting characteristics of modern malware and the current trends. Furthermore, we focus on botnets as one of the latest malware incarnations.

Malware represents the main carrier of malicious activities on the Internet. Malware implements a variety of malicious and illegal activities that disrupt the use of compromised computers and jeopardize the security of the end users. In parallel with the development and expansion of Internet-based services, malware has evolved by improving the mechanisms of propagation, malicious activities, and resilience to take down efforts. Modern malware targets a variety of client platforms, compromising millions of computers worldwide, deploying sophisticated attack campaigns and causing great financial damages to both industry and governments.

Modern malware covers a variety of platforms from mobile operating systems [5] to industrial control systems [6]. Although often perceived as a problem exclusively tied to Windows platform malware has spread out to other operating system as well such as Apple Mac OS and Linux [7]. One of the latest trends is the shift towards mobile operating systems due to the popularity of smartphones and their use for different services such as e-banking, online shopping, etc. Symantec reports that over 1 million distinct mobile malware samples were observed in 2014 where the majority of them were targeting Android operating system [8].

Estimations of the number of novel malware indicate that in 2015 over 390,000 new malware samples were observed daily [9]. Furthermore, the number of new malware variants has seen increase of 26% in 2014 reporting staggering 317 million of new malware variants [8]. The number of infected machines worldwide has been increasing over the last 10 years with the latest estimation from 2014 that indicates that 14% of all residential and 0.68% of mobile Internet users are compromised with some kind of malware [10].

Malware is used to implement a variety of malicious activities such as sending SPAM messages, deploying DDoS attacks, information theft, mining digital currencies, ransomware, etc. All of this activities cause a significant financial damage to individuals, companies and governments. Some reports estimate that the annual global cyber-crime costs are more than 300 billion US dollars [11]. The majority of these costs is directly or indirectly related to malware. Furthermore, the recent study by Ponemon Institute outlines the cost of malware containment commercial companies are faced with [12]. The report indicates the great financial expenses of effectively protecting company infrastructure from malware threat.

Based on the presented, malware is rightfully regarded as one of the biggest cyber security threats today. As such malware requires efficient and

effective neutralization techniques. Malware detection represents a key element of any successful neutralization techniques. In the following we put more light on botnets as the latest incarnation of malware and opportunities for their detection.

1.1 Botnets - the connected malware

One of the most capable types of malware is the notorious bot malware. Bot malware represents a program that allows the creator to control infected computers remotely. This class of malware is commonly considered as one of the most advanced malware classes as it incorporates sophisticated propagation, resilience and attack techniques used by other malware classes [13, 14]. The main advantage in comparison to other malware types and the main trait of bot malware is the ability to facilitate remote control of compromised clients by an attacker through a specially deployed *Command and Control (C&C)* communication channel [15–17]. Once loaded onto a client machine the bot malware compromises the vulnerable machine and, using the C&C channel, puts it under the remote control by the attacker. The attacker is popularly referred to as the *botmaster*, while compromised hosts are known as *bots*. Using a deployed C&C channel the botmaster can remotely control the behavior of bots and transfer the data to and from the compromised machine. This way the attacker can make the operation of bots more flexible and consequently more effective in implementing their malicious agenda.

A botnet is a usually large collection of computers that are infected with the specific bot malware. Controlled and coordinated by the botmaster, botnets represent a collaborative and highly distributed platform for the implementation of a wide range of malicious and illegal activities. Botnets may range in size from a couple of hundred to several million bots [18, 19]. In addition, botnets can span over home, corporate and educational networks, while covering numerous autonomous systems operated by different Internet Service Providers (ISPs). Since botnets include such a large number of bots, they often have enormous bandwidth and computational power at their disposal. Furthermore, botnets are capable of implementing diverse malicious activities such as: information theft, spam distribution, DDoS attacks, malware distribution, click fraud, mining digital currencies, etc.

Botnet threat - real-world examples

The threat of botnets is best illustrated by the examples of botnets observed in the wild over the past decade. Some of the most notorious botnets ever encountered [20] are:

Storm - Storm botnet was one of the first wide-scale botnets captured in the wild. The Storm botnet was first detected in 2007 and it is notable

for being one of the first peer-to-peer botnets. Estimates of Storm's size ranged anywhere from 250,000 to 50 million compromised computers. This botnet was known for enabling share price fraud and identity theft but portions of it were often leased for other malicious activities as well. Storm was partially shut down in 2008.

Conficker - Conficker represents one of the widest spread malware of the last decade. At its peak in 2009, the Conficker worm have infected 15 million computers, but the total number of machines under its botnet control was between 3 and 4 million. This makes the Conficker one of the largest botnets ever.

Cutwail - Cutwail represent one of the biggest spam botnets to date. At its peak in 2009 the botnet controlled up to 2 million compromised computers, sending 74 billion spam emails per day which is equivalent to nearly a million e-mails per minute. This made up 46.5% of the global spam volume at the time. In 2010, two-thirds of Cutwails's control servers were disabled.

ZeroAccess - ZeroAccess botnet is one of the more recent botnets to be detected. The size estimates indicate that it was controlling over 1.9 million compromised computers around the world. This botnet is known for implementing click fraud and bitcoin mining. Due to the latter, this botnet was reported to be consuming enough energy to power 111,000 homes every single day from all its infected computers.

Windigo - Windigo botnet was discovered in 2014 after operating undetected for three years. In this time, it had infected 10,000 Linux servers enabling it to send 35 million spam emails a day. The threat posed by Windigo is ongoing and as more than 60% of all web servers use Linux servers the potential risk is huge.

Botnet operational life-cycle

Botnet operation can be described through the analysis of *botnet life-cycle* i.e. the set of botnet operational phases [14, 21]. The botnet life-cycle is commonly generalized as consisting of three distinct phases: *the infection phase*, *the C&C communication phase* and *the attack phase*.

The infection phase is the first phase of the botnet life-cycle in which vulnerable computers are compromised with bot malware, thus becoming a member of a specific botnet. This phase is commonly divided into two sub-phases i.e. initial infection and secondary infection. During the initial infection sub-phase computers are infected with a malicious piece of software also known as the "dropper". The initial infection can be realized in different ways, for instance, through the unwanted download of malware

from malicious websites, through the download of infected files attached to email messages, by propagation of malware from infected removable disks, etc. The dropper assists in obtaining the bot malware binary. Upon successful initial infection, the dropper downloads the malware binary over the network and installs it on the vulnerable machine as a part of the secondary infection sub-phase. Bot malware binaries can be downloaded using diverse communication protocols, such as FTP (File Transfer Protocol), HTTP/HTTPS (Hypertext Transfer Protocol) or P2P (Peer-to-Peer) protocols.

The second phase of the life-cycle is the C&C communication phase that covers communication between compromised computers and malicious network infrastructure. This phase covers several communication actions such as: initial connection attempts to the C&C infrastructure upon successful infection phase, connection attempts by the bot after reboot of the compromised machine, periodical connection attempts in order to report the status of the infected machine and the connection attempts initiated by the attacker in order to update malware code or propagate instructions to bots. The communication channel established between bots and C&C servers i.e. C&C channel can be implemented in different ways.

The third phase of botnet life-cycle is the attack phase that includes bot operation aimed at implementing attackers' malicious agenda. This phase includes malicious and illegal activities outlined above but also malware propagation mechanism such as scanning for vulnerable computers. The second and the third phase are functionally linked so they are usually altering one after another, once a vulnerable computer is successfully infected.

Infection vectors

Modern malware relies on a number of *infection vectors* i.e. methods used by the perpetrators for propagating the malware to other machines or networks within the initial infection operational phase. Initial infection is realized using a variety of infection vectors, such as:

- Trojan horse - represent a propagation method in which the user is tricked into installing the malicious software without understanding its true nature.
- Network scanning - represents a common method to exploit vulnerable network services of client machines. If client machines provide a vulnerable service over the network, it can be used by an attacker to attack the system by network scanning for vulnerabilities.
- Drive-by-download - represents a method that targets user's web browser by exploiting vulnerabilities in the browser or browser plug-ins. In this case malware is able to fetch code from the malicious web

sites using connections which was initiated by the user himself and then execute it on the victim's machine.

The outlined infection vectors heavily rely on social engineering in order to lure the user into performing a set of actions that could lead to successful infection. Usually, a user is targeted through spam e-mails or social network campaigns that commonly involve clicking on URLs, downloading malicious files and in some cases installing malicious programs. In order to achieve this, attackers often misuse governmental institutions' or companies' insignia thus trying to recreate look and feel of legitimate e-mails and web pages. Finally, human impact should not be forgotten, as in many cases it is crucial for successful infection. Human impact commonly reefers to the susceptibility to social engineering and phishing scams as well as the lack of security awareness and knowledge about sound security practices.

C&C communication

The C&C channel is one of the defining traits of bot malware and the main carrier of botnet functionality. The C&C channel facilitates remote coordination of compromised computers, and introduces a level of flexibility in botnet operations by offering the ability to change and update malicious code. Attackers rely on several control mechanisms in terms of communication protocol and network architecture for deploying the C&C channel [15–17, 22–24]. Based on the topology of the C&C network, botnets are commonly classified as centralized, decentralized and hybrid botnets.

Centralized botnets have centralized C&C network architecture, where bots contact one or several C&C servers owned by the botmaster. Centralized C&C channels are commonly realized using IRC (Internet Relay Chat) and HTTP/HTTPS protocols. IRC-based botnets are created by deploying IRC servers or by using IRC servers in public IRC networks. In this case, the botmaster specifies a chat channel on a IRC server to which bots connect to in order to receive commands. HTTP-based botnets rely on HTTP/HTTPS protocols to transfer C&C messages. In contrast to IRC-based botnets, bots in HTTP-based botnets contact a web-based C&C server notifying their existence with system-identifying information via HTTP/HTTPS requests. As a response, the malicious server sends back commands or updates via counterpart response messages. IRC- and HTTP-based botnets are characterized with low latency of command messages and they are easy to deploy and manage. For this reason, they have been widely used. However, the main drawback of centralized botnets is that they are vulnerable take down due to the single point of failure. That is, once the C&C servers have been identified and disabled, the entire botnet could be taken down.

Decentralized botnets represent a class of botnets developed with the goal of being more resilient to take down efforts. Botnets with decentralized C&C

infrastructure have adopted P2P communication protocols as the mean of communicating within a botnet [17, 22]. This implies that bots belonging to the P2P botnet form an overlay network and that the botmaster can use any of the bots (P2P nodes) to distribute commands to other peers or to collect information from them. P2P botnets are realized either by using some of the existing P2P transfer protocols, such as Kademlia [25], Bittorent [26] and Overnet [27], or by custom P2P protocols. While more complex and perhaps more costly to manage and operate compared to centralized botnets, P2P botnets offer higher resiliency, since even if the significant portion of the botnet is taken down the remaining bots may still be able to communicate with each other and with the botmaster. However, P2P botnets are commonly characterized with high latency and low reliability of C&C communication, which severely limits the overall efficiency of orchestrating attacks.

Some of the recent botnets [28] have adopted more advanced hybrid network architectures, that combine the principles of centralized and decentralized botnets. The hybrid botnets use advanced hybrid P2P communication protocols in order to combine the resiliency of P2P botnets with the low latency of centralized botnets. The hybrid botnet architecture has been investigated by several authors [23, 24] suggesting that in order to provide both resilience and low latency of communication hybrid botnets should be realized as networks in which bots are interconnected in P2P fashion and organized in two distinct groups i.e. the group of proxy bots and the group of working bots. Working bots would implement the malicious agenda while proxy bots would relay C&C messages between bots and the botmaster. Working bots would periodically connect to the proxy bots in order to receive commands. Based on the work presented in [23, 24] this topology provides higher resilience to take down efforts and improvements in the latency of C&C messages comparing to traditional P2P botnets.

Malicious activities

As already partly illustrated malware can facilitate a variety of sophisticated malicious and illegal activities. Some of the most prominent include identity theft, information stealing, pay-per-install (PPI), click fraud, adware, malware distribution, spam distribution, DDoS attacks, mining digital currencies and the attacks targeted at industrial control systems and critical infrastructure.

The presented attack strategies produce more or less distinguishable behavior at both client- and network-level. The attack strategies rely on the network communication to different degrees. Identity theft and Information stealing involve transferring sensitive client data over the network. A number of recent data breaches where realized using sophisticated malware that was able to steal an enormous amount of data over the network [29]. As an example, the hacker group responsible for the Sony Pictures hacking case [30]

has claimed that they stole over 100 TB of sensitive data, from which 200GB was publicly released [31]. Spam distribution without any doubt represents one of the malicious activities with the largest network footprint. Report by Cisco SecurityWorks [32] from 2008 indicates that top botnets are capable of sending over 100 billion spam e-mails per day. Some of the most famous spamming botnets such as Grum was responsible for 26% of world's spam email traffic in 2012 and during its peak it could send 39.6 billion spam messages daily [20]. Finally, DDoS attacks pose a serious challenge to the existing Internet infrastructure. The DDoS attacks are usually implemented by botnets and their power is commonly measured in Gb/s. Arbor Network reports that the largest monitored and verified attack in 2014 was 325.05 Gb/s [33]. It should also be noted that the attacks have been growing in their power and sophistication over the last decade. Other attack strategies also include network activities such as downloading malware payload, network scanning for vulnerabilities, etc.

Resilience techniques

One of the primary goals of the malware operation is flying under the radar of detection and neutralization systems. Therefore, malware is equipped with a diversity of resilience techniques capable of providing the stealthiness and robustness of operation. Resilience techniques can be implemented both on client and network levels.

Client-level resilience techniques provide the robustness of malware to detection at the client machines and hinder both static and dynamic analysis of malicious code [34–36]. Some of the most prominent client-level resilience techniques are:

- Packing - represents the techniques of forming a binary file composed of compressed versions of executable files. The use of packing within the binary file hides parts of their content thus preventing the analysis.
- Polymorphic and metamorphic code - represent code obfuscation techniques that enable the malware code to mutate without changing the functions or the semantics of its payload. Hence, malware binaries of the same botnet are commonly different from each other. Using these techniques malware evades conventional detection solutions that depend on the signatures of malware binaries.
- Obfuscation of behavioral patterns - represent resilience techniques that obfuscate malware behavior at the client computer and thus hamper the system for monitoring client-level forensics [37].
- Rootkit ability - represents one of the most challenging resilience techniques deployed by malware at the client-level as it provides the mal-

ware with the ability to operate on kernel-level [38, 39]. Having the rootkit ability, the malware is able to defeat the majority of malware tracking systems implemented at client machines.

The client-level resilience techniques have proved to be very effective in avoiding modern detection systems, thus posing the great challenges to automatized detection at client-level. As a result, the majority of the contemporary detection methods focus on the analysis of network traffic produced by compromised computers [13, 14]. The following section presents more on the existing detection solutions.

Network-level resilience techniques have a goal of hampering detection of malware based on network traffic analysis by providing the secrecy and integrity of communication between compromised machines and the attacker, preserving the anonymity of the attacker, and facilitating the robustness of the C&C channel to take down efforts. Some of the most important means of providing secrecy of C&C communication are obfuscation of existing and development of custom communication protocols, as well as the encryption of the communication channel. Using these techniques, the security and the integrity of communication are preserved, thus efficiently defeating detection methods that rely on content of the traffic payloads for detection. Other commonly used techniques that provide resilience of malware network operation are DNS-based resilience techniques such as Fast-flux [40] and Domain-flux [41]. These techniques are characterized with the ability to dynamically change domain names and IP addresses associated with a particular service over time and they are commonly referred to as "agile" DNS traffic [42]. Agile DNS is widely abused by cyber criminals in order to avoid existing detection methods and take down techniques, thus providing the resilience of malicious services and C&C communication.

Fast-flux refers to the constant changing of IP address information related to a particular domain name [40]. Botnet operators abuse this ability to change IP address information associated with a host name by linking multiple IP addresses with a specific host name and rapidly changing the linked addresses. Fast-flux [40] is widely used by the botnets to hide phishing and malware delivery sites behind a dynamic network of compromised hosts acting as proxies. This way the anonymity of C&C servers and the attacker is protected, while providing more reliable malicious service.

Domain-flux is effectively the inverse of Fast-flux and refers to the constant changing and allocation of multiple domains to a single or multiple IP addresses. DGA (Domain Generation Algorithm) [41] is one of the most prominent Domain-flux techniques. DGA periodically generates a large number of domain names that can be used to reach C&C communication infrastructure. Bots using the DGA generate large number of pseudo-random domain names that are queried to determine addresses of the C&C servers.

The large number of domains generated each day makes their blacklisting difficult. Using DGA as a backup strategy higher resilience and robustness of C&C communication is achieved.

1.2 ZeroAccess botnet - the case study

ZeroAccess represents a sophisticated malware that targets Microsoft Windows operating systems. Computers compromised with this malware become a part of a notorious ZeroAccess botnet, which is one of the most advanced botnets observed during the last decade [43]. The ZeroAccess botnet was first detected in May 2011, while in 2012 at its peak it had an estimated size of over 1 million bots. This botnet is predominantly involved in click fraud and Bitcoin mining but it also has the capability of implementing a number of other attack campaigns. In December 2013 Microsoft led a coalition aimed at taking down ZeroAccess C&C network. The take down campaign was only partially effective as not all C&C servers were seized. As a result, the botnet was able to resurrect through its peer-to-peer command and control infrastructure. However, some of the latest studies show that the ZeroAccess botnet is only a shadow of former self, numbering 50.000 compromised machines globally [44].

The ZeroAccess botnet relies on a number of advanced propagation, resilience and attack techniques that are summarized below:

Infection vectors - ZeroAccess botnet utilizes different infection vectors where the most common is using exploit kits such as Blackhole [45], where the users are lured into vising the web page with a malicious script build in. This script tries to compromise the client by different software vulnerabilities and infecting it with a dropper program. The dropper program then downloads the ZeroAccess malware. Alternatively, the ZeroAccess malware is distributed through a number of trojan programs such as keygens, cracks and similar. Finally, the ZeroAccess malware is often downloaded by other malicious software as it has a very lucrative pay-per-install affiliate program.

C&C communication - This botnet employs sophisticated C&C infrastructure realized using custom P2P communication protocol. The C&C infrastructure has a hierarchical topology with number of super nodes that have a public IP address and working nodes behind the NAT. The P2P protocol relies on distributed list of peers between which UDP and TCP communication is realized. The ZeroAccess malware comes with hard-coded list of IP addresses and UDP and TCP port numbers. Furthermore, this malware relies on HTTP to report back to the attacker. Here the malware is using DGA as a resilience technique for discover-

ing the rendezvous point. Finally, all network communication used by the botnet is encrypted.

Attack campaigns - ZeroAccess botnet is predominantly implementing click fraud and Bitcoin mining as attack campaigns. These malicious campaigns are deployed by plug-ins programs downloaded by the ZeroAccess malware. The fact that the botnet is relying on malicious plug-ins indicates that it offers the possibility of easily extending its malicious capabilities. Each of the plug-ins have its own C&C and update mechanisms. These mechanisms are often related to the ZeroAccess C&C infrastructure indicating that the same people are behind the malicious plug-ins and the botnet itself.

Detection opportunities

As illustrated in the previous modern malware represents complex phenomena that manifests itself in different aspects and thus offering various opportunities for detection. Table 1 summarizes the characteristics of ZeroAccess botnet and the type of detection methods that could target each of the particular characteristics. Similarly, to any other malicious software ZeroAccess can be tackled both by client and network-level detection, targeting the behavior of malware at client machine and its network activity, respectively.

Table 1: Zero Access botnet - the analysis of detection opportunities.

Operation phase	Characteristics	Detection methods
Infection vectors	Exploit kits (with droppers)	Client-level, Network-level
	Trojan horses (keygens, cracks, games)	Client-level
	Downloaded by other malicious software	Client-level, Network-level
C&C communication	P2P network	Network-level
	Hard-coded UDP and TCP ports	Network-level
	Phone home via HTTP	Network-level
Attack campaigns	Click fraud	Network-level
	Bitcoin mining	Client-level, Network-level
	Crypto ransomware	Client-level
	Search engine redirection	Client-level, Network-level
	Sending SPAM	Network-level
	Arbitrary file download	Network-level
Resilience techniques	Rootkit ability	Static analysis
	Malware packer (dropper)	Static analysis
	Anti-debugging	Static analysis
	Encrypted traffic	Network-level
	DGA (phone home)	Network-level

Client-level detection has a number of challenges in the case of ZeroAccess

malware. First, certain variations of the malware have rootkit ability and operate on kernel-level. Furthermore, the dropper uses different resilience techniques such as code packing while ZeroAccess malware is equipment with anti-debugging techniques. These techniques significantly harden the use of static and dynamic code analysis. However, it should be noted that client-level analysis and especially static analysis could still provide very important information as the malware comes with hard coded list of IP addresses and TCP/UDP ports that are used for C&C communication.

Network-level detection could target different traffic characteristics and could be implemented at different parts of network. First, as the ZeroAccess botnet relies on a hard coded list of peer IP addresses and UDP and TCP ports it can be tackled using relatively trivial IP address and port blacklisting techniques as well as port number based classifiers. However, the malware has mechanisms for updating its infrastructure by periodically changing the peers list and the port numbers, thus limiting the use of above mentioned detection methods. Alternatively, the ZeroAccess network activity could be tackled by targeting different traits of botnet traffic, such as periodicity of network traffic, traffic distribution, etc. In addition, the malware could be targeted based on the principles of Deep Packet Inspection (DPI) but only with a limited impact as the botnet encrypts all C&C communication. Finally, as the botnet is relying on DGA it is possible to use DNS traffic analysis in order to identify pseudo-random domain names used by the botnet. The network-level detection can be realized both closer to client machines at local and enterprise networks as well as in the higher network tiers depending on the chosen principles of detection. The analysis of DNS traffic could be suiting for detection even in ISP networks while other approaches would preferably be implemented to implementation at local/enterprise networks.

Based on the presented we can conclude that different detection methods could be used in order to discover comprised machines in the case of the ZeroAccess botnet. The detection methods target different botnet characteristics and are often complementary. The following section examines different approaches to malware detection specially focusing on network-based detection and the use of machine learning for identifying malware network activities.

2 Network-based detection

This section provides an overview of the existing malware detection approaches. The section focuses on the use of network traffic analysis for the detection of botnets.

Conventional malware detection approaches are deployed at client computers targeting malware operating at compromised machines [46–52]. These methods are usually referred to as *client-based* detection approaches. The client-based detection approaches typically rely on the signatures of malicious software as in the case of conventional Anti-Virus (AV) solutions. In addition to matching signature of malicious binaries the client-based approaches can perform behavioral analysis by examining different client-level forensics, for instance, application and system logs, active processes, key-logs and the usage of system resources [46–52]. Finally, the client-based detection can also include examination of traffic visible on the computer's network interfaces in order to identify some of the signs of malicious network use [53–55].

As already indicated in the Section 1, modern malware relies on the Internet for different actions such as the propagation, the communication with the attacker and the deployment of different attack strategies. It could even be claimed that all modern malware relies on Internet communication in some phases of operation. Malware that would not produce any network activity would consequently severely limit their malicious potential. Such malware could only be used in tailored denial of service attacks towards specially selected targets. Network activity produced by malware is an important indicator of their operation and it is often seen as one of the most important resource for malware detection. As a result, many authors have turned their attention to *network-based* detection that relies on the analysis of network traffic for identifying compromised computers [14, 56]. Network-based detection approaches are deployed at an "edge" of the network (usually in routers or firewalls), providing detection of computers compromised by malware by analyzing network traffic. This class of methods identifies compromised computers by recognizing network traffic produced by them within all three phases of their life-cycle i.e. the infection phase, the C&C communication phase and the attack phase. These approaches are commonly referred to as intrusion detection systems (IDS) [14, 21].

In parallel with network- and client-based detection methods a novel class of collaborative detection methods has emerged [55, 57–59]. This class of methods concludes about the existence of malware on the basis of observations gathered at both client and network levels. The main hypothesis behind collaborative detection is that it is possible to provide robust and accurate detection by correlating findings of independent client- and network-based

detection systems. This class of detection approaches embraces the idea that there is no "silver bullet" in security as all of the detection solutions have their challenges and drawbacks and they could be avoided by malware if a sufficient effort is invested by the attacker. On the other side, the collaborative detection solutions that integrate the principles of diverse detection systems would require substantial effort in order to be avoided. In order to avoid such a collaborative detection method, the attacker would need to either significantly limit the attack potentials in order to be stealthier or make the malware operation more dynamic thus investing additional time and effort. The motivation for a collaborative detection can be found in the analysis of the ZeroAcess botnet presented in the previous section. The ZeroAccess is characterized with a number of resilience techniques that harden both client- and network-level detection. On the client-level there are rootkit ability, anti-debugging techniques and code packing, while on network-level traffic there are traffic encryption and the use of DGA as a relaying technique. This indicates that correlating findings from client- and network-based detection solutions could greatly contribute to effective detection.

2.1 Opportunities of network-based detection

There are several conceptual differences between client- and network-based detection because of which network-based detection is often seen as a more promising solution. Network-based detection is targeting the essential aspects of botnet and the functioning of modern malware, i.e. network traffic produced as the result of their operation. Network-based approaches assume that in order to implement their malicious functions botnets and malware in general have to exhibit certain network activity. They could make their operation stealthier by limiting the intensity of attack campaigns (sending spam, launching DDoS attacks, scanning for vulnerabilities, etc.) and by tainting and obfuscating C&C communication. However, this often contradicts the goal of providing the most prompt, powerful and efficient implementation of malicious campaigns. On the other side, attackers invest great efforts in making the presence of malware undetectable at compromised machines through a number of client level resilience techniques such as rootkit ability and code obfuscation [34–36]. Attackers also try to deploy a number of network based resilience techniques such as Fast-flux, Domain-flux and encryption but these techniques often introduce additional traffic traits that can be used for detection [60, 61]. Furthermore, as network-based detection is primarily based on the passive analysis of network traffic it is more stealthy in its operation and even undetectable to attackers in comparison to the client-based detection which could be detected by the malware operating at the compromised machine. Finally, depending of the point of traffic monitoring network-based detection can have a wider scope then the client-level detection systems. When

deployed in core and ISP networks network-based detection approaches are able to capture traffic from a larger number of client machines. This provides the ability of capturing additional aspects of botnet phenomena, for instance, group behavior of bots within the botnet, time regularities of bots' activity and diurnal propagation characteristics of botnets.

The point of traffic monitoring

Based on the point of traffic monitoring the approaches can target malware at client machines, local and enterprise networks and large-scale ISP networks. The main difference between different types of methods is in the network scope they cover. By analyzing traffic at the client machine only one compromised machine can be detected while implementing the detection system further from client machines would include traffic from multiple potentially compromised machines. However, implementing traffic monitoring in the higher network tiers implies the need for processing larger amount of data.

Detection of malware at local and enterprise networks is implemented closer to client machines usually in the routers or gateways connecting certain enterprise network to the Internet. Enterprise or campus networks are usually realized as a set of LANs (Local Area Networks) where some of them can be geographically separated. These networks are usually based on heterogeneous communication technologies while relying on VLAN (Virtual LAN) for the networking of geographically distanced LANs. A typical example of such network is university campus network or enterprise network.

The main opportunities for traffic monitoring at enterprise networks are following. First, traffic is monitored closer to client machines thus having the capabilities of more precisely pinpointing potentially compromised clients. In enterprise network one organization is usually the owner of the infrastructure thus having the ability of identifying compromised machines in more details. It should not be forgotten that NAT (Network Address Translation) is also used within enterprise networks so it could possibly pose some challenges in identifying compromised clients. However, at least the network is owned by the same organization so the problematic clients could be more easily identified. Second, the enterprise networks are usually characterized by a relatively manageable amount of traffic, opening possibilities for more detailed analysis of network traffic in on-line scenarios.

The main drawbacks of monitoring traffic at enterprise networks is the fact that this does not give a "bigger" picture on the operation of botnets. Botnets are characterized by a usually large set of compromised machines distributed over different countries and networks of different ISPs. Furthermore, these machines are relying on the same C&C infrastructure thus contacting same C&C servers, using the same sets of DGA generated domains, etc. Finally, botnets implement often distributed attack campaigns such as

DDoS attacks, click fraud and spam distribution. This creates a significant amount of distinguishable characteristics that can be used for identifying botnets. Monitoring traffic at enterprise network would not be able to identify the majority of these characteristics such as group behavior of bots, diurnal nature of bot activity, C&C infrastructure shared by many bots, etc.

Detection of malware in ISP networks is implemented further away from client machines usually in the backbone routers or at DNS servers. Monitoring traffic in these networks is fundamentally different from traffic monitoring at local/enterprise networks. However, some of the differences at the same time define the main opportunities of these approaches. First, by monitoring traffic in ISP network there is possibility of capturing a series of botnet characteristics not visible from the local perimeter. However, the main drawbacks of monitoring traffic at ISP network is the fact that there is a vast amount of traffic that need to be processed. This often requires the use of costly network traffic sniffers, while processing such a large amount of data in on-line fashion represents a great computational challenge. Furthermore, the use of NAT in this case poses more critical challenge as such systems would usually only be able to identify public IP addresses of the networks of large companies or organization in which compromised computers "hide".

The principles of operation

The existing network-based detection approaches rely on various principles of operation. Based on the stealthiness of functioning detection methods can be classified as *passive* or *active*. The passive detection approaches do not interfere with malware operation directly, but operate based on observation only, which makes them stealthy in their operation and undetectable by the attacker. The active detection methods, on the other hand, are more invasive methods that actively interfere with malicious activities or C&C communication [62]. Additionally, these techniques often target specific heuristics of the C&C communication or the attack campaign, providing higher precision of detection at the expense of flexibility and generality of the approach.

In parallel with the classification of botnet detection based on the place of implementation or the stealthiness of functioning the methods can be classified based on their functional characteristics. Typically, network-based detection approaches can be classified as *signature-based* or *anomaly-based* detection.

Signature-based detection

Signature-based methods identify malicious network traffic based on a set of signatures and rules on how does malicious traffic look like [63–67]. This class of detection approaches draws its functional principles on conventional IDS/IPS solutions that are usually based on DPI and matching signatures and

rules of anomalous network traffic. The signatures can have different form and commonly include regular expressions of payload strings, defined rules regarding malicious ports and suspicious IP addresses and rules regarding common sequence of communication actions within the botnet life-cycle. This class of detection techniques covers all three phases of botnet life-cycle and it is able to detect known traffic anomalies with high precision and commonly low number of false positives. The main drawback of signature-based approaches is that they are only able to detect known threats, and that efficient use of these approaches requires constant update of signatures. Additionally, these techniques are vulnerable to various evasion techniques that change the characteristics of malicious traffic, such as encryption and obfuscation of C&C channel, Fast-flux, Domain-flux, etc.

Anomaly-based detection

Anomaly-based detection is a class of detection methods that is devoted to the detection of traffic anomalies that can indicate existence of compromised machines within the network [68–78]. The traffic anomalies that could be used for detection differ from easily detectable as changes in traffic rate, latency, to more finite anomalies in flow patterns. Some of the most prominent anomaly-based approaches detect anomalies in packet payloads [69], DNS traffic [77, 78], botnet group behavior [76, 79], etc. The anomaly-based detection can be realized using different algorithms ranging from the statistical approaches, machine learning techniques, graph analysis, etc. In contrast to the signature-based approaches, the anomaly detection is generally able to detect new forms of malicious activity and it is more robust to existing botnet resilience techniques. However, some challenges in using anomaly-based detection still exist. This class of techniques requires the knowledge of anomalies that characterize botnet traffic. Additionally, traffic produced by modern botnets is often similar to the "normal" traffic, resulting in many false positives. Finally, the anomaly detection methods often have to analyze a vast amount of data, which is difficult to perform in real-time, making the detection of a fine-grained anomalies in large-scale networks often a prohibitive task. One of the novel and the most promising anomaly-based methods is the group of detection methods that rely on machine learning algorithms (MLAs) for detection of malware-related traffic patterns. This class of detection methods promises automated detection that can infer on how does malicious and benign traffic look like from available traffic observations.

2.2 Machine learning-based detection

MLAs have found their use for network-based malware detection due to several reasons. MLAs have the ability to find similarities i.e. "patterns" within a large amount of data. The main assumption of machine learning-based approaches is that malware create distinguishable patterns within the network traffic and that these patterns could be efficiently detected using MLA [46, 80]. Malware detection approaches promise automated data-driven detection that infers knowledge about malicious network traffic from the vast amount of available traffic traces. One of the major appeals of using MLAs is the possibility of automated operation with limited operator's involvement. Furthermore, MLAs can be used to discover anomalies from data, without the need for prior knowledge about them. Finally, combining the characteristics of MLAs with "big data" techniques and capable computing units we get a powerful weapon against the malware threat.

Machine Learning (ML), is a branch of artificial intelligence, that has the goal of constructing and studying systems that can learn from data [81]. Learning in this context implies ability to recognize complex patterns and make qualified decisions based on previously seen data. The main challenge of machine learning is generalizing knowledge derived from the limited set of previous experiences, in order to produce a useful decision for new, previously unseen, events. To tackle this problem, the field of machine learning develops an array of algorithms that discover knowledge from data, based on sound statistical and computational principles. Machine learning algorithms can be coarsely classified based on the desired outcome of the algorithm as *supervised* MLAs and *unsupervised* MLAs.

Supervised MLAs [82] is the class of well-defined machine learning algorithms that generate a function (i.e., model) that maps inputs to desired outputs. These algorithms are trained by examples of inputs and their corresponding outputs, and then they are used to predict output for some future inputs. The supervised MLAs are used for classifying input data into some defined class and for the regression i.e. predicting continuous valued output. In the context of network-based malware and botnet detection, supervised MLAs are commonly used for implementing network traffic classifiers that are able to classify malicious from benign traffic or identify traffic belonging to different malware families. Some of the most popular supervised MLA for network-based detection are: SVM (Support Vector Machines), ANN (Artificial Neural Networks), Decision tree classifiers and Bayesian classifier.

Unsupervised MLAs [83] is the class of machine learning algorithms where training data consists of a set of inputs without any corresponding output values. The goal of unsupervised learning may be to discover groups of similar examples within the input data, referred to as clustering, to determine the distribution of data within the input space, known as density estimation,

or to project the data from a high-dimensional space down to two or three dimensions for the purpose of visualization. In the context of network-based malware and botnet detection, unsupervised MLAs are commonly used for discovering similarities that characterize malicious network traffic. The main characteristic of unsupervised MLAs is that they do not need to be trained beforehand. The most popular unsupervised MLAs used for network-based detection are: K-means, X-means and Hierarchical clustering.

In both learning scenarios traffic is analyzed from a certain analysis perspective that entails how do traffic instances, that will be classified or clustered by MLAs, look like. For each of traffic instances a set of features is extracted and used within the MLAs to represent them. Choosing the right features representation is one of the most challenging task of practical deployment of MLAs. The chosen features should capture targeted botnet traffic characteristics and pose balanced requirements in terms of feature extraction and selection. Finally, in parallel with the two learning problems outlined here modern machine learning-based approaches commonly implement detection through several phases, using the combination of different MLAs or by deploying MLAs in an adaptive manner. This way fine-grained, flexible, and adaptable detection can be achieved.

Various detection methods have been developed using an array of MLAs deployed in diverse setups [56]. An extensive overview of the contemporary detection approaches based on machine learning can be found in Section 4.

3 Problem Statement

This thesis tackles the problem of network-based malware detection by proposing novel detection approaches that aim at achieving the detection of malicious network activity in less time and expense. The thesis explores the possibilities of developing a novel collaborative approach to botnet protection that would utilize insights from various detection sensors. We focus on network-based detection aspects of the collaborative framework by devising detection approaches that are aimed at identifying malware network activity at different points in the network and based on different, mutually complementary, principles of traffic analysis. The detection approaches proposed by the thesis rely on machine learning for identifying malicious network traffic evaluating the capabilities of machine learning algorithms to provide accurate and time-efficient detection. The proposed approaches are developed in order to cover different aspects of malware network activity and thus be suitable candidates for a future collaborative protection system. Finally, the thesis provides a comprehensive overview of existing network-based detection approaches with the focus on machine learning-based approaches.

The overall problem statement addressed by this thesis can be formulated as the following:

Can malware network activity be detected in accurate and time-efficient manner within the operational network environment using machine learning techniques?

We address the outlined problem by answering the following research questions:

- *How can a collaborative approach to botnet detection be realized?*
- *What are the main challenges of using machine learning for network-based malware detection and how can they be overcome?*
- *How can malicious network activities be identified at enterprise networks?*
- *How can malicious network activities be identified in ISP networks?*

Research question 1: How can a collaborative approach to botnet detection be realized?

Modern malware represents complex phenomena that can be tackled using a diverse set of detection methods that target different aspects of the malware operation. Malware is targeted by both client- and network-based detection solutions, while detection based on network traffic analysis can be implemented in different parts of the network targeting different traits of malicious traffic. One of the challenging research questions is how can the information

from diverse detection systems be used in order to develop more effective detection in comparison with the independent detection systems.

We answer the presented research question by proposing a novel framework for collaborative botnet protection based on indications that originate from a multiple set of sensors deployed at both client machines as well as in different parts of network in order to achieve accurate and reliable detection. We stress the crucial role of network-based detection is such a framework and the possibility of complementary network-based detection solutions that are based on different principles of traffic analysis and deployed at different points in the network.

Research question 2: What are the main challenges of using machine learning for network-based botnet detection and how can they be overcome?

MLAs have been successfully used for identifying malware network activity by a number of detection approaches over the last decade [14, 56]. However, the popularity of this class of detection approaches within scientific community is not necessarily followed by their wide-spread use within operational networks [84]. This discrepancy is often based on the characteristics of MLAs and should be properly understood in order to develop detection approaches that will have a good outlook of being implemented in real-world detection solutions.

We answer the presented research question by providing a comprehensive overview of the existing work on the use of MLA for identifying botnet network traffic. Our goal is to clarify the opportunities and the challenges of using MLAs for network-based detection, in order to be able to develop more robust and reliable detection solutions. Furthermore, we contribute to solving the "ground truth" problem as one of the most crucial challenges of the effective use of MLAs, by proposing a novel approach to labeling agile DNS traffic. Finally, we elaborate on a number of pitfalls that should be avoided in order to make machine learning-based detection approaches more suitable for the use in operational networks.

Research question 3: How can malicious network activities be identified at enterprise networks?

The analysis the network traffic on local and enterprise networks have an advantage of being able to capture more precise patterns of malicious traffic and having full visibility of client machines within the network. The traffic load on local network is arguably suited to even complex traffic analysis approaches. A number of detection approaches target malicious traffic at local networks using MLAs. The approaches cover different types of malware

network activity and report varying detection performance, thus opening the space for novel detection approaches that would provide more accurate and efficient detection.

We address the presented research question by devising a series of detection methods for identifying botnet network traffic at local and enterprise networks. The proposed detection methods evaluate several traffic analysis scenarios in order to conclude on the ability of providing accurate and timely detection using MLA.

Research question 4: How can malicious network activities be identified in ISP networks?

The analysis of network traffic in ISP networks offers both a number of opportunities and challenges. These approaches have the advantage of having a more comprehensive overview on the network thus being able to capture fundamental characteristics of botnet such as group behavior of compromised client machines. The main challenge of these approaches is the high traffic load that need to be processed. This is commonly solved by filtering or sampling the traffic. Alternatively, some authors rely on DNS traffic as it represents only a subset of total amount of traffic, and it is widely used by malware in different phases of their life-cycle.

We answer the presented research question by proposing a novel approach for identifying potentially compromised clients based on DNS traffic analysis in large-scale ISP networks. The proposed detection method relies on MLA for identifying malicious agile domains-to-IPs mappings i.e. Fast-flux and Domain-flux as resilience techniques often used by malware. Furthermore, the approach is able to pinpoint potentially compromised clients that have contributed to identified malicious mappings.

4 The state of the art

This section presents the state of the art on network-based detection approaches with the focus on the use of machine learning for identifying malware traffic activity. The section presents the contemporary scientific efforts on botnet detection by answering a series of relevant questions regarding the existing work. The section provides a comprehensive outlook on the characteristics of the existing methods, opportunities they offer and their challenges and limitations.

4.1 Collaborative detection

In this subsection we answer a series of relevant questions regarding existing collaborative detection methods pinpointing the role of network-based detection within them.

- **What are the most well-regarded collaborative detection approaches?**

  Faced with many challenges of detecting modern malware, scientific community turned their efforts to the development of novel collaborative class of detection approaches that integrate diverse principles of botnet detection. The general approach of correlating aspects of behavior observed by various sensors is elaborated by Oliner et al. [85] and Flaglien et al. [86] extending older proposals for the correlation of alerts in IDS systems made by Cuppens and Miège [87] and Ning et al. [88]. Building on these concepts several authors have proposed botnet detection systems that correlate alerts from diverse detection entities in order to achieve more accurate and reliable detection [55, 57–59]. The basic characteristics of the existing collaborative detection methods are presented in Table 2.

Table 2: Contemporary collaborative botnet detection methods

Detection Method	Client-level Analysis	Network-level Analysis	Traffic Monitoring Point	Correlation of Findings
Zeng et al. [55]	Classification of client-level behavior	Clustering of traffic flows	Client, LAN, Campus	Score-based correlation
Wang et al. [57]	Support for various client-level detection methods	Support for various network-level detection methods	Client, LAN, ISP	Dempster-Shafer theory
Muthumanickam et al. [58]	Clustering of client-level behavior	Clustering of P2P nodes	LAN, Campus	Score-based correlation
He et al. [59]	Conventional AV solutions	Clustering of traffic flows	LAN	Score-based correlation

- **What are the basic principles of collaborative detection presented by existing work?**

  The contemporary detection methods propose collaborative detection using diverse principles of client- and network-level analysis as well as correlation of findings from different detection sensors. The client-level analysis typically includes the analysis of malware behavioral at client machines by analyzing changes to file system, registry keys, function calls, and so forth [55, 58]. Some approaches rely on conventional AV solutions for providing input on potentially compromised clients [59]. The analyzed approaches commonly rely on MLA as a tool for identifying client-level malware behavior [55, 58]. The existing work analyzes network traffic by deploying network sensors at different points in network, commonly local and campus networks. Furthermore, in the majority of addressed approaches MLAs are used in order to cluster network observations. Framework proposed by Wang et al. [57] provides the possibility of correlating findings from various client- and network-based analysis solutions.

  Network-based detection plays a crucial role in the existing collaborative detection methods as modern malware is characterized by network activity used for propagating, C&C communication and implementing its malicious agenda. As an illustration, Zeng et al. [55] proposed the framework that first identifies suspicious clients by discovering similar behavior among clients using flow-based traffic analysis, and then validates the identified suspects to be malicious or not by scrutinizing their client-level behavior. In order to take the full advantage of network traffic analysis some existing approaches [55, 57] envision the use of several network analysis entities employed at different points in network starting from client machine to backbone networks.

  The correlation of findings from independent analysis systems is commonly realized by custom scoring systems where detection indications are contributing to detection score that can indicated existence of compromised clients [55, 58, 59], while Wang et al. [57] have envisioned the use of Dempster-Shafer (D-S) theory for fusing the indication from different detection entities.

- **What is the state of the operational deployment of these methods?**

  To the best of our knowledge existing collaborative detection methods still have not seen a wider operational deployment, and are mainly proofs of concept demonstrating that systems which combine information from multiple sources can achieve increased accuracy in recognizing the presence of compromised clients. In addition to technical difficulties in realizing such detection systems there are legal and soci-

ological difficulties as well. The realization of such detection solutions requires an active role of ISPs followed by collaboration from end users. Generally, ISPs should implement network traffic analysis in the core network as they are the owners of the infrastructure. They should also be the one to persuade clients in adopting suitable client-level detection solutions. An alternative approach would be to establish collaboration between conventional AV companies that already provide client-level detection solutions with ISPs. However, this solution would open some concerns regarding sharing of clients' information, that are yet to be clarified. Finally, current legislation in the many countries does not find ISPs accountable for any malicious activity in their network thus not providing enough incentives for ISPs to invest in often complex and expensive collaborative detection solutions. However, in the future we should expect to see ISPs offering complete security solutions based on the principles of collaborative detection as the service to their customers.

4.2 Signature-based detection

In this subsection we answer a series of relevant questions regarding existing signature-based detection methods.

- **What are the most well-regarded signature-based detection approaches?**

  The signature-based detection represents one of the oldest classes of network-based detection approaches that has its roots in conventional IDS/IPS systems. Some of the most well-known contemporary signature-based detection approaches are presented in Table 3. Snort [63], Bro [64] and Suricata [65] represent some of the most widely used IDS/IPS systems that have also found the use in identifying botnet network traffic. Gu et al. [66] and Goebel et al. [67] on the other hand proposed detection approaches that are specially developed in order to target botnets.

- **What are the capabilities of existing detection methods?**

  The contemporary detection methods typically contribute to the identification of compromised client machines i.e. bots [63–67] but can also contribute to identification of malicious C&C servers [63–66]. The majority of the analyzed detection approaches encompasses all phases of operation while some as Goebel et al. [67] target C&C communication as the main characteristic of botnet operation. Furthermore, the approach proposed by Goebel et al. is limited to IRC botnets while other approaches are able to capture botnet detection independent of C&C

Table 3: Contemporary signature-based botnet detection methods

Detection Method	Traffic Monitoring Point	Detection Target	Botnet Type	Operational Phase	Real-time Operation	Signatures and Rules
Snort [63] Bro [64] Suricata [65]	Client, LAN, ISP	Bots, C&C Servers	Generic	Propagation, C&C, Attack	*	IP and port rules Communication sequence Payload content
Gu et al. [66]	Client, LAN, ISP	Bots, C&C Servers	Generic	Propagation, C&C, Attack	*	Botnet life cycle
Goebel et al. [67]	LAN, Campus	Bots	IRC	C&C	-	Port, IRC nicknames

communication protocol. Finally, real-time operation is provided by the majority of the approaches [63–66] facilitating on-line operation even in large-scale networks.

- **What are the basic principles of operation of existing signature-based approaches?**

  This class of detection approaches draws from the principles of functioning of conventional IDS/IPS solutions that are usually based on DPI and matching signatures and rules regarding anomalous network traffic. The signatures and rules can have different form ranging from regular expressions of payload strings, defined rules regarding malicious ports, IP addresses and following common sequence of communication actions within the botnet life-cycle. Snort [63], Bro [64] and Suricata [65] rely on the libraries of rules on anomalous malicious traffic as well as signatures of malicious payloads. Gu et al. [66] and Goebel et al. [67] approaches are on the other hand, specially developed to target the signatures of botnet traffic. Gu et al. [66] approach represents botnet detection approach developed as the extension of Snort. The proposed platform ensures efficient botnet detection by adding the two anomaly detection plug-ins on top of existing Snort's signature database. The approach defines a model of botnet infection dialog process and then uses it as a guideline for the recognition of infection processes within the network. Goebel et al. approach [67] is one of the first detection techniques to tackle IRC botnets. This approach provides botnet detection by matching IRC nicknames with nickname patterns of IRC bots.

- **What are the main drawbacks of signature-based approaches?**

  The signature-based detection approaches rely on signatures and rules regarding malicious network traffic so they only capture "known" anomalies of malicious traffic and known attacks and intrusion. This is the main drawback of this class of detection approaches as the attacker is able to evade them by slightly changing network actions by obfus-

cating and randomizing connection attempts. Furthermore, as these detection techniques heavily rely on DPI they can be defeated using payload encryption. Finally, although promising detection with zero False Positives (FP), these approaches are often associated with a number of misidentified traffic events due to build-in statistical approaches and imprecise signatures and rules on malicious traffic.

- **What is the state of the operational deployment of these methods?**

  These methods have a wide operational implementation and they are an integral part of many real-world detection systems. These approaches are often used within enterprise networks while the latest implementations also promise the use of these approaches on high speed network links. Some existing commercial solutions offer the implementations of Snort and Gu et al. approach on top of high speed network sniffers that can be implemented in core networks at network links with speeds higher than 10Gb/s [89]. The great adoption of signature-based approaches in operational networks can be explained due to one of the main requirements of operational networks which is low FP rates. Operational networks do not tolerate high number of FP rates as that would put additional stress on network operators and would deemed the detection approach in question unreliable and consequently ignored.

4.3 Anomaly-based detection

In this subsection we answer a series of relevant questions regarding existing anomaly-based detection methods.

- **What are the most well-regarded anomaly-based detection approaches?**

  The anomaly-based detection approaches target network traffic anomalies that characterize malware network activity by relying on different anomaly detection algorithms. Some of the traffic anomalies targeted by these approaches are: group activities of compromised clients, the periodicity of botnet network activity, scanning activity, etc. The overview of some of the most well-known contemporary anomaly-based detection approaches is presented in Table 4.

- **What are the capabilities of existing detection methods?**

  The contemporary anomaly-based detection methods are typically developed for identifying compromised client machines by monitoring traffic at local networks. Only a handful of approaches are capable of facilitating detection at ISP networks [76, 78]. The majority of approaches

Table 4: Contemporary anomaly-based botnet detection methods

Detection Method	Traffic Monitoring Point	Detection Target	Botnet Type	Operational Phase	Real-time Operation
Binkley et al. [68]	LAN	Bots	IRC	C&C, Attack	-
Gu et al. [69]	LAN, Campus	Bots, C&C Servers	IRC, HTTP	Propagation, C&C, Attack	-
Kang et al. [70]	LAN	Bots	P2P	C&C, Attack	-
Wang [71]	LAN, Campus	Bots	IRC	C&C	-
Ma et al. [72]	LAN	Bots, C&C Servers	IRC	C&C	-
Al-Duwairi et al. [73]	LAN	Bots	IRC, HTTP	C&C, Attack	-
Wang et al. [74]	LAN	Bots	Generic	C&C, Attack	-
Yu et al. [75]	LAN	Bots	Generic	C&C, Attack	*
Karasaridis et al. [76]	ISP	C&C Servers	IRC	C&C	-
Villamarín-Salomón et al. [77]	LAN, Campus	Bots	Generic	C&C	-
Ramachandran et al. [78]	ISP	Bots	Generic	C&C	*

target specific types of botnets based on used C&C communication such as IRC, HTTP and P2P, while some of the approaches are able to identify botnets independent of C&C communication protocol [74, 75, 77, 78]. All analyzed detection approaches capture C&C phased while some are also able to capture attack activities [68–70, 73–75]. Finally, real-time operation is provided by handful of approaches such as Ramachandran et al. [78] and Yu et al. [75].

- **What are the basic principles of operation of existing anomaly-based approaches?**

  This class of approaches targets a number of traffic anomalies that characterize botnets. Table 5 presents traffic anomalies targeted by the existing methods and anomaly detection algorithms used to capture them. Some of the most popular traffic anomalies are group behavior [69, 76], packet size [70, 72–74], periodicity of behavior [72, 73] and DNS request/response dynamics [77, 78]. The used anomaly detection algorithms greatly vary where some of the most well-known ones are: threshold-based [68, 72, 78], fuzzy logic [73, 74], Discrete Fourier Transformation [75] and Bayesian approach [76, 77]. In addition to the presented anomaly detection approaches there is a prominent sub-class of detection approaches that rely on MLA for detecting anomalous malware network activity. This sub-class of detection approaches is thoroughly analyzed in Subsection 4.4.

- **What are the main drawbacks of anomaly-based approaches?**

  The main drawback of the anomaly-based approaches is the fact that they commonly target a specific "known" anomaly of malicious network activity. This class of detection approaches can commonly be successfully avoided by changing the characteristics of malicious traffic.

Table 5: Contemporary anomaly-based botnet detection methods: targeted anomalies and anomaly detection algorithms

Detection Method	Targeted Anomaly	Anomaly Detection Algorithm
Binkley et al. [68]	TCP scanning activity	Threshold-based
Gu et al. [69]	Long-lived C&C sessions Active response by bots Bots group behavior	Correlation analysis
Kang et al. [70]	Number of ICMP, UDP, SMTP packets	Multi-chart CUSUM
Wang [71]	Similarity of IRC nicknames	Euclidean distance
Ma et al. [72]	Packet size, Periodical repeatability	Threshold-based, Ukkonen algorithm
Al-Duwairi et al. [73]	Packet size Exchange ratio Periodical repeatability	Fuzzy Logic
Wang et al. [74]	Failed DNS queries DNS query intervals Failed network connections Payload sizes	Fuzzy Logic Pattern Recognition
Yu et al. [75]	Flow duration Number of packets/bytes per flow	Incremental Discrete Fourier Transform
Karasaridis et al. [76]	Bots group behavior	Bayesian approach / Modified K-means algorithm
Villamarín-Salomón et al. [77]	DNS traffic similarities	Bayesian approach
Ramachandran et al. [78]	DNSBL lookups by botmasters	Threshold-based

This problem is solved by anomaly detection algorithms that are able to learn from observations and adapt to the chaining nature of network traffic such as MLAs.

- **What is the state of the operational deployment of these methods?**

  The principles of the anomaly-based detection are commonly used for detecting malware related network traffic within some of the existing IDS/IPS systems. As an example, BotHunter [66] relies on the principles of anomaly-based detection for identifying payload distribution and scanning activities. In addition, many anomaly-based approaches are capable of operating on large-scale networks [76, 78] and promise real-time operation [75, 78] thus fulfilling some of the most important requirements of operational deployment. Finally, machine learning-based detection as a sub-class of anomaly-based detection has been a part of many existing operational detection solutions.

4.4 Machine learning-based detection

In this subsection we answer a series of relevant questions regarding existing machine learning-based detection methods. The findings presented in the following are based on a survey of contemporary machine learning-based botnet detection approaches presented in Paper II.

- **What are the most well-regarded machine learning-based detection approaches?**

  Since 2006 when Livadas et al. [90] proposed the first machine learning-based approach targeted at IRC-based botnets a number of authors have chosen to rely on MLAs for identifying patterns of malware related traffic [53–55, 90–106]. The existing approaches employ different principles of network traffic analysis, they rely on different MLAs and they provide varying performance of identifying malicious network traffic. The overview of the 20 most well-known contemporary machine learning-based detection approaches is presented in Table 6.

Table 6: Contemporary machine learning-based botnet detection methods

Detection Method	Traffic Monitoring Point	Detection Target	Botnet Type	Operational Phase	Communication Protocol	Real-time Operation
Masud et al. [53]	Client	C&C Servers	IRC	C&C, Attack	TCP	-
Shin et al. [54]	Client	Bots	Generic	C&C, Attack	TCP, UDP, DNS	-
Zeng et al. [55]	Client, LAN, Campus	Bots	Generic	C&C, Attack	TCP, UDP	-
Livadas et al. [90]	LAN, Campus	Bots	IRC	C&C	TCP, IRC	-
Strayer et al. [91]	LAN, Campus	Bots	IRC	C&C	TCP, IRC	potentially
Gu et al. [92]	LAN, Campus	Bots	Generic	Propagation, C&C, Attack	TCP, UDP	potentially
Choi et al. [93]	Campus, ISP	Bots	Generic	C&C	DNS	advertised
Saad et al. [94]	LAN	Bots	P2P	C&C	TCP, UDP	-
Zhao et al. [95]	LAN	Bots	P2P	C&C, Attack	TCP, UDP	potentially
Zhang et al. [96]	LAN, Campus	Bots	P2P	C&C	TCP, UDP	potentially
Lu et al. [97]	LAN	Bots	IRC	C&C	TCP, UDP	potentially
Bilge et al. [98]	ISP	C&C Servers	Generic	C&C	TCP, UDP	advertised
Bilge et al. [99]	ISP	Bots, C&C Servers	Generic	C&C	DNS	advertised
Antonakakis et al. [100]	ISP	Bots, C&C Servers	Generic	C&C	DNS	potentially
Antonakakis et al. [101]	ISP	Bots, C&C Servers	Generic	C&C	DNS	potentially
Perdisci et al. [102]	ISP	Bots, C&C Servers	Generic	C&C	DNS	potentially
Tegeler et al. [103]	LAN, ISP	C&C servers	Generic	C&C	TCP, UDP	advertised
Zhao et al. [104]	ISP	Bots	P2P	C&C	DNS	-
Zhang et al. [105]	LAN, Campus	Bots	P2P	C&C	TCP, UDP	potentially
Haddadi et al. [106]	LAN	Bots	HTTP	C&C	HTTP	-

- **What are the capabilities of existing detection methods?**

  The contemporary machine learning-based detection methods typically contribute to the identification of compromised client machines i.e. bots [54, 55, 90–97, 104–106] or malicious C&C servers [53, 98, 103]. Detection methods that target malicious DNS traffic [99–102] commonly provide identification of malicious domains which can contribute to the detection of both bots that try to resolve them as well as C&C servers.

The majority of the analyzed detection approaches target C&C communication as the main characteristics of botnet operation, while some also include the ability to capture botnet attack campaigns [53–55, 92, 95]. The propagation phase is covered by only one detection method [92], most likely as the propagation could be effectively tackled by existing IDS/IPS systems.

The methods analyze different communication protocols in order to perform botnet detection. The existing detection methods commonly analyze TCP, UDP and DNS protocols as the main carriers of botnet network activity. The majority of detection approaches rely on the analysis of TCP and UDP traffic while some more specifically cover IRC [90, 91] and HTTP [106] protocols as they are targeting IRC and HTTP botnets. One of the approaches analyzes all three protocols in order to capture the majority of the botnet network activities [54].

A number of existing detection methods are independent of C&C communication [54, 55, 92, 93, 98–103], while others target specific types of botnets, such as IRC-based [53, 90, 91, 97], HTTP-based [106] and P2P-based [94–96, 104, 105] botnets by relying on specific traits of IRC, HTTP and P2P C&C channels, respectively. It should be noted that we consider that DNS-based detection methods [99–102] can contribute to the detection of botnets independent of the used C&C communication technology.

The real-time operation is promised by only a subset of approaches [93, 98, 99, 103]. Some of the contemporary detection approaches show the potential of providing real-time detection as they operate in a time window and they could be periodically re-trained using the new training set or by periodically updating the clusters of the observation [91, 92, 95–97, 100–102, 105]. Finally, some methods such as [99] have proved their ability of real-time operation through a real-world operational deployment.

- **At what point in network existing methods monitor traffic?**

  The machine learning-based approaches can be implemented at client computers, local/enterprise networks and ISP networks. The majority of contemporary detection approaches addressed by this survey monitor traffic at local [94, 95, 97, 106] and possibly campus/enterprise networks [90–93, 96, 105], while others can be implemented in core and ISP networks [93, 98–104]. Finally, there are several approaches that target malware at client computers by strongly relying on network traffic analysis [53–55]. As already indicated in Section 2 the point of traffic monitoring defines the visibility of network space but also the principles of traffic analysis.

- **What are the principles of traffic analysis employed by contemporary machine learning-based detection approaches?**

  The existing detection methods analyze network traffic from different perspectives i.e. based on different principles of analysis. Furthermore, different methods rely on MLAs to different degree where in some MLAs play only a minor role while in others MLAs are the key element of the detection approach. Table 7 provides an overview of the principles of traffic analysis used by the contemporary machine-learning detection approaches.

  The existing methods use several perspectives of traffic analysis. The approaches that analyze TCP and UDP traffic generally analyze it from the perspective of traffic "flows". It should be noted that definition of a flow varies from the approach to the approach so some use NetFlow flows [55, 98, 103] while others use a conventional definition of traffic flows where a flow is defined as traffic on a certain 5-tuple i.e. $<ip_{src}, port_{src}, ip_{dst}, port_{dst}, protocol>$. Furthermore, some approaches consider bi-directional flows in order to capture the differences in incoming and outgoing traffic [95]. DNS-based detection approaches commonly target agile DNS i.e. IP-flux and Domain-flux techniques. They do so by analyzing DNS traffic from the perspective of DNS query responses (i.e. domain names and their resolving IPs) [93, 99–101, 104], while some analyze it from the perspective of domain clusters [102].

  Traffic instances are represented as sets of traffic features in MLAs. As already indicated, feature engineering is a challenging task as the chosen feature representation should capture targeted characteristics of malicious traffic. The analyzed detection approaches greatly vary in employed feature representation. The TCP/UDP based approaches use features that are generally independent from the payload content, relying on the information that can be gathered from packets headers as well as different traffic statistics. Several techniques [53, 92, 97, 106] rely on the content of payloads thus being easily defeated by the encryption or the obfuscation of the packet payload. Furthermore, some approaches rely on IP addresses as features [94, 95] opening the possibility of introducing bias in the evaluation of the detection performance. In the case of the DNS analysis approaches typically rely on information extracted from the DNS query responses, such as: lexical domain name features, IP-based features, geo-location features, etc.

- **What are the most common learning principles used by the existing methods?**

  As illustrated by Table 7, the existing methods use a variety of machine

Table 7: Contemporary machine learning-based botnet detection methods: traffic analysis perspective and machine-learning algorithms

Detection Method	Analysis Perspective	**S**upervised / **U**nsupervised	MLAs
Masud et al. [53]	Flow	S	SVM, C4.5, Naive Bayes, Bayes Network, and Boosted decision tree classifiers
Shin et al. [54]	Flow	S	Correlation of the findings of two MLAs: SVM and One Class SVM (OCSVM)
Zeng et al. [55]	Flow	S,U	Correlation of the findings of two MLAs: Hierarchical clustering and SVM
Livadas et al. [90]	Flow	S	C4.5 Tree, Naive Bayes and Bayesian Network classifiers
Strayer et al. [91]	Flow	S	C4.5 Tree, Naive Bayes and Bayesian Network classifiers
Gu et al. [92]	Client	U	Two level clustering using X-means clustering
Choi et al. [93]	DNS query/response	U	X-means clustering
Saad et al. [94]	Flow	S	SVM, ANN, Nearest Neighbours, Gaussian and Naive Bayes classifiers
Zhao et al. [95]	Flow	S	Naive Bayes and REPTree (Reduced Error Pruning) Decision Tree
Zhang et al. [96]	Flow	U	Two level clustering using BIRCH algorithm and Hierarchical clustering
Lu et al. [97]	Flow	U	K-means, Un-merged X-means and Merged X-means clustering
Bilge et al. [98]	Flow	S	C4.5, SVM, and Random Forest classifiers
Bilge et al. [99]	DNS query/response	S	C4.5 classifier
Antonakakis et al. [100]	DNS query/response	S,U	X-Means clustering and Decision Tree using Logit-Boost strategy (LAD)
Antonakakis et al. [101]	DNS query/response	S	Random Forest classifier
Perdisci et al. [102]	Clusters of domain names	S	C4.5 classifier
Tegeler et al. [103]	Flow	U	CLUES (CLUstEring based on local Shrinking) algorithm
Zhao et al. [104]	DNS query/response	S	REPTree (Reduced Error Pruning) Decision Tree
Zhang et al. [105]	Flow	U	Two level clustering using K-means algorithm and Hierarchical clustering
Haddadi et al. [106]	Flow	S	C4.5 classifier

learning algorithms deployed in diverse setups. In total 15 different MLAs were considered by the analyzed approaches. Supervised and unsupervised MLAs are evenly represented in the analyzed methods. Some of the authors experimented with more than one MLA providing the good insight on how the assumed traffic representation holds in different learning scenarios as well as what are the performance of different MLAs [53, 90, 91, 94, 95, 97, 98]. Additionally, some authors used MLAs in more advanced setups, where clustering of obser-

vation is realized through two level clustering schemes [92, 96, 105] or where the findings of independent MLAs were correlated in order to pinpoint the malicious traffic pattern [54, 55, 92, 100]. Finally, several authors used the same MLAs within their detection systems [53, 90, 91, 94, 95, 98, 99, 101, 102] providing us with the opportunity to assess their capability of capturing network traffic anomalies in different commonly independent data sets.

- **What MLAs are best suited for identifying malware network traffic?**

  As already mentioned, a number of MLAs have been used in order to develop the existing detection methods. Some of the most popular supervised MLAs are Artificial Neural Networks, Tree Classifiers, Naive Bayes Classifier, Bayesian Network Classifiers, Nearest Neighbors Classifier and a number of ensemble classifiers. In parallel a number of unsupervised approaches have been used where some of the most often used ones are K-means, X-means and Hierarchical clustering.

  Based on the existing work, some of the best performing supervised MLAs are decision tree classifiers including C4.5, Random Forests, REPTree classifiers. The tree classifiers have shown to provide overall good performance in both terms of accuracy of classification as well as the time needed to perform training and classification tasks. The latter should not be overlooked as having time-efficient machine learning algorithm is often one of the most important factors for operational implementation. The most popular unsupervised MLAs are Hierarchical clustering and X-means clustering. The reason for this is that these algorithms do not need to be provided with a number of expected clusters such in case of k-means clustering.

- **How good are the performance of the existing machine learning-based detection approaches?**

  The contemporary detection approaches have reported mostly affirmative detection performance that confirm the potential of using MLAs for the task of identifying malware related network activity. Several detection methods indicate TPR of 100% and overall low FPR [96, 104, 105]. Furthermore, a number of approaches is characterized with a FPR less than 1%. These results indicate the possibility of using some of the approaches in real-world operational networks.

  However, when assessing performance of detection methods, it is crucial to understand the used evaluation procedure. The existing methods are commonly evaluated using malicious and benign traffic traces recorded at different networks and at different times and contributed by diverse types and number of malware samples as well as traffic from

diverse types of benign applications. As a result, contemporary methods cannot be directly compared based on the reported performance alone.

In Paper II we have compared a number of approaches based on the evaluation procedure used and the reported performance of detecting malicious traffic. The evaluation indicates several things. First, benign traffic is obtained at the point in the network corresponding to the monitoring point the methods are developed for, most commonly on campus or LAN networks with relatively limited number of client machines. Second, the malicious traffic samples are usually recorded for a limited number of malware samples. For instance, the performance of only five detection approaches were evaluated on the traffic traces produced by more than 5 malware samples [54, 55, 92, 103, 106], while the maximal number of samples used for evaluation was 188 in case of [103]. The rest of the methods were tested with less than 4 bot malware samples. Finally, the diversity of the used malware samples is poor as the majority of the analyzed approaches rely on less than 3 distinct families of botnets.

- **What are the main challenges and pitfalls of using MLA for identifying malware network activity?**

  Some of the biggest challenges of using MLA for identifying malware network activity are evaluation challenge and the high cost of errors. The evaluation challenge characterizes all data-driven approaches and is related to the challenges of obtaining training and testing data [107]. The high cost of errors can be attributed to the network security application domain where misidentified events can have significant consequence on security and integrity of safety critical system [84].

Evaluation challenges

The evaluation challenges can be differentiated into two problems i.e. obtaining the evaluation data and the ground truth problem. As already indicated, the existing detection methods are developed and evaluated using various data sets of malicious and benign traffic. The used data sets are often sparse consisting of only a handful of botnet traces that are obtained in a nontransparent way. Furthermore, the approaches often rely on data sets that are artificially formed by overlaying and merging data sets recorded at different monitoring points in network. Obtaining the "quality" data for evaluation of the proposed machine learning-based detection approach is crucial to reliable evaluation. Under quality we mean a substantial amount of data that successfully captures both malicious and benign traffic characteristics. Obtaining the

high volume of traffic traces is usually not the main problem as there is an abundance of network traffic that can be recorded at diverse points in network. Depending on the principles of traffic analysis used by the proposed detection system traffic can be recorded in different parts of network from local level to higher network tiers. However, once the traffic is obtained the "true" nature of the traffic should be determined which is usually referred to as the ground truth problem.

As MLAs are data-driven methods (either supervised and unsupervised), they are dependent on the accuracy of the data set used for their development, optimization and evaluation. In case of supervised learning inaccuracy in the ground truth will consequently lead to inaccuracy in the results of classification performed by the learning technique. Furthermore, the ground truth also has a relevant role in the context of unsupervised learning, for what concerns performance evaluation.

High cost of errors

In contrast to some other MLA application domains, malware detection is more sensitive to detection errors. Generally, malware detection is affected by the false negatives as failing to identified a threat could potentially lead to the loss of sensitive information or compromising often safety critical systems. False alarms on the other hand, as in case of many other anomaly detection systems, directly affect the operational usability of the detection approaches. In case that a detection system is producing too much false alarms the operator and end-users would be burdened by it and consequently forced to ignore the detection indications altogether. This should be taken in consideration as many existing detection approaches report on paper good result with false positive rates less than 5% [54, 94, 95, 108]. However, many fail to mention the fact that such systems when faced with high number of testing samples would result in a high number of false positives. As an example, detection approaches are typically used for flow classification where on enterprise networks these systems would easily be able to observe more than 1 million flows per day. If a detection system has false positive rate of 1% this corresponds to 10000 false positives which is in any regards too much. Such a high number of false positives would deem any detection system unusable in operational environment.

- **How is the "ground truth" problem solved by the existing work?**

  One of the biggest challenges of using machine learning-based approaches is the lack of the ground truth on malicious and benign network traffic. The existing methods solve this problem by relying on

honeypots and malware testing environments for obtaining the malicious network traffic or by relying on FQDNs and IPs blacklists and whitelisting of popular domains for the labeling of pre-recorded traffic as malicious or benign. The use of domain and IPs blacklists has been one of the most criticized but yet widely used labeling practice [93, 99–102, 109]. Many authors have indicated the drawbacks of such labeling approaches indicating that blindly relying on them could lead to wrong conclusions regarding malicious and benign network traffic [110–113]. Other authors rely on Honeypots and malware testing environments for obtaining the malicious traffic traces that are then usually merged with the benign traffic recorded at an equivalent point in network. Finally, some authors combine the aforementioned practices with manual validation by the network operator [102]. Although tedious this practice often is able to eliminate the majority of wrongly labeled network traffic instances.

- **What is the state of the operational deployment of these methods?**

  Existing have shown promising performance within experimental environments but many of them have difficulties bridging the gap between experimental and operational deployment. Some of the reasons for this are elaborated by Sommer et al. [84] and Aviv et al. [107] and include the cost of errors and the lack of quality data sets used for the development and the evaluation of detection approaches.

  However, it should be noted that some companies have developed effective detection approaches that are directly based on the scientific findings in regards to machine-learning botnet detection such as Damballa [114] that has successfully deployed concepts of DNS-based detection presented by Antonakakis et al. [100, 101, 115] in real-world detection solutions. Furthermore, some MLAs have found an efficient use suitable for operational network such as Naive Bayes classifier for the classification of SPAM messages and similar. The positive example of the use of MLA in real-world detection solutions indicates the great potential of this class of anomaly detection methods.

4.5 Opportunities for future work

As indicated by the presented state of the art, botnets and malware network activity can be effectively and efficiently identified using network traffic analysis. The existing collaborative detection methods indicate the opportunities of utilizing diverse sources of information regarding malicious network activity to achieve more accurate detection. In addition, the machine learning-based approaches have proved to be able to facilitate accurate detection in operational real-world solutions.

The opportunities for future work are numerous. First, the collaborative detection it is still in its early days and there is a significant space for the development of novel collaborative detection solutions. The novel solutions should operate across different networks in order to encompass commonly geographically disperse compromised clients and thus get a more comprehensive overview of botnet operation. These solutions should rely on both the client-based and the network-based detection where the network-based detection should be implemented in different points of the network from local to backbone networks targeting different aspects of malicious traffic. The novel collaborative solutions should also rely on conventional AV providers in order to facilitate prompt update of knowledge about client- and network-level malware behavior. Finally, the novel solutions should facilitate secure and trustworthy exchange of malware related information and generated alerts. Second, the future network-based detection methods should rely MLAs as prominent class of anomaly detection methods. These methods should rely on the principles of network traffic analysis that would successfully encompass targeted botnet network characteristics and would carry the context that is understandable to the operator of the system. Furthermore, one of the important goals of future detection systems is to provide detection in less time and expense comparing to the existing solutions. The novel detection approaches should provide accurate detection in real-time thus facilitating both effective and timely detection. The future methods should be evaluated using an extensive set of network traces originating from diverse malware families and types. Finally, special attention should be placed on minimizing the number of errors in identifying malicious network traffic so the proposed methods would performance-wise be suitable for being used in operational networks.

5 Main Contributions

This section outlines main contributions of the thesis by outlining contributions of work presented in the attached papers. Each paper is presented, addressing its position in the thesis as well as its contributions and limitations.

5.1 An overview of thesis contributions

In this section we summarize the scientific contributions of 6 appended research papers.

Paper I **A collaborative approach to botnet protection.**
This paper proposes a novel collaborative approach to botnet protection that integrates findings of diverse botnet detection solutions in order to achieve effective detection and mitigation of botnets.

Paper II **On the use of machine learning for identifying botnet network traffic.**
This paper provides a comparative analysis of the existing methods that rely on MLA for identifying botnet network traffic. The paper analyzes the characteristics, the capabilities and the challenges of this class of detection methods.

Paper III **On the ground truth problem of malicious DNS traffic analysis.**
This paper addresses the "ground truth" problem as one of the fundamental challenges of machine learning-based detection. The paper addresses the case study of agile DNS by proposing a novel approach for obtaining the ground truth on malicious and benign agile DNS traffic.

Paper IV **An efficient flow-based botnet detection using supervised machine learning.**
This paper proposes a novel approach for identifying botnet traffic at local and enterprise networks based on network traffic classification. The paper explores if detection can be achieved in less time and expense in comparison to the existing work.

Paper V **An analysis of network traffic classification for botnet detection.**
This paper proposes three novel network traffic classification approaches for identifying botnet network activity at local and enterprise networks. The paper evaluates the capabilities of the three approaches to obtain accurate and timely detection.

Paper VI A method for identifying compromised clients based on DNS traffic analysis.
This paper proposes a novel method for identifying potentially compromised clients based on DNS traffic analysis at large-scale ISP networks. The approach identifies malicious agile domains-to-IPs mappings and tracks potentially compromised clients that have contributed to them.

The contributions of the thesis can be differentiated into four categories in accordance to the four research questions addressed by the thesis. More detailed summary of the attached papers and their contributions is presented in the following.

5.2 Collaborative approach to botnet detection

The first group of contributions is related to the development of a novel collaborative botnet detection approach with the goal of achieving effective and efficient detection of botnets. These contributions are covered by Paper I.

Paper I - A collaborative approach to botnet protection

Motivation - Modern malware and botnets as their latest incarnation represent complex phenomena that can be addressed using different detection solutions. As illustrated on the ZeroAccess botnet case study presented in Section 1, modern malware can be tackled based on both client-level and network-level information while network-based detection solutions can be implemented in different parts of the network and based on diverse traffic analysis principles. Various detection approaches have different often non-overlapping scope thus various detection capabilities and vulnerabilities to evasion techniques. This indicates the possibility of correlating findings from independent detection entities in order to achieve effective detection with wider scope and higher resilience to evasion techniques.

Research Question - How can a collaborative botnet detection be realized in order to achieve efficient and effective detection?

Paper Summary - Paper I presents ContraBot framework, a novel systematic approach to the detection and mitigation of botnets. ContraBot belongs to the emerging class of collaborative botnet detection approaches that integrate diverse principles of botnet detection, in order to provide more efficient and effective detection. The ContraBot framework consists of a number of network traffic pre-analysis and client activity pre-analysis entities that target botnets on network- and client-level,

respectively. Network traffic pre-analysis entities represent a set of network sniffers placed within the network to collect and pre-process network traffic data, while client activity pre-analysis entities represent a set of activity monitors within the clients that collect and pre-process information about client activity. The output of this pre-processing is passed to a set of one or more correlators, where it is analyzed to reveal patterns of similar behavior in different clients and different parts of the network. Unusual patterns of activity, will lead to the harvesting of portions of code from the clients and the associated network traffic, so that these can be further analyzed by entities which investigate the code for malicious elements. In addition, the framework includes client distribution analysis entities that analyze modules, applications and other forms of software fetched by the clients from the network so that well-known malicious software can be disabled or removed as in a conventional AV system. As the ContraBot framework includes input from a wide range of sources, including sensors installed by end users, ISPs and backbone network providers it is important that all parties can evaluate the trustworthiness of the input they receive. The correlation framework will therefore include a trust management component that aims to establish the trustworthiness of input based on both direct experiences with the individual input provider and reputation ratings exchanged between the different correlators in the correlation framework. If the correlators, distribution analysis entities or code analysis entities detect signs of malicious software, they pass this information to a sub-system which generates warnings for distribution to subscribers of the anti-botnet service. This allows the subscribers to initiate various counter measures, such as walled-garden, disinfection, etc.

Scientific Contribution - The main contribution of the paper is a novel framework for botnet protection that combines existing detection efforts into a collaborative botnet protection framework.

Related Work - To the best of our knowledge the ContraBot framework could possibly be the first extensive attempt to take counter botnet research to a systematic level, providing the basis for a more comprehensive botnet defense system. The proposed framework builds on the same principles as several existing approaches [55, 57–59] but also has a number of advantages comparing to them. First, the ContraBot will employ traffic analysis in the core network, providing protection for a broader set of end-users. Secondly, the proposed set-up will combine information not only from network and client levels but also from in depth analysis of harvested code in order to improve the detection accuracy even further. Third, the proposed system will provide the flexibility of including diverse end-user platforms through the development of

appropriate client-based analysis entities. Fourth, our system will also introduce the feed-back mechanism. This will provide the adaptability of network- and client-based pre-analysis entities to the bot-related information generated by correlating findings from other sources. This information allows the system to dynamically adapt to changes in the behavior of bots and botnets.

5.3 Machine learning for network-based botnet detection

The second group of contributions is related to evaluating the use of MLAs for identifying malware network activities and addressing some of the main challenges of this class of detection methods. These contributions are covered by Paper II and Paper III.

Paper II - On the use of machine learning for identifying botnet network traffic

Motivation - A number of detection methods based on MLAs have been proposed over the last decade. These methods rely on diverse MLAs, they employ different principle of network traffic analysis, they are evaluated using a range of network traffic traces and consequently they provide varying detection performance. As a result, there is a need for providing a comprehensive overview of existing scientific efforts, outlining their capabilities and challenges.

Research Question - What are the capabilities and limitations of the existing machine-learning based approaches?

Paper Summary - Paper II presents a survey on the use of machine learning for network-based botnet detection. The paper evaluates 20 prominent contemporary detection approaches by analyzing their characteristics, capabilities and limitations. The paper analyzes the characteristics of existing methods by analyzing the principles of traffic analysis, MLAs and feature representation used by the approaches. Furthermore, the capabilities of the approaches are analyzed by comparing methodologies used for performance evaluation as well as by analyzing the reported performance. The paper also evaluates the vulnerability of existing detection methods to a series of evasion techniques. Finally, the paper outlines the best practices and opportunities for future work.

Scientific Contribution - The main contribution of the paper is a comprehensive insight into the characteristics, the capabilities and the challenges of existing detection methods.

Results and Conclusions - The paper indicates that existing detection methods based on MLA show a great perspective for being used for identifying botnets i.e. malware network activity. The existing methods target traffic at different points in network where the majority of the approaches target botnets at local and campus network. Overall the methods mostly target C&C traffic by targeting TCP, UDP and DNS protocols as the main carriers of botnet activity. The traffic is commonly analyzed from the perspective of transport layer flows and DNS query responses. Furthermore, the study indicates that some of the most popular MLAs are decision tree classifiers (C4.5, Random Forests, REPTree) for classification and Hierarchical clustering and X-means clustering for grouping traffic observations.

The proposed methods have been evaluated using diverse evaluation schemes and malicious network traffic obtained in different scenarios that commonly include running malware binaries. However, the majority of existing method have been evaluated with traffic traces from a limited set of malware samples/families, thus requiring a more thorough evaluation. The reported detection performance varies where a few methods achieve performance characterized with high true positive rate and low false positive rates. Finally, real-time operation is promised by a handful of the approaches opening the space for further improvements.

Related Work - To the best of our knowledge this paper presents one of the first and the most comprehensive overviews of the machine-learning based detection methods. Garcia et al. [116] and Karim et al. [56] have also provided comprehensive surveys on existing network-based botnet detection approaches indicating the crucial place of approaches that use MLAs for identifying botnet network activity. However, Garcia et al. compared 8 detection methods based on MLAs, while Karim et al. analyzed only 3 approaches thus covering only a sub-set of the research area.

Paper III - On the ground truth problem of malicious DNS traffic analysis

Motivation - One of the main challenges of machine learning-based detection methods is the lack of the "ground truth" on the malicious and benign network traffic that is needed in order to evaluate the proposed methods and train methods based on supervised MLAs. Detection methods that target agile malicious DNS traffic i.e. Fast-flux and Domain-flux often have MLAs at their core and therefore suffer from the same problem. The existing methods commonly rely on conventional methods for obtaining the ground truth such as domain and IP

address blacklists and the whitelisting of popular domains. Blacklists are often formed in non-transparent way and based on different criteria regarding what constitutes malicious domains and IP addresses, so relying on them can lead to sub-optimal and unreliable labeling. Furthermore, whitelisting popular domains can lead to the exclusion of malicious popular domains. All of this indicates the need for a novel approach for labeling agile DNS traffic.

Research Question - How can reliable and time-efficient labeling of agile DNS traffic be achieved?

Paper Summary - Paper III elaborates the ground truth problem on the use case of agile DNS traffic. The paper gives a critical overview of the DNS labeling approaches used by contemporary DNS-based detection approaches, that are often solely based on 3rd party domain and IP blacklists. The paper also introduces a novel semi-manual labeling approach that has a goal of providing reliable and time-efficient labeling of agile DNS traffic by incorporating operator's insight in efficient manner. The labeling method analyzes DNS traffic from the perspective of domains-to-IPs mappings. We rely on DNSMap [42] for extracting agile mappings from the recorded DNS traffic. The agile domains-to-IPs mappings are analyzed as bipartite graphs where for each of them a set of distinguishable features is extracted. The used perspective of traffic analysis is suitable for manual analysis as it groups the vast amount of DNS traffic to a number of mappings suitable for manual evaluation.

The proposed labeling method combines an automated approach for labeling agile DNS traffic based on a set of analysis with a manual validation step. The automated approach covers 6 following analysis entities analyzing different characteristics of domains-to-IPs mappings: the graph analysis, the analysis of domain names, the analysis of IP addresses, the analysis of blacklisted domains, the analysis of blacklisted IP addresses and the analysis of whitelisted domains. The analysis entities extract a set of distinguishable features that are used in order to represent the domains-to-IPs mappings within clustering algorithm. The clustering algorithm pre-labels extracted mappings as malicious and benign and presents them to the operator for validation.

The proposed semi-manual labeling approach is evaluated using an extensive set of DNS traffic traces from a regional ISP. Comparison of the proposed labeling approach to conventional approaches indicates that the proposed approach is able to discover a much wider set of malicious agile domain-to-IPs mappings than conventional methods. Finally, the approach has proved to be time-efficient by incorporating operator's insight in an efficient manner.

Scientific Contribution - The main contribution of the paper is a novel method for obtaining the ground truth on malicious agile DNS traffic that incorporates operator's insight in time-efficient manner.

Results and Conclusions - The proposed labeling approach is evaluated using DNS traffic traces from a regional ISP. The labeling results are compared with existing labeling practices that predominantly rely on 3rd party domain and IP address blacklists. The experimental results confirm the importance of domain and IP address blacklists as well as domain whitelisting for labeling of DNS traffic but also indicate that the blind reliance on them may lead to misleading conclusions about analyzed DNS traffic. Comparing the proposed semi-manual labeling approach with automated labeling approaches that rely on domain/IP blacklists and domain whitelisting, the proposed approach has shown better coverage as it discovers suspicious domains/IP addresses based on their association with other rogue domains/IP addresses. Furthermore, the automated solutions lead to a number of false positives, requiring human insight in order to safeguard against them. Finally, the proposed labeling approach has proved to incorporate operator's insight in efficient manner requiring between 4-6 man-hours for evaluating labels for a week long network trace. The reported performance makes the proposed labeling approach a viable candidate for deployment within operational ISP networks.

Related Work - The majority of existing methods for identifying malicious DNS traffic [93, 99–102] rely on domain and IP address blacklists and whitelisting of popular domain for labeling DNS traffic. To the best of our knowledge the proposed method is one of the first to introduce an approach for labeling agile DNS traffic that incorporates operator in time-efficient manner.

5.4 Detection of malicious network activities at enterprise networks

The third group of contributions is related to the development of novel detection methods for identifying botnets at local and enterprise networks based on network traffic classification. These contributions are covered by Paper IV and Paper V.

Paper IV - An efficient flow-based botnet detection using supervised machine learning

Motivation - Existing methods rely on a number of different supervised MLAs for identifying botnet network activities. Furthermore, several

approaches rely on flow-level traffic analysis. This indicates the need for a thorough evaluation of the capabilities of the flow-level analysis and supervised MLAs to facilitate accurate and time-efficient identification of botnet network traffic.

Research Questions - Can the flow-level analysis and the supervised MLAs facilitate detection of botnet network traffic in less time and expense in comparison to the contemporary approaches? What supervised MLA shows the best performance in classifying botnet network traffic? What is the minimal amount of traffic per flow that needs to be considered in order to perform accurate detection?

Paper Summary - Paper IV proposes a novel botnet detection approach that analyzes network traffic from the perspective of traffic flows. The proposed method is capable of targeting botnets at local and enterprise networks by covering all phases of botnet operation and identifying botnet traffic regardless of the underlying C&C communication protocol and botnet topology. The proposed approach relies on flow-level analysis, where we define flows such that they encompass bidirectional communication via TCP, UDP and ICMP protocols. Furthermore, the paper evaluated eight different supervised MLAs thus representing one of the most comprehensive studies of the use of different supervised MLAs for the task of botnet traffic classification. The paper also analyzes how much traffic need to be analyzed per flow so botnet traffic could be accurately detected. The results of the evaluation indicate the possibilities of detecting malicious network traffic using only 10 packets per flow while monitoring flows for only a period of 60 seconds. The achieved accuracy of traffic classification is in line with results reported by the existing work. However, it should be noted that the proposed approach achieved it for a limited amount of traffic analyzed per flow.

Scientific Contribution - The main contribution of the paper is a novel detection approach that evaluates the performance of identifying botnet network activity at local and enterprise networks using the flow-level analysis and an array of MLAs.

Results and Conclusions - The proposed detection approach is evaluated using botnet traffic traces captured by honeypots and non-malicious traffic originating from diverse benign applications. For the evaluation we use the same data set as Saad et al. [94] approach thus a suitable comparison is possible. The proposed detection system has proved to be accurate in detecting botnet traffic using simple flow-level feature representation and Random Tree classifier. Additionally, the experiments showed that in order to provide a high accuracy of detection the

traffic flows need to be monitored for only a limited duration of time and a limited number of packets per flow. The obtained classification results are comparable with ones reported by Saad et al. but with the note that our approach used limited amount of traffic per flow and was able to obtain accurate results for only 10 packets per flow and 60 seconds of flow monitoring time. The results indicate the possibilities of using the presented approach in a more adaptive set-up that could facilitate on-line detection.

Related Work - The proposed method draws from the experiences and findings of several detection methods that rely on flow-level analysis [90, 94, 95]. Our solution covers all phases of botnet network activity and it is independent from C&C protocol in contrast to some existing approaches [90, 94]. Furthermore, as already indicated the proposed method is able to provide comparable detection performance by minimizing amount of traffic analyzed per flow. Finally, in contrast to such as Saad et al. approach our detection method does not rely on IP addresses or any other client identifiers as features thus avoiding the possibility of over optimistic detection using biased data sets.

Paper V - An analysis of network traffic classification for botnet detection

Motivation - As concluded in Paper IV, promising detection performance of botnet traffic can be achieved using supervised MLAs. However, the flow-level analysis used in Paper IV has limitation in capturing more detailed characteristics of traffic such as the state of TCP connections, DNS traffic queries, etc. Therefore, in order to improve classification performance more advanced traffic analysis is required. Furthermore, detection methods should be evaluated with more extensive traffic data sets in order to obtain more reliable evaluation of the performance of the method.

Research Question - Can accurate and time-efficient classification of botnet TCP, UDP and DNS traffic be realized using supervised MLAs?

Paper Summary - Paper V proposes three novel methods for network traffic classification targeting three protocols often seen as the main carriers of botnet network activity namely TCP, UDP and DNS. The proposed classifiers are capable of being used for identifying botnet traffic at local and enterprise networks covering all phases of botnet network operation regardless of the underlying C&C communication protocol. The three classifiers are developed using a capable Random Forests classifier. In contrast to Paper IV, the work presented in this paper brings more advanced traffic analysis by separating the analysis of TCP, UDP

and DNS traffic where TCP and UDP are analyzed from the perspective of bi-directional transport layer conversations while DNS is analyzed from the perspective of queries/responses for a particular domain name. Furthermore, the analysis is performed in time window thus opening the possibility of applying the proposed detection method in on-line fashion, where the traffic classifiers would be periodically retrained. Traffic instances extracted for the proposed classifiers rely on novel feature representations that should better leverage the theoretical and practical knowledge about botnet traffic anomalies. The detection methods have been evaluated using one of the most extensive botnet data sets. For the evaluation of classifiers, we considered different length of the analysis window and different number of packets per TCP/UDP conversations. The results of evaluation indicate that all three classifiers are able to achieve accurate classification (accuracy > 98%) in reasonable classification time.

Scientific Contribution - The main contribution of the paper is development of three new classifiers that provide an overall improvement in classification performance in comparison to our previous work.

Results and Conclusions - The proposed method has been evaluated using benign traffic traces recorded at local/campus networks and malicious traffic traces obtained using Honeypots and malware testing environments. It should be noted that we evaluated the presented classifiers using one of the most extensive set of botnet network traces to date. The detection performance obtained with the proposed classification methods are on pair with some of the most prominent detection methods, with precision and recall over 0.98 for all three classifiers. However, we believe that our approach has a slight advantage as the results were obtained using one of the most extensive data sets.

Related Work - The three proposed classifiers provide significant improvements in the accuracy of botnet traffic classification comparing to the classifier presented in Paper IV. Furthermore, similarly to the work presented in Paper IV the three classifiers have several advantages over the existing approaches. First, our approach is evaluated with one of the most extensive botnet data sets. Second, our solution covers all phases of botnet network operation in contrast to some existing approaches [90, 94]. Third, our detection methods do not consider the use of IP addresses or any other client identifiers as features in contrast to the existing work [94, 95].

5.5 Detection of malicious network activities in ISP networks

The forth group of contributions is related to the detection of malware network activity in ISP networks. More precisely identifying potentially compromised clients based on DNS traffic analysis. These contributions are covered by Paper VI.

Paper VI - A method for identifying compromised clients based on DNS traffic analysis

Motivation - DNS is often abused by cyber criminals in order to host malicious services and facilitate the discovery of malicious network infrastructure. Furthermore, agile DNS strategies such as Fast-flux and Domain-flux are often used by miscreants as they provide resilient relaying and reliable hosting. DNS traffic analysis is suitable for the deployment in ISP networks due to relatively small amount of DNS traffic in comparison to the total amount of network traffic. All of this indicates the possibility of using DNS traffic analysis for identifying potentially compromised clients in large-scale ISP network.

Research Question - How can an efficient identification of potentially compromised clients based on DNS traffic be implemented in large-scale ISP networks?

Paper Summary - Paper VI introduces a novel method for identifying potentially compromised clients based on DNS traffic analysis. The method targets malicious agile DNS traffic as the main carrier of illicit DNS activity. The proposed method analyzes traffic from the perspective of domains-to-IPs mappings, as in the case of the labeling approach proposed in Paper III. The extracted agile domains-to-IPs mappings are classified as malicious or benign using a novel classifier based on Random Forests classifier and a novel feature representation for agile domain-to-IPs mappings. The identified malicious mappings are then used to trace back to potentially compromised clients that have produced them. The proposed approach provides the operator with the ability of analyzing the identified malicious domains-to-IPs mappings and discovering clients that have queried domains within them. This way the operator can pinpoint the malicious agile DNS traffic of interest and can discover clients that have contributed to it. The method was evaluated using an extensive DNS traffic traces recorded at diverse ISP networks. The evaluation indicates a great potential of accurately identifying malicious agile domain-to-IPs mappings and clients that have contributed to it.

Scientific Contribution - The main contribution of this paper is a novel detection method that discovers potentially compromised clients based on DNS traffic analysis in large-scale ISP networks.

Results and Conclusions - The novel DNS-based detection approach is evaluated using extensive traffic traces from the networks of mobile and fiber ISPs. The traces were recorded over the total period of 15 weeks during the last 5 years. The obtained traffic traces are labeled using the semi-manual labeling approach proposed in Paper III. The evaluation shows promising results in identifying malicious agile DNS mappings and clients that have tried to resolve domains within them. The accuracy of identifying agile malicious domain-to-IPs mappings is just under 87% opening possibilities for future improvements.

Related Work - As in Paper III, the proposed method analyzes DNS traffic from the perspective of domains-to-IPs mappings as defined by DNSMap [42]. The analysis of domains-to-IPs mappings is recognized as a valuable mean of identifying malicious agile DNS by other authors such as Schiavoni et al. [117]. Although based on the similar DNS analysis principles Schiavoni et al. approach is only able to identify DGA based domains while our solution covers both Fast- and Domain-flux. Furthermore, the proposed method belongs to the small group of DNS-based detection methods that identify potentially compromised clients [93, 118, 119]. The main advantage of our approach comparing to the existing work is the ability to target both Fast-flux and Domain-flux and operate on large-scale ISP networks.

6 Conclusions

This section outlines how the solutions presented in this thesis can contribute to tackling the malware threat. The section also summarizes the main conclusions for each of the research questions addressed by the thesis. Furthermore, the section discusses the possibilities of applying the presented methods in real-world operational networks. Finally, the section outlines the opportunities for future work.

The solutions presented in this thesis contribute to solving the malware problem in the following ways. Paper I presents a collaborative framework for botnet protection, that represents a comprehensive solution that envisions the use of various detection and mitigation approaches in order to achieve effective protection against botnets. The proposed solution could be implemented at the network of one or multiple ISPs thus providing the protection against botnets for all clients within the network. Paper II contributes to solving the malware problem by clarifying the opportunities and challenges of using MLAs for identifying botnet network traffic through the analysis of the existing work. Paper III solves the ground truth problem as one of the biggest challenges of machine learning-based detection approaches on the case study of agile DNS traffic. The proposed solution provides the labeling of data sets needed for the training and the evaluation of detection solutions in reliable and time-efficient manner. Paper IV, Paper V and Paper VI propose detection solutions that can be used in identifying malicious network traffic at different points in network and based on diverse traffic analysis principles. The solutions presented in Paper IV and Paper V can be used for identifying botnet network traffic at local and enterprise networks while the solution presented in Paper VI can be used for identifying potentially compromised clients in large-scale ISP networks. The solution presented in Paper VI captures a wider subset of malicious traffic by covering DNS traffic used by malware and botnets but also DNS traffic used for facilitating scam and spamming campaigns. As the proposed detection solutions target different traits of malicious traffic and as they are developed to monitor traffic at different points in network they could be used within a future collaborative botnet protection approach that would be developed based on the principles presented in Paper I.

6.1 Summary

Research question 1 - The first research question highlights the need for a collaborative multifaceted approach to botnet protection. We have addressed the research question by introducing ContraBot - a novel framework for collaborative botnet protection in Paper I.

Paper I stresses that complex threats such as modern malware manifest them self in a number of forms and that there are various opportunities for identifying existence of compromised computers. Furthermore, the paper highlights the fact that there is no "silver bullet" in botnet detection and that all detection approaches are vulnerable to evasion by the attacker to smaller or larger degree. Therefore, the paper concludes that effective detection should incorporate a number of available analysis solutions in order to cover different aspects of botnet operation and thus limit the possibilities of evading detection. The proposed system should include different detection entities varying from network traffic analysis, behavioral analysis of malware to static code analysis.

Research question 2 - The second research question addresses the challenges of using machine learning-based approaches and the ways of overcoming them. We addressed the research question by Paper II and Paper III that have goals of putting more light on the use of MLAs in existing detection methods and solving ground truth problem as one of the crucial challenges of the use of MLAs.

Paper II brings a number of conclusions regarding the use of MLAs by the existing detection methods. First, detection solutions should specially consider analysis perspective so that the results of detection would provide the operator with an insightful outlook in the state of the network, instead of reporting a yet another alarm. Second, detection solutions should put emphasis on limiting detection errors and especially tackling the problem of high number of false positives. Third, there is a need for more thorough evaluation of existing detection methods using traffic traces from more diverse malware samples and diverse benign applications as well as the need for reliable methods and tools for obtaining the ground truth on malicious and benign traffic.

Paper III concludes that labeling used by existing DNS-based solutions often produces sub-optimal results and that there is a clear need for more reliable approach for obtaining the ground truth on agile DNS traffic. Furthermore, the used domain-to-IPs analysis perspective contributes to the better understanding of the nature of analyzed DNS traffic and the discovery of a wider set of potentially malicious domains-to-IPs mappings. Finally, the paper concludes that human insight is invaluable for obtaining reliable ground truth and that one of the goals of novel labeling approaches should be including the human insight in time-efficient manner.

Research question 3 - The third research question tackles the problem of identifying botnets at local and enterprise networks using the principle of network traffic classification. We have proposed novel approaches for identi-

fying botnet network activity based on network traffic classification in Paper IV and Paper V.

Paper IV evaluates the use of eight supervised MLAs and the flow-based traffic analysis for the identification of botnet traffic at local and enterprise networks. The paper concludes that the employed principles of traffic analysis can provide classification performance in line with the contemporary approaches but with limited amount of traffic analyzed per flow. Furthermore, the paper concludes that the optimal detection performance and time requirements of classification can be achieved using tree based classifiers.

Paper V evaluated three traffic classifiers targeted at identifying botnet TCP, UDP and DNS traffic. We evaluated the three classifiers with some of the most extensive botnet data sets achieving promising classification results. The main conclusion of the paper is that by using separate classifiers for the three protocols it is possible to obtain more fine grained classification that consequently leads to more accurate classification in comparison to work presented in Paper IV.

Research question 4 - The fourth research question tackles the problem of detecting malicious network activity in ISP networks. We address this research question by introducing a novel method for identifying potentially compromised clients based on DNS traffic analysis at large-scale ISP network. The method is presented in Paper VI.

Paper VI concludes on several points. First, the paper concludes on the great benefit of domains-to-IPs analysis perspective that offers both better contextualization of the detection results and the possibility for network operator to manually analyze detection results and correct any errors that may have occurred. Second, the paper concludes on the promising ability of the proposed domains-to-IPs mappings classifier to accurately identify malicious mappings. Third, the paper concludes on the possibilities of efficiently pinpointing the potentially compromised clients based on particular malicious domains-to-IPs mappings whose domain names clients resolved.

6.2 Discussion

The methods presented in this thesis have promising perspectives of being implemented in operational networks. However, the methods also come with challenges that need to be thoroughly understood in order for methods to be effectively used.

The novel DNS traffic labeling approach proposed by Paper III is developed considering the use in operational networks. The approach relies on domains-to-IPs mappings perceptive that is suitable for analysis by a human operator as it yields a reasonable number of mappings when analyzing DNS

traffic from an ISP network. The approach incorporates operator's insight in the labeling process in time-efficient manner which makes it a great tool for security practitioners that aim at obtaining reliable ground truth on analyzed DNS traffic. Finally, the method has been evaluated using network trace from a regional ISP operator and based on the analysis the labeling approach could be scaled to network several times bigger still keeping the operator's insight at a reasonable scale.

Network traffic classifiers presented by Paper IV and Paper V show encouraging perspectives in being used for botnet detection at local and enterprise networks. The performance of the proposed approaches in terms of computational requirements and time-efficiency indicate the possibilities of using the proposed concepts for real-time detection at traffic load that could be expected at enterprise networks using even of-the-shelf computers. Classification performance are also promising but still require further improvements in order for the classifiers to be effectively used in operational environment. For classification methods presented in Paper V the number of false positives averages at 1-2% which is on pair with existing work. However, this needs to be addressed aiming at zero false positives before the classifiers could be moved into operational environments.

Finally, detection approach proposed by Paper VI is based on the similar principles of traffic analysis as labeling approach presented in Paper III and thus was developed with the operational use in mind. The performance of the systems is suitable for carrying out per-week analysis of ISP network traffic and extracting a set of client machines (Internet endpoints) from which problematic domains have been queried. The performance of the system was evaluated using an off-the-shelf computer indicating possibilities for further performance improvements. Regarding the identification performance, the detection system still has a noticeable number of falsely identified domains-to-IPs mappings that need to be further minimized in order to use the full potential of the system. However, even if the proposed system produces false positives due to the nature of the used analysis perspective and the relatively low number of agile mappings these errors could be noticed and eliminated by the operator of the system.

6.3 Future Work

The future work will be devoted to several tracks. First, one of the primary goals should be bringing to life the collaborative detection frameworks presented in Paper I. The collaborative approach could be based on the solutions for detecting botnets at enterprise and ISP networks proposed by Paper IV, Paper V and Paper VI, as well as additional client-based detection solutions. For the realization of the client-based detection solution we can rely on some of our work on identifying malware types and families [120, 121] based on

client-level behavioral analysis. However, such a collaborative system would require a wide coalition of ISPs, AV vendors and end users in order to fulfill its potentials. This could potentially be done through future nationwide or EU projects. Second, the detection approaches proposed in papers Paper IV, Paper V and Paper VI should be further developed in order to provide more precise detection. This could be done by optimizing the principles of network traffic analysis through feature engineering and optimization of used MLAs. Furthermore, as these methods rely on supervised MLAs that is dependent on the training data sets additional traffic traces should be used for training the classifiers. This is especially important for the approach presented in Paper VI as we attribute the majority of falsely classified instances to the lack of training data. Third, the labeling approach proposed in Paper III should be further improved by optimizing the traffic analysis used by it in order to further minimize human involvement in the process of DNS traffic labeling.

References

[1] ITU-International Telecommunication Union, "Measuring the Information Society Report 2015," ITU, Tech. Rep., 2015.

[2] P. Cerwall (ed.), "Ericsson Mobility Report on the Pulse of the Networked Society," Tech. Rep., 2015.

[3] Juniper Research, "The Internet of Things: Consumer, Industrial & Public Services 2015-2020," Tech. Rep., Sep. 2015. [Online]. Available: www.juniperresearch.com/press/press-releases/iot-connected-devices-to-triple-to-38-bn-by-2020

[4] M. Stevanovic and J. Pedersen, "Machine learning for identifying botnet network traffic," Aalborg University, Tech. Rep., 2013.

[5] A. P. Felt, M. Finifter, E. Chin, S. Hanna, and D. Wagner, "A survey of mobile malware in the wild," in *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*. ACM, 2011, pp. 3–14. [Online]. Available: http://doi.acm.org/10.1145/2046614.2046618

[6] N. Falliere, L. O. Murchu, and E. Chien, "W32. stuxnet dossier," *White paper, Symantec Corp., Security Response*, vol. 5, 2011.

[7] Kaspersky Global Research and Analysis Team, "Kaspersky Security Bulletin 2014," Tech. Rep., 2014.

[8] Wood, P. (ed.), "Internet security threat report 2015," Tech. Rep., 2015.

[9] AV-test, "Number of new malware samples," aug 2015. [Online]. Available: https://www.av-test.org/en/statistics/malware/

[10] Alcatel-Lucent Motive Security Labs, "Motive Security Labs Malware Report – H2 2014," Tech. Rep., 2014.

[11] Intel Security, "Net Losses: Estimating the Global Cost of Cybercrime - Economic impact of cybercrime II," Tech. Rep., jun 2014.

[12] Ponemon Institute LLC, "The Cost of Malware Containment," Tech. Rep., jan 2015.

[13] Hogben, G. (ed.), "Botnets: Detection, measurement, disinfection and defence," ENISA, Tech. Rep., 2011.

[14] S. S. Silva, R. M. Silva, R. C. Pinto, and R. M. Salles, "Botnets: A survey," *Computer Networks*, vol. 57, no. 2, pp. 378–403, 2013.

[15] P. R. Marupally and V. Paruchuri, "Comparative analysis and evaluation of botnet command and control models," in *Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications*, ser. AINA '10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 82–89.

[16] H. Zeidanloo and A. Manaf, "Botnet command and control mechanisms," in *Computer and Electrical Engineering, 2009. ICCEE '09. Second International Conference on*, vol. 1, dec. 2009, pp. 564 –568.

[17] D. Dittrich and S. Dietrich, "P2P as botnet command and control: a deeper insight," in *Proceedings of the 3rd International Conference On Malicious and Unwanted Software (Malware 2008)*, 2008, pp. 46–63.

[18] M. A. Rajab, J. Zarfoss, F. Monrose, and A. Terzis, "My botnet is bigger than yours (maybe, better than yours): why size estimates remain challenging," in *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, ser. HotBots'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 5–5.

[19] S. Liu, J. Gong, W. Yang, and A. Jakalan, "A survey of botnet size measurement," in *Networking and Distributed Computing (ICNDC), 2011 Second International Conference on*, sept. 2011, pp. 36 –40.

[20] K. Thomas. (2015, Feb.) Nine bad botnets and the damage they did. [Online]. Available: http://www.welivesecurity.com/2015/02/25/nine-bad-botnets-damage/

[21] M. Feily and Shahrestani, "A survey of botnet and botnet detection," *Emerging Security Information, Systems and Technologies, 2009. SECURWARE '09. Third International Conference on*, pp. 268–273, 2009.

[22] J. B. Grizzard, V. Sharma, C. Nunnery, B. B. Kang, and D. Dagon, "Peer-to-peer botnets: overview and case study," in *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, ser. HotBots'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 1–1.

[23] P. Wang, S. Sparks, and C. C. Zou, "An advanced hybrid peer-to-peer botnet," in *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, ser. HotBots'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 2–2.

[24] Z. Zhang, B. Lu, P. Liao, C. Liu, and X. Cui, "A hierarchical hybrid structure for botnet control and command," in *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on*, vol. 1, june 2011, pp. 483 –489.

[25] P. Maymounkov and D. Mazières, "Kademlia: A peer-to-peer information system based on the xor metric," in *Revised Papers from the First International Workshop on Peer-to-Peer Systems*, ser. IPTPS '01. London, UK, UK: Springer-Verlag, 2002, pp. 53–65.

[26] J. Pouwelse, P. Garbacki, D. Epema, and H. Sips, "The Bittorrent P2P File-Sharing System: Measurements and Analysis," in *Peer-to-Peer Systems IV*, ser. Lecture Notes in Computer Science, M. Castro and R. Renesse, Eds. Springer Berlin Heidelberg, 2005, vol. 3640, pp. 205–216.

[27] K. Kutzner and T. Fuhrmann, "Measuring large overlay networks — the overnet example," in *Kommunikation in Verteilten Systemen (KiVS)*, ser. Informatik aktuell, P. Müller, R. Gotzhein, and J. Schmitt, Eds. Springer Berlin Heidelberg, 2005, pp. 193–204.

[28] G. Sinclair, C. Nunnery, and B.-H. Kang, "The waledac protocol: The how and why," in *Malicious and Unwanted Software (MALWARE), 2009 4th International Conference on*, oct. 2009, pp. 69 –77.

[29] Forbes. (2014, Jan.) The Big Data Breaches of 2014. [Online]. Available: http://www.forbes.com/sites/moneybuilder/2015/01/13/the-big-data-breaches-of-2014/

[30] Wikipedia. (2016, Jan.) Sony pictures entertainment hack. [Online]. Available: https://en.wikipedia.org/wiki/Sony_Pictures_Entertainment_hack

[31] Business Insider UK. (2014, Dec.) Sony Hackers Have Over 100 Terabytes Of Documents. Only Released 200 Gigabytes So Far. [Online]. Available: http://uk.businessinsider.com/the-sony-hackers-still-have-a-massive-amount-of-data-that-hasnt-been-leaked-yet-2014-12

[32] J. Stewart. (2008, Apr.) Top spam botnets exposed. [Online]. Available: http://www.secureworks.com/cyber-threat-intelligence/threats/topbotnets/

[33] Arbor Networks, "Worldwide Infrastructure Security Report 2015," Tech. Rep., jan 2015.

[34] M. Egele, T. Scholte, E. Kirda, and C. Kruegel, "A survey on automated dynamic malware-analysis techniques and tools," *ACM Comput. Surv.*, vol. 44, no. 2, pp. 6:1–6:42, Mar. 2008.

[35] I. You and K. Yim, "Malware obfuscation techniques: A brief survey," in *Broadband, Wireless Computing, Communication and Applications (BWCCA), 2010 International Conference on*, nov. 2010, pp. 297 –300.

[36] J. Marpaung, M. Sain, and H.-J. Lee, "Survey on malware evasion techniques: State of the art and challenges," in *Advanced Communication Technology (ICACT), 2012 14th International Conference on*, feb. 2012, pp. 744 –749.

[37] E. Stinson and J. C. Mitchell, "Towards systematic evaluation of the evadability of bot/botnet detection methods," in *Proceedings of the 2nd conference on USENIX Workshop on offensive technologies*, ser. WOOT'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 5:1–5:9.

[38] B. Blunden, *The Rootkit Arsenal: Escape and Evasion in the Dark Corners of the System*. USA: Jones and Bartlett Publishers, Inc., 2009.

[39] T. Arnold and T. A. Yang, "Rootkit attacks and protection: a case study of teaching network security," *J. Comput. Sci. Coll.*, vol. 26, no. 5, pp. 122–129, May 2011.

[40] T. Holz, C. Gorecki, K. Rieck, and F. C. Freiling, "Measuring and detecting fast-flux service networks," in *Proceedings of the Network and Distributed System Security Symposium, NDSS 2008, San Diego, California, USA, 10th February - 13th February 2008*, 2008.

[41] M. Antonakakis, J. Demar, C. Elisan, and J. Jerrim, "DGAs and Cyber-Criminals: A Case Study," Damballa Inc, Tech. Rep., 2012.

[42] A. Berger and W. N. Gansterer, "Modeling DNS agility with DNSMap," in *INFOCOM*, 2013, pp. 3153–3158.

[43] J. Wyke, "The ZeroAccess botnet - Mining and fraud for massive financial gain," *Sophos Technical Paper*, 2012.

[44] M. Stockley. (2015, january) ZeroAccess click fraud botnet coughs back to life. [Online]. Available: https://nakedsecurity.sophos.com/2015/01/31/zeroaccess-click-fraud-botnet-coughs-back-to-life/

[45] F. Howard, "Exploring the Blackhole exploit kit," *Sophos Technical Paper*, 2012.

[46] M. Masud, L. Khan, and B. Thuraisingham, *Data Mining Tools for Malware Detection*. Taylor & Francis Group, 2011.

[47] M. Bailey, J. Oberheide, J. Andersen, Z. M. Mao, F. Jahanian, and J. Nazario, "Automated classification and analysis of internet malware." in *RAID*, ser. Lecture Notes in Computer Science, C. Krügel, R. Lippmann, and A. Clark, Eds., vol. 4637. Springer, 2007, pp. 178–197.

[48] E. Stinson and J. C. Mitchell, "Characterizing bots' remote control behavior," in *Botnet Detection*, ser. Advances in Information Security, W. Lee, C. Wang, and D. Dagon, Eds. Springer, 2008, vol. 36, pp. 45–64.

[49] L. Liu, S. Chen, G. Yan, and Z. Zhang, "BotTracer: Execution-Based Bot-Like Malware Detection," in *Proceedings of the 11th international conference on Information Security*, ser. ISC '08. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 97–113.

[50] C. Kolbitsch, P. M. Comparetti, C. Kruegel, E. Kirda, X. Zhou, and X. Wang, "Effective and efficient malware detection at the end host," in *Proceedings of the 18th conference on USENIX security symposium*, ser. SSYM'09. Berkeley, CA, USA: USENIX Association, 2009, pp. 351–366.

[51] U. Bayer, P. M. Comparetti, C. Hlauschek, C. Krügel, and E. Kirda, "Scalable, behavior-based malware clustering." in *NDSS*. The Internet Society, 2009, pp. 5–5.

[52] Y. Park, Q. Zhang, D. Reeves, and V. Mulukutla, "AntiBot: Clustering Common Semantic Patterns for Bot Detection," in *Computer Software and Applications Conference (COMPSAC), 2010 IEEE 34th Annual*, july 2010, pp. 262 –272.

[53] M. Masud, T. Al-khateeb, L. Khan, B. Thuraisingham, and K. Hamlen, "Flow-based identification of botnet traffic by mining multiple log files," in *Distributed Framework and Applications, 2008. DFmA 2008. First International Conference on*, oct. 2008, pp. 200–206.

[54] S. Shin, Z. Xu, and G. Gu, "Effort: Efficient and effective bot malware detection," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 2846–2850.

[55] Y. Zeng, X. Hu, and K. Shin, "Detection of botnets using combined host- and network-level information," in *Dependable Systems and Networks (DSN), 2010 IEEE/IFIP International Conference on*, 28 2010-july 1 2010, pp. 291 –300.

[56] S. García, A. Zunino, and M. Campo, "Survey on network-based botnet detection methods," *Security and Communication Networks*, vol. 7, no. 5, pp. 878–903, 2014.

[57] H. Wang, J. Hou, and Z. Gong, "Botnet detection architecture based on heterogeneous multi-sensor information fusion," *Journal of Networks*, vol. 6, no. 12, pp. 1655–1661, Dec. 2011.

[58] K. Muthumanickam and E. Ilavarasan, "P2P botnet detection: combined host- and network-level analysis," in *Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on*. IEEE, 2012, pp. 1–5.

[59] Y. He, Q. Li, Y. Ji, and D. Guo, "BotInfer: A Bot Inference Approach by Correlating Host and Network Information," in *Network and Parallel Computing*. Springer, 2013, pp. 356–367.

[60] T. Holz, C. Gorecki, K. Rieck, and F. C. Freiling, "Detection and mitigation of fast-flux service networks," in *NDSS'08: Proceedings of the 15th Annual Network and Distributed System Security Symposium, San Diego*. Internet Society, Feb. 2008.

[61] Damballa Inc, "A New iteration of the TDSS/TDL4 Malware Using DGA-based command-and-control," Damballa, Tech. Rep., 2012.

[62] G. Gu, V. Yegneswaran, P. Porras, J. Stoll, and W. Lee, "Active botnet probing to identify obscure command and control channels," in *Proceedings of the 2009 Annual Computer Security Applications Conference*, ser. ACSAC '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 241–253.

[63] M. Roesch, "Snort - lightweight intrusion detection for networks," in *Proceedings of the 13th USENIX conference on System administration*, ser. LISA '99. Berkeley, CA, USA: USENIX Association, 1999, pp. 229–238.

[64] V. Paxson, "Bro: a system for detecting network intruders in real-time," *Computer Networks*, vol. 31, no. 23-24, pp. 2435 – 2463, 1999.

[65] Open Information Security Foundation, "Suricata: open-source IDS/IPS/NSM engine," 2015.

[66] G. Gu, P. Porras, V. Yegneswaran, M. Fong, and W. Lee, "BotHunter: Detecting malware infection through IDS-driven dialog correlation," in *Proceedings of the 16th USENIX Security Symposium, San Jose, California*. USENIX Association, Jul. 2007, pp. 167–182.

[67] J. Goebel and T. Holz, "Rishi: identify bot contaminated hosts by IRC nickname evaluation," in *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, ser. HotBots'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 8–8.

[68] J. R. Binkley and S. Singh, "An algorithm for anomaly-based botnet detection," in *Proceedings of USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI)*, 2006, pp. 43–48.

[69] G. Gu, J. Zhang, and W. Lee, "BotSniffer: Detecting botnet command and control channels in network traffic," in *Proceedings of the 15th Annual Network and Distributed System Security Symposium (NDSS'08)*, February 2008, pp. 1–1.

[70] J. Kang, J.-Y. Zhang, Q. Li, and Z. Li, "Detecting new P2P botnet with multi-chart CUSUM," in *Networks Security, Wireless Communications and Trusted Computing, 2009. NSWCTC'09. International Conference on*, vol. 1. IEEE, 2009, pp. 688–691.

[71] W. Wang, B. Fang, Z. Zhang, and C. Li, "A novel approach to detect IRC-based botnets," in *Networks Security, Wireless Communications and Trusted Computing, 2009. NSWCTC'09. International Conference on*, vol. 1. IEEE, 2009, pp. 408–411.

[72] X. Ma, X. Guan, J. Tao, Q. Zheng, Y. Guo, L. Liu, and S. Zhao, "A novel IRC botnet detection method based on packet size sequence," in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[73] B. Al-Duwairi and L. Al-Ebbini, "BotDigger: a fuzzy inference system for botnet detection," in *Internet Monitoring and Protection (ICIMP), 2010 Fifth International Conference on*. IEEE, 2010, pp. 16–21.

[74] K. Wang, C.-Y. Huang, S.-J. Lin, and Y.-D. Lin, "A fuzzy pattern-based filtering algorithm for botnet detection," *Computer Networks*, vol. 55, no. 15, pp. 3275–3286, 2011.

[75] X. Yu, X. Dong, G. Yu, Y. Qin, D. Yue, and Y. Zhao, "Online botnet detection based on incremental discrete fourier transform," *JNW*, vol. 5, no. 5, pp. 568–576, 2010.

[76] A. Karasaridis, B. Rexroad, and D. Hoeflin, "Wide-scale botnet detection and characterization," in *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, ser. HotBots'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 7–7.

[77] R. Villamarín-Salomón and J. C. Brustoloni, "Bayesian bot detection based on DNS traffic similarity," in *Proceedings of the 2009 ACM symposium on Applied Computing*, ser. SAC '09. New York, NY, USA: ACM, 2009, pp. 2035–2041.

[78] A. Ramachandran, N. Feamster, and D. Dagon, "Revealing botnet membership using DNSBL counter-intelligence," in *Proceedings of the 2nd conference on Steps to Reducing Unwanted Traffic on the Internet - Volume 2*, ser. SRUTI'06. Berkeley, CA, USA: USENIX Association, 2006, pp. 49–54.

[79] D. Dagon, C. Zou, and W. Lee, "Modeling botnet propagation using time zones," in *In Proceedings of the 13 th Network and Distributed System Security Symposium NDSS*, 2006, pp. 7–7.

[80] S. Dua and X. Du, *Data mining and machine learning in cybersecurity.* Boca Raton, FL: CRC Press. xxii, 234 p. $ 89.95 , 2011.

[81] T. M. Mitchell, *Machine Learning*, 1st ed. New York, NY, USA: McGraw-Hill, Inc., 1997.

[82] S. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Frontiers in Artificial Intelligence and Applications*, vol. 160, p. 3, 2007.

[83] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.

[84] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Security and Privacy (SP), 2010 IEEE Symposium on*. IEEE, 2010, pp. 305–316.

[85] A. J. Oliner, A. V. Kulkarni, and A. Aiken, "Community epidemic detection using time-correlated anomalies," in *Recent Advances in Intrusion Detection*. Springer, 2010, pp. 360–381.

[86] A. Flaglien, K. Franke, and A. Årnes, "Identifying malware using cross-evidence correlation," in *Advances in Digital Forensics VII*, ser. IFIP ACIT, G. Peterson and S. Shenoi, Eds. IFIP, 2011, vol. 361, ch. 13, pp. 169–182.

[87] F. Cuppens and A. Miège, "Alert correlation in a cooperative intrusion detection framework"," in *"Proceedings of IEEE Symposium on Security and Privacy"*, "may" "2002", pp. "202–215".

[88] P. Ning, Y. Cui, and D. S. Reeves, "Constructing attack scenarios through correlation of intrusion alerts," in *Proceedings of CCS'02*. ACM, Nov. 2002, pp. 245–254.

[89] Metaflows. (2015, Sep.). [Online]. Available: https://www.metaflows.com/

[90] C. Livadas, R. Walsh, D. Lapsley, and W. Strayer, "Usilng machine learning technliques to identify botnet traffic," in *Local Computer Networks, Proceedings 2006 31st IEEE Conference on*, nov. 2006, pp. 967 –974.

[91] W. T. Strayer, D. Lapsely, R. Walsh, and C. Livadas, "Botnet detection based on network behaviour," in *Botnet Detection*, ser. Advances in Information Security, W. Lee, C. Wang, and D. Dagon, Eds. Springer, 2008, vol. 36, pp. 1–24.

[92] G. Gu, R. Perdisci, J. Zhang, and W. Lee, "BotMiner: Clustering analysis of network traffic for protocol- and structure-independent botnet detection," in *Proceedings of the 17th conference on Security symposium*, 2008, pp. 139–154.

[93] H. Choi and H. Lee, "Identifying botnets by capturing group activities in DNS traffic," *Computer Networks*, vol. 56, no. 1, pp. 20–33, 2012.

[94] S. Saad, I. Traore, A. Ghorbani, B. Sayed, D. Zhao, W. Lu, J. Felix, and P. Hakimian, "Detecting P2P botnets through network behavior analysis and machine learning," in *Privacy, Security and Trust (PST), 2011 Ninth Annual International Conference on*, july 2011, pp. 174 –180.

[95] D. Zhao, I. Traore, B. Sayed, W. Lu, S. Saad, A. Ghorbani, and D. Garant, "Botnet detection based on traffic behavior analysis and flow intervals," *Computers & Security*, vol. 39, pp. 2–16, 2013.

[96] J. Zhang, R. Perdisci, W. Lee, U. Sarfraz, and X. Luo, "Detecting stealthy P2P botnets using statistical traffic fingerprints," in *2011 IEEE/IFIP 41st International Conference on Dependable Systems and Networks (DSN), Hong Kong*. IEEE/IFIP, Jun. 2011, pp. 121–132.

[97] W. Lu, G. Rammidi, and A. A. Ghorbani, "Clustering botnet communication traffic based on n-gram feature selection," *Computer Communications*, vol. 34, pp. 502–514, 2011.

[98] L. Bilge, D. Balzarotti, W. Robertson, E. Kirda, and C. Kruegel, "Disclosure: detecting botnet command and control servers through large-scale netflow analysis," in *Proceedings of the 28th Annual Computer Security Applications Conference*, ser. ACSAC '12. New York, NY, USA: ACM, 2012, pp. 129–138.

[99] L. Bilge, S. Sen, D. Balzarotti, E. Kirda, and C. Kruegel, "EXPOSURE: a passive DNS analysis service to detect and report malicious domains," *ACM Transactions on Information and System Security (TISSEC)*, vol. 16, no. 4, p. 14, 2014.

[100] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster, "Building a dynamic reputation system for DNS," in *Proceedings of the 19th USENIX conference on Security*, ser. USENIX Security'10. Berkeley, CA, USA: USENIX Association, 2010, pp. 273–290.

[101] M. Antonakakis, R. Perdisci, W. Lee, N. Vasiloglou II, and D. Dagon, "Detecting Malware Domains at the Upper DNS Hierarchy," in *USENIX Security Symposium*, 2011, p. 16.

[102] R. Perdisci, I. Corona, and G. Giacinto, "Early Detection of Malicious Flux Networks via Large-Scale Passive DNS Traffic Analysis," *Dependable and Secure Computing, IEEE Transactions on*, vol. 9, no. 5, pp. 714–726, 2012.

[103] F. Tegeler, X. Fu, G. Vigna, and C. Kruegel, "BotFinder: finding bots in network traffic without deep packet inspection," in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*. ACM, 2012, pp. 349–360.

[104] D. Zhao and I. Traore, "P2P botnet detection through malicious fast flux network identification," in *P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2012 Seventh International Conference on*. IEEE, 2012, pp. 170–175.

[105] J. Zhang, R. Perdisci, W. Lee, X. Luo, and U. Sarfraz, "Building a scalable system for stealthy P2P-botnet detection," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 1, pp. 27–38, 2014.

[106] F. Haddadi, D. Runkel, A. N. Zincir-Heywood, and M. I. Heywood, "On botnet behaviour analysis using GP and C4.5," in *Proceedings of the 2014 conference companion on Genetic and evolutionary computation companion*. ACM, 2014, pp. 1253–1260.

[107] A. J. Aviv and A. Haeberlen, "Challenges in experimenting with botnet detection systems," in *Proceedings of the 4th conference on Cyber security experimentation and test*, ser. CSET'11. Berkeley, CA, USA: USENIX Association, 2011, pp. 6–6.

[108] D. Zhao, I. Traoré, A. A. Ghorbani, B. Sayed, S. Saad, and W. Lu, "Peer to peer botnet detection based on flow intervals," in *SEC*, 2012, pp. 87–102.

[109] R. Perdisci, I. Corona, D. Dagon, and W. Lee, "Detecting Malicious Flux Service Networks through Passive Analysis of Recursive DNS Traces," in *Proceedings of the 2009 Annual Computer Security Applications Conference*, ser. ACSAC '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 311–320.

[110] S. Sinha, M. Bailey, and F. Jahanian, "Shades of grey: On the effectiveness of reputation-based "blacklists"," in *Malicious and Unwanted Software, 2008. MALWARE 2008. 3rd International Conference on*. IEEE, 2008, pp. 57–64.

[111] S. Sheng, B. Wardman, G. Warner, L. Cranor, J. Hong, and C. Zhang, "An empirical analysis of phishing blacklists," in *Sixth Conference on Email and Anti-Spam (CEAS)*, 2009.

[112] M. Kührer, C. Rossow, and T. Holz, "Paint it black: Evaluating the effectiveness of malware blacklists," in *Research in Attacks, Intrusions and Defenses*. Springer, 2014, pp. 1–21.

[113] C. J. Dietrich and C. Rossow, "Empirical research of ip blacklists," in *ISSE 2008 Securing Electronic Business Processes*. Springer, 2009, pp. 163–171.

[114] Damballa Inc. (2015, Sep.). [Online]. Available: https://www.damballa.com/

[115] M. Antonakakis, R. Perdisci, Y. Nadji, N. Vasiloglou, S. Abu-Nimeh, W. Lee, and D. Dagon, "From throw-away traffic to bots: Detecting the rise of DGA-based malware," in *Proceedings of the 21st USENIX security symposium*, 2012.

[116] S. García, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods," *Computers & Security*, vol. 45, pp. 100–123, 2014.

[117] S. Schiavoni, F. Maggi, L. Cavallaro, and S. Zanero, "Phoenix: DGA-based botnet tracking and intelligence," in *Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer, 2014, pp. 192–211.

[118] P. Luo, R. Torres, Z.-L. Zhang, S. Saha, S.-J. Lee, A. Nucci, and M. Mellia, "Leveraging client-side DNS failure patterns to identify malicious behaviors," in *Communications and Network Security (CNS), 2015 IEEE Conference on*. IEEE, 2015, pp. 406–414.

[119] R. Sharifnya and M. Abadi, "DFBotKiller: Domain-flux botnet detection based on the history of group activities and failures in DNS traffic," *Digital Investigation*, vol. 12, pp. 15–26, 2015.

[120] R. S. Pirscoveanu, S. S. Hansen, T. M. T. Larsen, M. Stevanovic, J. M. Pedersen, and A. Czech, "Analysis of malware behavior: Type classification using machine learning," in *Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), 2015 International Conference on*, June 2015, pp. 1–7.

[121] S. S. Hansen, T. M. T. Larsen, M. Stevanovic, J. M. Pedersen, and A. Czech, "An Approach for Detection and Family Classification of Malware Based on Behavioral Analysis," in *Computing, Networking and Communications (ICNC), 2016 International Conference on*. IEEE.

SUMMARY

This thesis explores how can network traffic analysis be used for accurate and efficient detection of malware network activities. The thesis focuses on botnet detection by devising novel detection approaches that are aimed at identifying malware network activity at different points in the network and based on different, mutually complementary, principles of traffic analysis. The proposed approaches rely on machine learning algorithms (MLAs) for automated and resource-efficient identification of the patterns of malicious network traffic. We evaluated the proposed methods through extensive evaluations using traffic traces from honeypots and malware testing environments as well as operational ISP networks. Based on the evaluation, the novel detection methods provide accurate and efficient identification of malicious network traffic, thus being promising in the light of operational deployment. Furthermore, the thesis provides an overview of some of the biggest challenges of using MLAs for identifying malicious network activities. The challenge specially addressed by the thesis is the "ground truth" problem, where we proposed a novel labeling approach for obtaining the ground truth on agile DNS traffic that provides reliable and time-efficient labeling. Finally, the thesis outlines the opportunities for future work on realizing robust and effective detection solutions.

ISSN (online): 2246-1248
ISBN (online): 978-87-7112-490-3

AALBORG **UNIVERSITETSFORLAG**