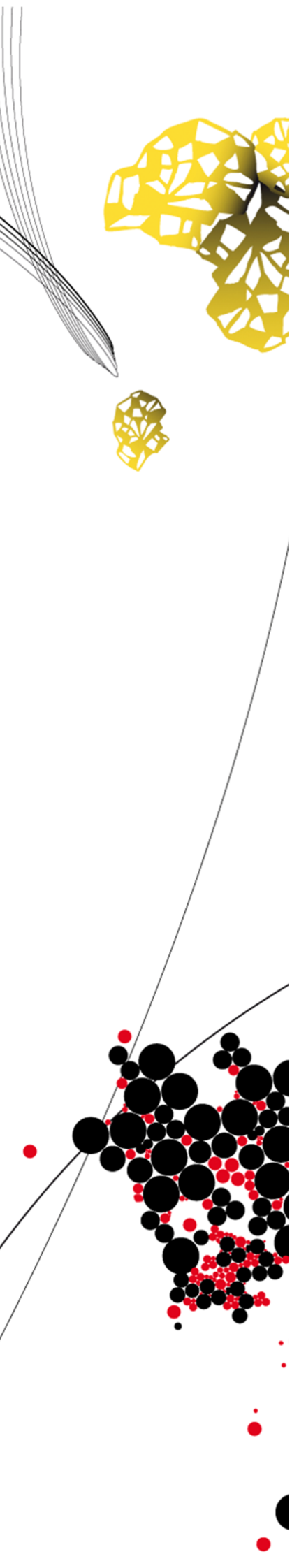


UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering, Mathematics & Computer Science

Analysis of Malicious Domains using Active DNS Data Provided by Blacklists

Raoul Tolud

MSc. Thesis
Feb. 2020

Supervisors:
prof. dr. ir. Aiko Pras
dr. Anna Sperotto
dr. Doina Bucur
Olivier van der Toorn, MSc

Design and Analysis of Communication Systems
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Preface

I want to thank dr. Anna Sperotto for allowing me to join the department of Design and Analysis of Communication Systems (DACS) to conduct research under her supervision. I had a fruitful time working with her and a chance to broaden my scope on the field of internet security. Her continuous support and knowledge guided me towards successfully finishing my master thesis. Furthermore, I would also like to thank Olivier van der Toorn who provided me with useful feedback and insight. Lastly, I would like to thank my parents for supporting me during this journey in pursuing my Master's Degree in Electrical Engineering. Without them, this would have not been possible.

Abstract

With the daily translation of millions of human-readable addresses into IP addresses and 4.5 billion users, the Domain Name System is a very crucial infrastructure. Though the DNS provides us with many benign services, this also comes paired with a lot of DNS abuse, such as: spreading malware, setting up command and control, distributing spam e-mail, hosting spam, and phishing domains. All these can be considered malicious or suspicious domains. In order to identify these malicious domains, many approaches have been proposed based on the use of DNS data. The collection of this DNS data can be separated into passive and active DNS data. The difference between these two methods is that one provides us user-generated DNS data and the other targeted DNS data. In this thesis, we make use of the OpenINTEL measurement platform, which provides active DNS data based on publicly available blacklists. This thesis aims to make a comparison between bad domains extracted from these publicly available blacklists, to see if there are shared properties at the DNS level that can make for a useful signature. This newly found signature or profile can then be used to assist in identifying unlisted malicious domains using the OpenINTEL data set. In this research, we present two main contributions, namely analyzing the difference between a set of DNS features on RBL and Alexa for active DNS data and the analysis of malicious clustering on IP level by adapting the bad neighborhood concept to use domains. In order to analyze the difference between active DNS features on ALEXA and RBL, a set of features were extracted from the active DNS data and analyzed. The results indicated that there is no statistical deviation on most of the features based on the available data set and method used. However, one set of features did show signs of deviation, but this alone is not sufficient to build a valid signature. As a result of this, we investigate if there is any clustering of malicious behavior at the IP level by using the bad neighborhood concept. The bad neighborhood concept is seen as a group of IP addresses that persistently perform malicious activities and are acquired by using a particular aggregation criterion. Due to the nature of our active DNS data, this concept is adapted to domains. In order to adapt this concept and detect the bad neighborhoods within the RBL data set, different approaches are analyzed. Our adapted model of bad neighborhood takes into consideration both the number of hosts in the bad neighborhood and the number of

malicious domains hosted. As a result of this adapted bad neighborhood concept, a detection method was built, which allows us to identify the bad neighborhoods using scatter plots based on host and domain count. To see if this method can function as a standalone method for the detection of malicious domains, a validation in real time was performed. The validation period shows a low number of true positives and a high number of false positives. This could be a result of the lack of validation data resulting in high number of false positives.

Contents

List of acronyms

ASN Autonomous Name System

CIDR Classes Inter-Domain Routing

CDF Cumulative Distribution Function

DNSWL DNS-based white list

DACS Design and Analysis of Communication Systems

DNS Domain Name System

FN False Negative

FP False Positive

IP Internet Protocol

ISP Internet Service Provider

KS KolmogorovSmirnov

RBL Real Time Blacklist

RFC Request For Comments

RHSBL Right Hand Side Blacklist

SLD Second Level Domain

TTL Time To Live

TLD Top Level Domain

TN True Negative

TP True Positive

URIBL Uniform Resource Identifier Blacklist

URL Uniform Resource Locator

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

With the rapid growth of technology and new applications being developed daily, the internet has not been far beyond and has been growing alongside at an exponential rate. Unfortunately, this has further increased the DNS-based attacks on hosts. DNS, or Domain Name System, is a hierarchical decentralized naming system, which maps the human-readable addresses into IP addresses, and is used by hosts to connect to the internet. The DNS has a current estimate of 330.8 million registered domains [3] and 4.5 billion users as of 2020. This makes it very difficult to keep track of all their actions. It is precisely this lack of overview that leads to an increase in the attacks on the hosts making these DNS servers a target of attacks. These domains could be called bad domains. These malicious domains can linked to various malicious activities, such as: spreading of malware, setting up command and control, distributing of spam emails, and hosting phishing websites [4]. In order to identify these bad domains, several approaches have been proposed. A prominent approach is using the passive DNS data [5], a system that monitors DNS queries to and from the authoritative name server. Another less common but promising approach would be using the active DNS data [6], since it provides a more complete view of the DNS, and so domains with malicious intent can be preemptively identified. To obtain this data, a collector is ordered to send DNS queries to a list of targeted domains and record the DNS answers it receives. This list of domains that are being queried in active DNS measurement projects generally use TLD zone files and in some specific cases provided by black and/or white lists. In this thesis multiple domain name blacklists have been merged together and actively queried for their active DNS data. By analyzing active DNS data collected from these blacklists, a comparison is made to identify if they share any properties, which can account for a useful profile or signature. This newly found signature or profile could then be used to assist in identifying unlisted malicious domains using the OpenINTEL data

set. After the initial analysis of the active DNS data, the possibility of adapting the bad neighborhood concept to domains was analyzed. This led to further investigation of disproportionately high behavior of particular subnets in these blacklists and how this concept can be used to detect future malicious domains.

1.2 Thesis goal and research questions

This research aims to make a comparison between bad domains extracted from publicly available blacklists, to see if there are shared properties at the DNS level that can make for a useful signature. If the signature exists it will be used with OpenINTEL data set to detect unlisted malicious domains. To pursue this goal, the following research questions are defined as the base of this thesis:

RQ1: How much statistical difference can be observed between the active DNS data features on ALEXA and RBL?

In order to answer this question, prominent features employed in research on detecting malicious domains using DNS data are analyzed. These features are categorized in DNS record based features (A, AAAA, TXT, NS), network based features, TTL value-based features, and domain name based features. Moreover, these features are analyzed using cumulative distribution plots to observe any significant deviation between the malicious and benign domains on ALEXA and RBL.

RQ2: Can the concept of bad neighborhood be adapted to domains? If yes can we witness any form of bad neighborhoods inside RBL data?

In case the DNS features show no result, we can investigate if there is any clustering of malicious behavior at the IP level. To measure this malicious clustering on the IP level, we use the concept of bad neighborhoods. Due to the use of active DNS data and domain blacklists this concept might have to be adapted in order to witness any form of bad neighborhoods.

RQ3: How effective is the use of domains originating from bad neighborhoods as a valid standalone method to detect future malicious domains ?

In order to see if the adapted bad neighborhood concept can perform as a standalone method, a validation period is performed. This can give insight if domains originating from these neighborhoods can be classified as malicious.

1.3 Overview thesis

After this Introduction, the structure of this research is as follows:

Chapter 2: This chapter provides background information on how the Domain Name System works. Furthermore, it elaborates on the different types of DNS data and the information it provides. This chapter concludes with blacklisting and the types of blacklists that are available.

Chapter 3: State of the art highlights the different techniques and information that can be extracted from DNS data to aid in detection of domains that could be potentially malicious. This chapter concludes with a section on the internet bad neighborhood concept which can be utilized to analyze clustering of malicious activity on IP level.

Chapter 4: This chapter focuses on statistically comparing DNS data features of the RBL (malicious ground truth) and ALEXA (benign ground truth). The features have been chosen from different detection methods mentioned in the state of the art. This chapter will conclude with which features are statically relevant.

Chapter 5: In this chapter, we attempt to verify the presence of bad neighborhood within the actively queried RBL. The presences of bad neighborhoods is validated using different approaches. Finally, the bad neighborhood is adapted to use domains name blacklist to detect unlisted malicious domains.

Chapter ??: In this chapter, the design of the model is described and system needed to extract the suspicious domains. Furthermore, the domains classified by the model as suspicious are validated to measure the standalone performance.

Chapter 7: This chapter presents a summary of the overall conclusions of the thesis.

Chapter 2

Background information

This chapter provides background information on the Domain Name System, DNS data, and blacklisting. Section 2.1 gives an overview of the DNS and how it functions. Section 2.2 elaborates on the use of DNS data and also provides a summary on the types of DNS data. The chapter is concluded with blacklisting in Section 2.3.

2.1 The Domain Name System

The Domain Name System is used to locate hosts on the internet by translating human-readable addresses to IP addresses, since it is easier for host users to remember domain names rather then long number sequences. The DNS does this by providing a mapping to the resources of a domain. These resources are called resource records and are elaborated on in the following section. The previous naming system used a text file called Host.txt, which faced many problems. The problems mainly consisted of scaling and reliability of this system, since it was initially built for a small group of users. This system was then replaced with the DNS in the 1970s under the RFC 1034 [7] and 1035 [8]. The DNS was a worthy successor of the Host.txt by fixing the scaling and reliability problems with its inherent features. Firstly, it is globally distributed, meaning that no single host contains all DNS data, and any device can access these records with the use of the DNS Lookups. Secondly, the data is locally cache-able, resulting in improved performance. Furthermore, having multiple masters and slaves allows for better resilience and load balancing resulting in the capability of handling a significantly higher number of queries. Finally, data is replicated from the master to multiple slaves and can be queried by all clients.

2.1.1 Architecture

The DNS environment is build up out of 3 components: a client, name server, and resolver; all three components together make the physical end of the DNS architec-

ture [9]. Application on the host can access the Domain Name System through the use of a resolver. The used resolver contacts the DNS name server that the host needs to access, and the DNS server then returns the IP address to the resolver and forwards it to the host. Such a schematic is shown below in Figure 2.1.

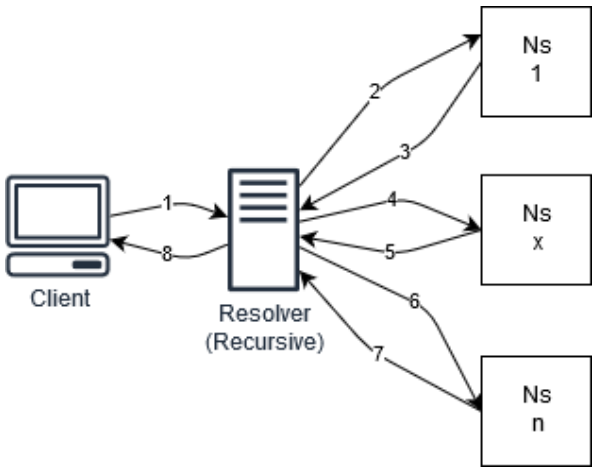


Figure 2.1: The DNS resolution process

Client An application on a host (client) accesses the DNS through a DNS client.

Resolver The Resolver contacts the DNS Server, also known as the name server. The DNS resolver is the first step in the DNS Lookup. The resolver will take the requested domain from the client and make a sequence of queries until the URL has been translated to an IP address.

DNS server The DNS server resolves the host name, which is passed along by the resolver to an IP address. The information of all the IP addresses in the DNS is held by 13 DNS root name servers run by different institutions.

In order to see how the human-readable addresses are translated into IP addresses, the resolution process is explained in Figure 2.1. The steps taken in the resolution process go as followed:

1. The client sends a query out to the resolver e.g. www.google.com.

2. The resolver redirects the query to one of the name servers in the root zone. The IP addresses of the root servers are static and hard coded inside the resolver.

3. The name server responds with a redirection to an authoritative name server in question in this case being .com. If the resolver is not recursive, it would directly respond to the client with the path.

4. A recursive resolver keeps on with finding the path needed for the client to gain access to the query sent.

5. The resolver, therefore, sends a query to the authoritative name server achieved in the previous step, which could be then followed by a number of queries until it obtains the authoritative name server of the domain in question.

6. If the resolver finds the authoritative name server of the domain in question it sends a final query.

7. The name server then responds with a query containing the IP address for the domain.

8. The resolver then redirects the response of the name server to the client.

2.1.2 DNS records

Each domain contains resource records that provide information and are analogous to files. These records are classified into different types depending on the information that is requested. An example of commonly used records is TXT, NS, A, and MX. The information that these records contain is defined in the zone files. The zones files are text-based files that are stored on the DNS server. The resource records will be crucial in active DNS analysis due to the collector actively probing domains for their DNS records and will provide data that can be used to identify the behavior of domains. The resource records mentioned in Tabel 2.1 are collected by the OpenINTEL measurement platform and will be sued in this thesis.

In , the resource records collected by the OpenINTEL measurement platform will be used in this thesis.

Table 2.1: Resource Records queried by OpenINTEL [2]

Resource Record	Description
SOA	The Start Of Authority record specifies key parameters for the DNS zone that reflect operational practices of the DNS operator.
A	Specifies the IPv4 address for a name.
AAAA	Specifies the IPv6 address for a name.
NS	Specifies the names of the authoritative name servers for a domain.
MX	Specifies the names of the hosts that handle e-mail for a domain.
TXT	It contains arbitrary text strings. This record type is used to convey among other things information required for spam filtering and is also often used to prove control over a domain to e.g. cloud and certificate authorities.
DNSKEY	Specifies public keys for validating DNSSEC signatures in the DNS zone.
DS	The Delegation Signer record references a DNSKEY using a cryptographic hash. It is part of the delegation in a parent zone, together with the NS and establishes the chain of trust from parent to child DNS zones in DNSSEC.
NSEC(3)	Used in DNSSEC to provide authenticated denial-of-existence, i.e. to cryptographically prove that a queried name and record type do not exist.
CAA	Specifies which certificate authorities are allowed to issue certificates to a domain.
CDS	Provides information about a signed zone file.
CDNSKEY	We only resolve these records for DNSSEC-signed domains for which at least a DNSKEY or DS record exists. All response records, including full CNAME expansions and RRSIG signature records, are stored.
SPF	Specifies spam filtering information for a domain. Note that this record type was deprecated in 2014 (RFC 7208), we query it to study the decline of an obsolete record type over time

2.2 The Domain Name System data

Detection of malicious domains can have various approaches, such as the analysis of the traffic network, an inspection of web content, URL scrutiny, or a hybrid of these methods. One method that has gained more popularity in the last decade is the use of DNS data. The use of this method proposes several benefits. First of all, it is very scale-able due to DNS data making only a small part of all the network traffic. Secondly, DNS data provides more insightful information on the domains linked to malicious activities. Thirdly, the features extracted from DNS data can be further enriched with the use of supplementary information.

2.2.1 Active DNS data

Active DNS data is obtained by a collector that sends out DNS queries to a targeted list of domains and then records the responses received. The list that is being queried is built out of different sources, including blacklists, ALEXA Top Sites, and zone files of different authoritative servers. The queries issued by the collector do not reflect behavior. Instead of capturing user-generated behavior, it captures the DNS records of domains which are targeted.

2.2.2 Passive DNS data

Passive DNS data collects data by deploying sensors in front of DNS servers or by monitoring DNS server logs to obtain queries and responses. Therefore, passive DNS data gives a more scoped view and is more focused on the user based activity.

2.2.3 Active vs passive DNS data

When collecting data for analyzing associations between DNS features, it can be done in two ways. One could be by actively querying a large group of domains to obtain information. Another way is to passively observe all requests send and receive by DNS servers and extract the necessary information. However, both methods present their own pros and cons, and each method has its place depending on the type of malicious activity that is being detected. Most detection methods use passive DNS data for detection of malicious activity [5]. It is shown that research relying on passive DNS data often focuses on security [2]. Passive DNS data provides us with data captured at the internal interface of the resolver, which provides detailed information about the queries and responses of users; this may directly link to certain types of malicious activities. This also allows for a more personalized detection method for the network being monitored. The downside of the approach is that it provides a scoped view of the malicious activity limited to internal interfaces. Having access to the internal interface of an ISP could partly solve the problem, but this kind of access is not easy to attain. To attain this broader view of the DNS, the use of active DNS data proves to be very beneficial. Although active DNS data does not reflect the usage of behavior, it does allow the collector to control which domains should be queried, giving it a more general view of the DNS. Another benefit is that the data is easy to use due to it not containing user-level behavior making it more accessible for research. The challenge that both methods face is that the setup of these collectors is not an easy task, especially when actively querying multiple domains daily. Even though setting up DNS traffic sensors is relatively more straightforward, the data collected only offers a limited view of the threat monitored.

2.3 Blacklisting

A blacklist is a reputation list of domains or IP addresses that are denied access to certain or all parts of the network. The listed domains or IP addresses are added because there have been multiple instances where malicious activity has been detected and reported. If a node or a set of nodes has been recorded to display malicious behavior, the administrator of such a network would isolate these by putting

them on a blacklist, so removing them from having access to the network. The black-listing can be done based on URLs, domain names, IP addresses, and is applied in different parts of the network such as the DNS-servers, mail-servers, and firewalls. One downside of using blacklist is that once an entity has been blacklisted, the same IP address or domain name cannot be reused until it is removed from the blacklist.

2.3.1 Types of blacklist

As previously mentioned, the type of blacklist used is dependant on the threat and on the access control available in the network. Blacklists generally consist of IP addresses or domain names, where either reported malicious domains, IP addresses,or URLs are listed. Based on these, the following blacklist has been defined:

DNSWL

DNSWL or DNS White List consists of a list of IP addresses or domain names that are known to display good behavior. White-listing prevents users from visiting websites outside the white list, e.g. DNSWL [10]. This type of listing can prove to be use full in defense against malware that deploys domain generation algorithms. It is safe to assume that these automatically generated domain names are not going to show up in the ALEXA Top 1m list [11].

RHSBL

RHSBL or Right Hand Side Blacklist or also more commonly known as Domain-based blacklist DNSBLs, instead of listing IP addresses it lists domains that have a bad reputation. These lists use the second level and top-level domains(e.g. .com is part of a top-level domain and the google before it forms google.com) of a given email address or fully qualified domain name [12].

URIBL

URIRBL or Uniform Resource Identifier Blacklist does not use only IP addresses or domain names but instead also makes use of URLs to list malicious behavior. URI DNSBLs were designed because too much spam made it past the spam filters in the time frame between the use of the suspected IP address and the moment it was listed in an IP based DNSBLs. The spam that made it past the IP spam filter contained a lot of domain names and IP addresses in their links (referred to as URI), where these URI were detected multiple times before in spam. However,

the URI was not yet found in non spam email, making it undetectable. Therefore, extracting the URIs from the messages and checking them against the URI DNSBL preemptively detects malicious activity if it is not yet listed by spam filters.

2.3.2 Proactive Blacklists using Active DNS data

A domain name blacklist allows users of a network to filter out unwanted traffic (mostly malicious) based on their domain features. As mentioned earlier, the Domain Name System provides a translating service that changes human-readable addresses to an IP address, so anytime a particular name is resolved to a specific domain that is on the blacklist at the interface of a network, the traffic will be discarded. The benefit of using active DNS data, is that these blacklists can preemptively detect domains as suspicious before they are proved to be malicious due to active data giving a bigger view of the Domain Name System compared to passive DNS data.

Chapter 3

State of the art

This chapter discusses the state of the art approaches on the detection of malicious domains using DNS data. Furthermore, we also discuss the state of the art on-internet bad neighborhood concept. Although many of the detection methods use passive DNS data or combination of active and passive, these features are still relevant when using active DNS data off course, excluding user-level features due to the nature of our data set. After discussing state of the art, we take a brief look at the Internet bad neighborhood concept. This is done in Section 3.2 and helps analyze any form of malicious clustering on the IP level.

3.1 Malicious domain identification using active and passive DNS features

Detection of malicious domains in between the millions of domains that are being generated daily has been a great challenge. The majority of these techniques utilize the use of passive DNS data to detect these malicious domains. These malicious domains can be categorized under the different categories, namely spam, DGA, botnets, phishing all these techniques can be classified on how and which data is processed into signature-based, anomaly-based, and DNS-based. The method presented in this thesis will be focused on the use of active DNS-based data. A few studies have already analyzed the use of active DNS data, in order to detect malicious domains; however, most methods presented are based on the use of passive DNS data. In order to find features that are relevant for analyzing the difference between active DNS data on ALEXA and RBL, different studies based on the detection of malicious domains using DNS data are analyzed. Though many studies used passive DNS data, the features that are not based on user-level data are still relevant when analyzing active DNS data.

3.1.1 Passive DNS data analysis

The use of passive DNS data was introduced Weimer et al. [5] in 2005, where he also presented several cases that can benefit from using this data collection method. This form of DNS data collection could be used for the containment of malware as it provides insight into query patterns, IP addresses, and host-names. Furthermore, it provides access to domain names, which can help distinguish typo domains from the legitimate domains preventing phishing domains from operating. This paper also showed that the use of filtering of unwanted traffic solely based on IP address could cause more damage than benefit because often multiple hostnames are resolved to the same IP address.

Antonakakis et al. [13] introduces NOTOS, a reputation-based system that assigns the status of the domain based on their DNS characteristics. These characteristics are established on network-based, zone-based, and evidence-based feature extraction. Network features provide how a domain has its resources such as domain names and IP addresses allocated. The zone-based features provide the list of IP addresses associated with the domain name and the history of the domain name itself. The evidence-based features are built upon the number of times a domain name has been associated with to be known malicious domain name or IP address and the number of blacklisted IP addresses that resolve to the domain name. The features used by NOTOS to detect these suspicious domains consist of the number of distinct BGP prefixes related to the IP addresses associated to the domain in question, the number of distinct countries, the number of Ases, the number of domains connecting to the same IP address and the length of the domain name. In order to use the previously mentioned features, NOTOS uses a clustering technique that, during the training, learns how to distinguish and identify network behavior. This results in an accurate detection system that has a TPR of 96.3% and 0.38% of FPR.
Exposure is a system developed by bilge et al [14], which makes use of passive DNS data to detect domains involved with malicious activity. EXPOSURE makes use of 15 features grouped into four categories: time-based, DNS-answer based, TTL value-based, and domain name-based. Time-based features split the time into intervals and are used for time two types of analysis. The first analysis is global, which is used to see if a domain is short-lived, and on the other hand, it is used to measure the behavior of the domains over time. DNS answer based features are used to measure the heterogeneity in the IP addresses associated with a suspicious domain. The features extracted from this category the number of different countries, the number of distinct domains that share the same IP address. Each queried resource record has registered time to live as malicious domains tend to have lower

TTL values making them harder to take down a group of these TTL features that could prove quite use full. From these categories, the following features have been derived: standard deviation of TTL values, number of different TTL values, the number of changes in TTL, and the ranges these malicious domains operate by. Domain name- based features analyze the number of digits or characters a domain name contains. The motivation for this that benign service more often uses easy to remember names. Exposure uses these features to build a classifier that achieves a detection rate of 98% and an FPR of 1%.

Passerini et al. [15]. introduced a system called FLUXOR that is used to identify and monitor fast-flux service networks used for malicious intent. This system uses a set of features that observes the domain name, availability of the network, and heterogeneity of the agents. Using DNS data and these three categories of features, it successfully attempts to identify fast-flux networks. The domain name features measure the age of the domain name and how it behaves over time. The availability of the network is measured by analyzing the number of distinct A records and the time to live due to these fast-flux domains operate under low TTL. The last category uses features characterizing the heterogeneity of the potential agents of the network. The features extracted from this category measure the number of distinct networks, distinct autonomous systems, distinct resolved qualified domains names and distinct organization a domain is associated with. The data was actively collected by sending out simple queries. This resulted in them detecting more than 390000 compromised machines in a short time, which originated from the 387 detect fast-flux service networks.

Furthermore, Holz et al. [16] analyses DNS data by looking at their IP diversity, the number of unique A records returned in DNS lookup, and the number of Name server records querying them actively. Using these features, he develops a metric with which Fast-flux service networks can be detected. The number of features implemented is minimal and does not add any features that have not been previously used. However, most of these features are not dependant on the use of user-level DNS features, making them use full as active DNS features.

Hao et al. [17] this system monitors the initial behavior of malicious domains by observing the DNS infrastructure the domain is associated with using the resource records and DNS lookup patterns. The resource records analyzed by this system consist of NS, MX, and A record of a domain. Each of these records is then enriched by observing the AS, the name of the AS and country it is associated with.

Chiba et al. [18]. describes domain profiler, which can discover domains that might become malicious in the future by analyzing temporal variation patterns. This

research proposes a set of 55 features when attempting to identify these possible upcoming malicious domains. These 55 features have been grouped into features using TVP (Legitimate), TVP (Malicious), BGP features, ASN features, registration features, and lexical features. The TVP (legitimate) analyses the behavior of the different versions of ALEXA over time to see if a domain name has risen, fell, or stayed stable. This is because newly registered domains are more likely not to exist in the ALEXA 1M list. Moreover, the TVP (malicious) is also analyzed, which provides domains that have already occurred in public blacklist. The BGP features measure the number of BGP prefixes, countries, and IP addresses that are associated with FQDN, 3LD, and 2LD. Similarly is performed for the ASN, registration, and domain name features. These features are then used to build a classifier that preemptively detects malicious domains a maximum of 220 days in advance with an accuracy of 98%.

Ma et al. [19] detects malicious websites using malicious URLs without needing to visit them. This approach classifies these malicious domains using lexical and host-based features. These two categories together are made up out of 17 features. The host-based category takes into consideration IP addresses properties, WHOIS properties, DNS properties, and geographic location. When analyzing the IP address, it is checked whether it is blacklisted or not, and if the IP addresses associated with the A, MX and NS records are within the same ASN. Another feature that is analyzed is the TTL value of the resource records belonging to the hostname. Furthermore, the geographic feature did not only include country, city, and continent but also up-link connection used by the host. The authors achieved a false negative rate of 7.6% while only falsely detection 0.1% of the there test data.

3.1.2 Active DNS data analysis

Though the use of passively collected DNS data provides good results, it does not provide the same view of the DNS active DNS data does. The use of actively queried DNS data with sufficient access to zone files provides a more comprehensive view of the threat being analyzed. Aside from being more comprehensive, it also provides the possibility of proactive blacklisting or preemptive detection of suspicious domains. All detection methods mentioned that use DNS or host name-based could be used as possible features in an approach using active DNS data. In the detection methods mentioned above, a few categories of DNS features keep being repeated, such as Network-based, TTL value-based, and domain name based. All of the features mentioned are applicable as Active DNS features aside from the user level once in the data set available in this thesis.

3.2 Internet Bad Neighborhoods

Daily malicious traffic comes from all parts of the internet, but there is evidence that suggests these are concentrated in certain parts of the IP space. Ward et al. [20] first introduced this idea when they were looking for a new way to filter spam, which did not have to analyze the entire email. Moura et al. [1] defines these clusters of malicious activity *"Internet Bad Neighborhood is a set of IP addresses clustered according to an aggregation criterion in which many IP addresses perform a certain malicious activity over a specified period."*. These bad neighborhoods are acquired by aggregating malicious IP addresses into clusters. This aggregation can be done using network prefixes, (e.g. /24, /8, /18), in Classless Inter-Domain Routing (CIDR). Usually, the aggregation is done /24 subnets because this is proven to be the most stable. In this paper, Moura et al. [1] characterizes the behavior of internet bad neighborhoods by separating them in high volume and low volume spammers. Among their findings, they found that ten percent of the spammers are responsible for a large part of the spam being sent. The detection of these bad neighborhoods was done using DNS blacklists by counting the number of spammers identified in certain IP space, and a fixed subnet of /24 was used. According to the author, there are three possible reasons the bad neighborhoods are occurring: some internet service providers keep a blind eye to malicious activities occurring on their network. Another reason could be that the ISP's are more malware tolerant, making the spreading of malware easier. Finally, non-technical factors like the absence of internet crime legislation which gives the ISP less incentive to pay attention to malicious activity on their network. The approach in this thesis is to analyze if there is any form of DNS abuse. This abuse could also present itself in the form of malicious clustering on an IP level; hence the bad neighborhood concept could prove useful.

Chapter 4

Analysis of active DNS features on ALEXA and RBL

In this chapter, we analyze the difference between active DNS data features on ALEXA and RBL. Section 4.1 gives an overview of the statistical methods that have been applied in this chapter. Section 4.2 gives a brief overview of the data set used thought this thesis. Section 4.3 gives a summary of the features that have been selected. Section 4.4 is where we take our features and analyze them for any deviation between the CDF plots. These features can be separated into four categories, namely DNS record based, Network based, TTL value-based, and Domain name based. This chapter will conclude with section 4.5, where we present the conclusion on the analysis.

4.1 Approach

This section will elaborate on the methods used to measure if there is any difference between the DNS features in RBL and ALEXA. The first part, briefly elaborates on Cumulative Distribution Function plots that are used to analyze the difference between ALEXA and RBL features. In order to validate the significance of the deviation between malicious and benign domain features, the Kolmogorov Smirnov test is used. This test assures that the two CDF plots are not from the same distribution and therefore the deviation is considered valid.

4.1.1 CDF plot

A cumulative distribution function (CDF) plot displays the cumulative distribution function of the data. This plot displays an F(x), which is defined as the proportion of x values less or equal to x. This is effective for analyzing the distribution of sample data and allows the comparison of empirical distribution (e.g. malicious domains) to the theoretical distributions e.g. the behavior of benign domains. In Figure 4.1, an example is given where the red CDF represents the empirical distribution (malicious database) and the blue the theoretical (benign database). The analysis of the graph indicate that there is a horizontal deviation at the 50Th percentile, which is represented by the black line. The value of the deviation is represented by the x value and can be any numerical feature. The deviations indicate that the RBL behaves differently than the ALEXA for certain values of x. The more significant the deviation between the two CDF plots, the more likely the feature could prove useful. The same reasoning will be used when analyzing the behavior of DNS features on ALEXA and RBL.

4.1.2 Kolmogorov Smirnov test

The CDF plots gives a good indication of the relationship between two data sets, but in certain situations this is not visually convincing. In order to measure the statistical difference, a second test is utilized. Using the KolmogorovSmirnov statistical test, the distance between samples of two CDF plots can be quantified. This is a test for the null hypothesis that two independent samples are drawn from the same distribution. In this chapter, the samples are DNS data features from ALEXA and RBL. The result of the test consists of a K-S statistic (Figure 4.1) and the p-value if the K-S statistic is small or the p-value is high, then the hypothesis cannot be rejected and two samples are from the same distribution. The K-S statistic is displayed by the red line, which shows the most significant vertical deviation between the two CDF plots. In the Table 4.1 below, the p and d values have been calculated for each feature to validate that deviation is not due to them being from the same data set.

4.2 Data set

In this section, a brief overview is given on the data set used in this thesis. The active DNS data used in this thesis is actively collected and provided by the OpenINTEL platform. OpenINTEL is a High-Performance, scalable Infrastructure for Large-Scale Active DNS Measurement [2], which measures over 60% of the domain name space daily. This high-performance measurement infrastructure outputs DNS data from

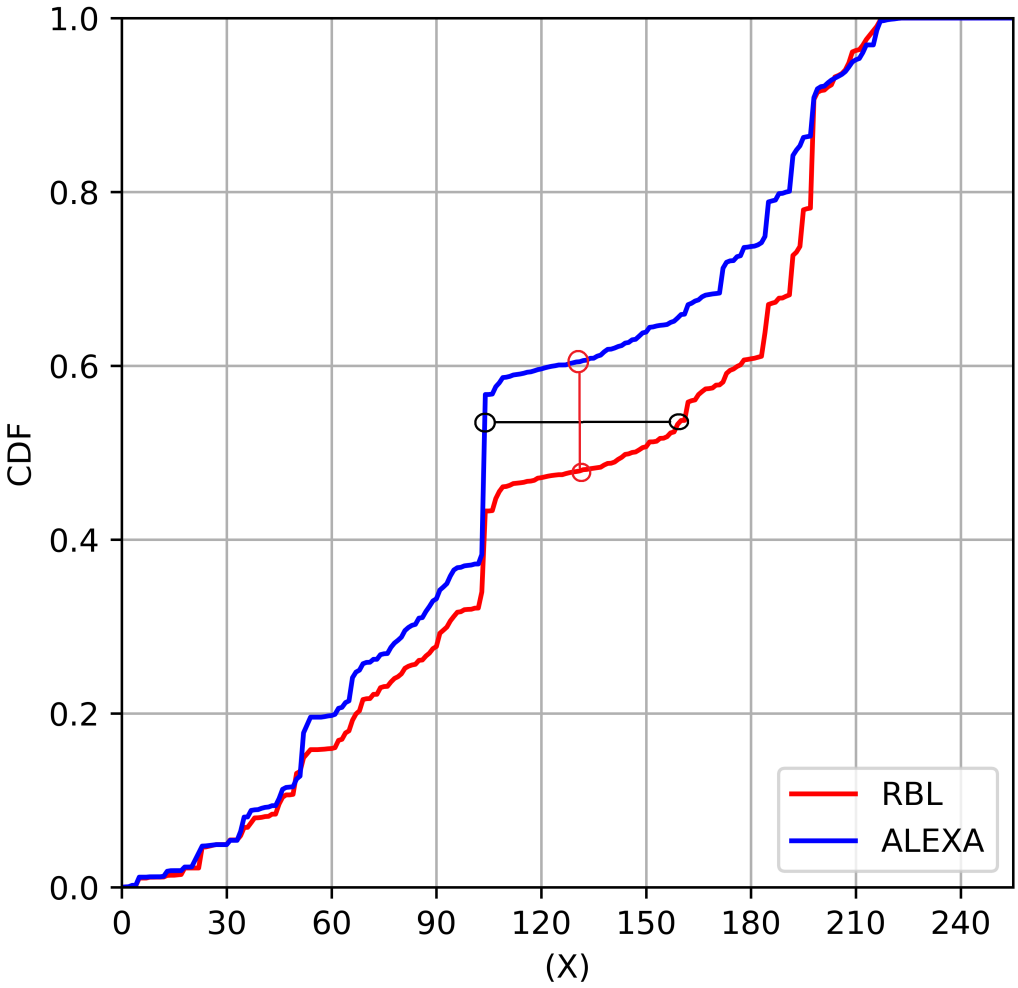


Figure 4.1: Cumulative Distribution Function

the records that can be seen in Table 2.1. Whenever the DNS queries are sent to the domains, the response is stored, and this repeats once a day due to the high number of domains to size of the list that has to be queried. In this thesis, two data sets are actively queried and used, namely ALEXA Top 1M and Real-time blacklist. The ALEXA Top 1M is a white list containing the top 1 million ranked domains in the world. This list is used as ground truth for comparison to malicious behavior of the RBL. The RBL consist of 22 publicly available blacklists, which are aggregated into second-level domains. The blacklists are aggregated together to create a more comprehensive and complete blacklist. These blacklists contain an average of 400000 domains per day. Furthermore, for the validation 64 antivirus databases are used, which are provided by virus total.

4.3 Feature selection

From the set of entries of each domain in our Active DNS database, 27 features are extracted. The resulting feature set are comprised of 4 groups, namely DNS record based (1-12), Network based (13-19), TTL based (20-24), and domain name based (25-27). In the following section, these features are analyzed based on the methods mentioned in Section 4.1.

Table 4.1: Analyzed features

Category	#	Features	D value	P value
DNS record based	1	A	0.99	0
	2	AAAA	0.37	0.28
	3	MX	0.25	0
	4	NS	1	0
	5	CAA	0.11	0
	6	CDS	0.03	0.68
	7	SOA	0.021	0
	8	TXT	0.9	0
	9	NSEC3PARAM	0.03	0
	10	NSEC3	0.3	0
	11	NSEC	0.022	0
	12	SOA	0.02	0
Network based	13	As_count per domain	0.003	0
	14	A_unique IP address	0.05	0
	15	Unique_MX_loud_addresses	0.01	0
	17	Number of VSPF	0.02	0
	18	Number of IPs in VSPF	0.05	0
	19	Number of IPs in vspf	0.06	0
TTL based	20	TTl of A	0.0680	0
	21	TTl of AAAA	0.18	0
	22	TTl of MX records	0.133	0
	23	TTl of NS records	0.09	0
	24	TTl of TXT records	0.083	0
Domain Name Based	25	Words in domain	0.03	0
	26	Length_domain_name_	0.07	0
	27	Query name length	0.1	0

4.4 Feature analysis

In this section, we will apply the statistical methods mentioned in Section 4.1 to features in Table 4.1. As a result of this concluding which features are useful for building a detection method based on the RBL data set.

4.4.1 DNS records

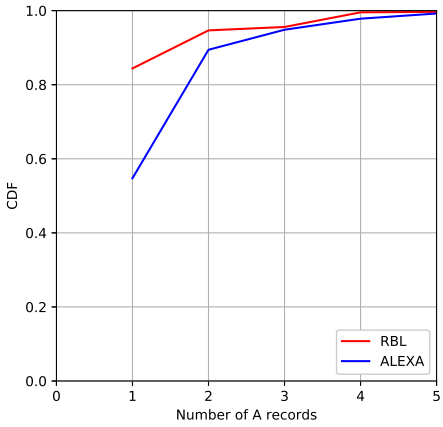
As previously mentioned, DNS records can be queried by hosts to provide information on a domain. The number of records could provide insight into the behavior of a domain. In order to do that, the number of records is counted for each domain. When an A record is queried, it provides the user with an IPv4-address. Higher levels of A records can be associated with the use of spam domains, as seen in [21]. In Figure 4.1a below, we compared the count of A records per domain between ALEXA and RBL. The analysis indicates that there is no significant deviation between A records on the ALEXA and RBL. This result shows that the counting of A records cannot be used as a feature to detect malicious domains in the RBL. This comparison has also been made for feature 2 - 12 with no significant deviation between any of the features listed (see Figure 4.1).

4.4.2 Network based

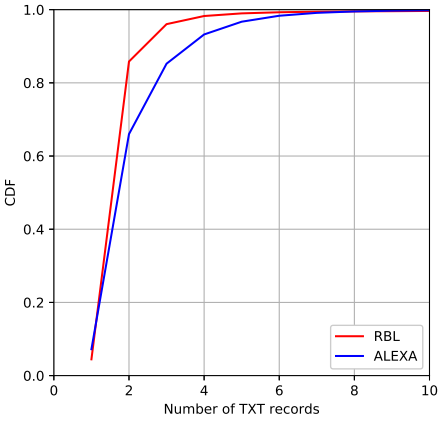
The DNS answer that is received by a server can consist of multiple A records mapping from a host to multiple IP addresses [16], [22]. These IP addresses may all lead to the same location, but is not considered the most effective technique due to load balancing. Typically malicious domains resolve using compromised computers that are located in different ASN, countries, IP ranges, and regions. For this reason, the network based feature might present some useful insight on the behavior of malicious domains.

Unique As count per domain

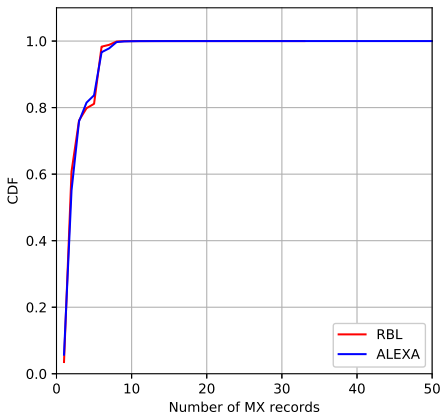
Malicious domains can be hosted by infected computers originating from different autonomous systems making them harder to trace the origin [14]. An autonomous system is a collection of routing prefixes maintained by one of our more entities. These entities set the routing rules, policies and the region of that network. In order to the analyze the difference between the RBL and ALEXA domains, we observed the number of (unique) ASNs per domain. Figure 4.2 shows the CDF plot of the number of (Unique) As counts for benign and malicious domains. The result shows that there is no significant deviation when comparing ALEXA and RBL. Similarly, the



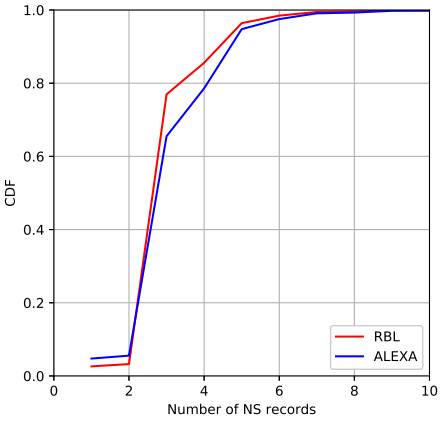
(a)



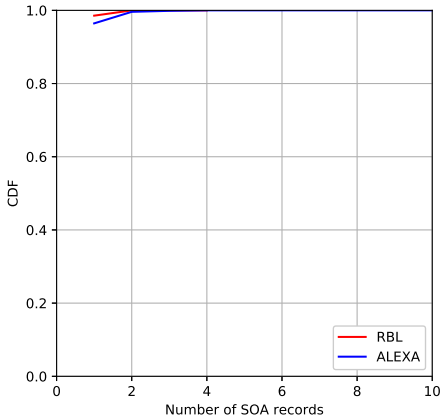
(b)



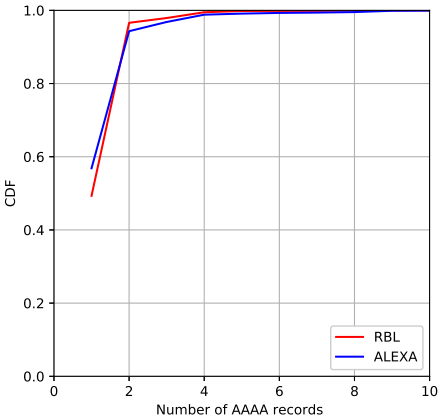
(c)



(d)



(e)



(f)

features [13,19] shows no significant deviation between the benign and malicious domains.

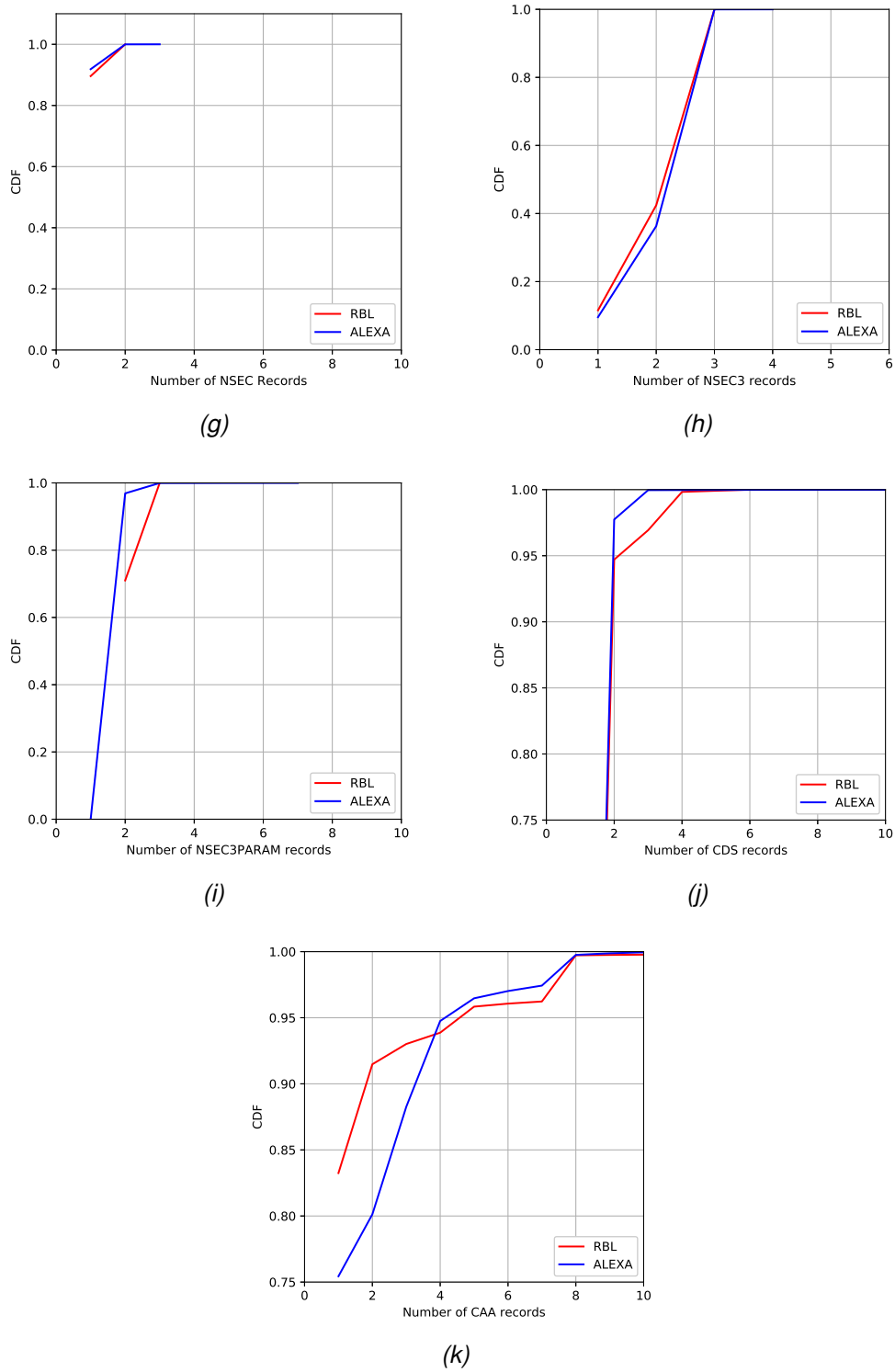


Figure 4.1: Cumulative distribution of DNS records for malicious and benign domains for the following records: *(a)* A , *(b)* TXT, *(c)* MX, *(d)* NS, *(e)* SOA, *(f)* AAAA, *(g)* NSEC, *(h)* NSEC3, *(i)* NSEC3PARAM, *(j)* CDS, (k) CAA

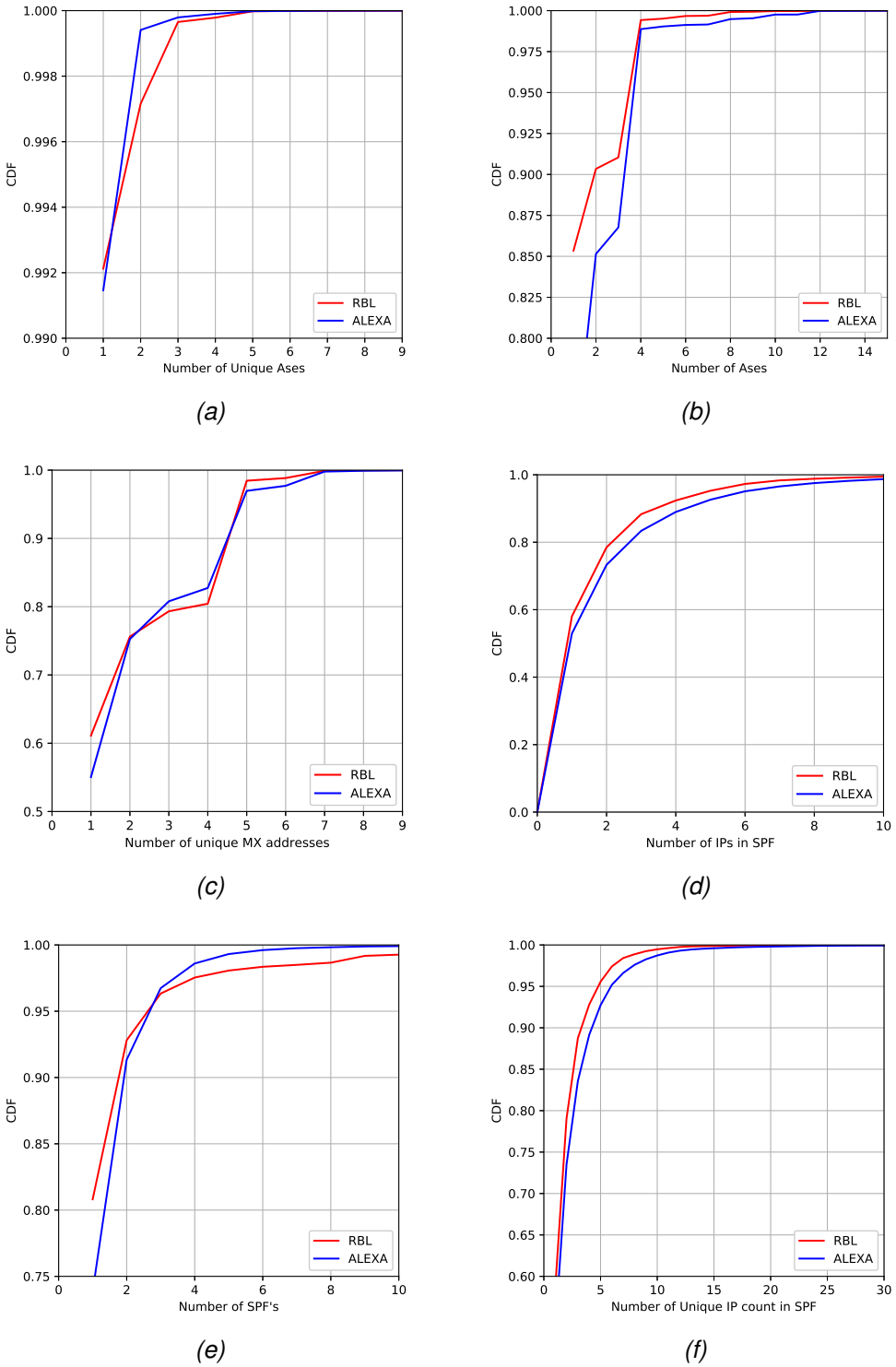


Figure 4.2: Cumulative distribution of DNS answered based features for malicious and benign domains analyzing the following features: *(a)* Unique Autonomous System count *(b)* Autonomous System count *(c)* Unique MX address count *(d)* Verification Sender Policy Framework IP count *(e)* Verification Sender Policy Framework count *(f)* Unique IP count in Verification Sender Policy Framework

4.4.3 TTL Based

Each DNS record is provided with a TTL(Time To Live) containing the time a record of a domain is allowed to stay in the cache. These TTL values benefit both the name server and DNS clients. For a system to have high availability, a low TTL value is preferred with the use of a round-robin. The problem with this approach is the constant switching of IP addresses due to short TTL. This configuration makes it easier for malicious systems to gain higher availability. An example of this Fast flux service networks, which generate many domains with short TTL making it harder for the blacklist to capture them. Analyzing this feature may provide insight on the behavior of malicious domains in comparison to benign.

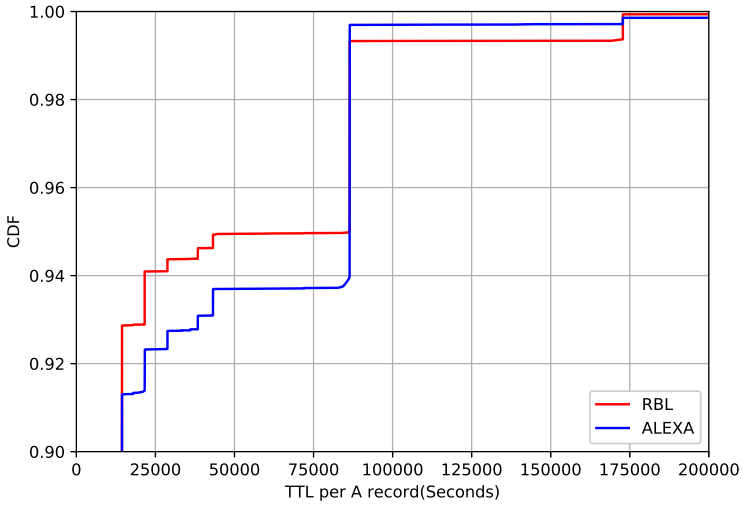
The TTL of the A and AAAA records

The TTL of the A and AAAA records provide us with the time the IPv4 and IPv6 addresses are valid for connecting to a server. However, there is no guarantee this address is valid, and the server returns whatever values are configured. If A or AAAA record have a short TTL, it can be associated with flux domains because they prefer to operate using low TTL values.These low value make it harder to track them due to the constant changing of the IP or domain they identify with. However, at the same time, services(e.g content distribution networks) prefer lower TTL because it ensures their users have the latest updates. Figure 4.3a shows a CDF of the A record time to live for malicious domains vs. benign domains. This analysis indicates that at the 94th percentile for the A records, TTL distribution of malicious domains has, on average 65000 seconds lower TTL value then benign domains. Similarly, in Figure 4.3b, at the 70th percentile of AAAA records, the TTL distribution of malicious domains have an average of 3000 seconds lower TTL value then the benign domains. These results show that the A and AAAA records TTL values of malicious domains deviate from the benign domains, meaning this could be used as a possible feature for the detection of malicious domains.

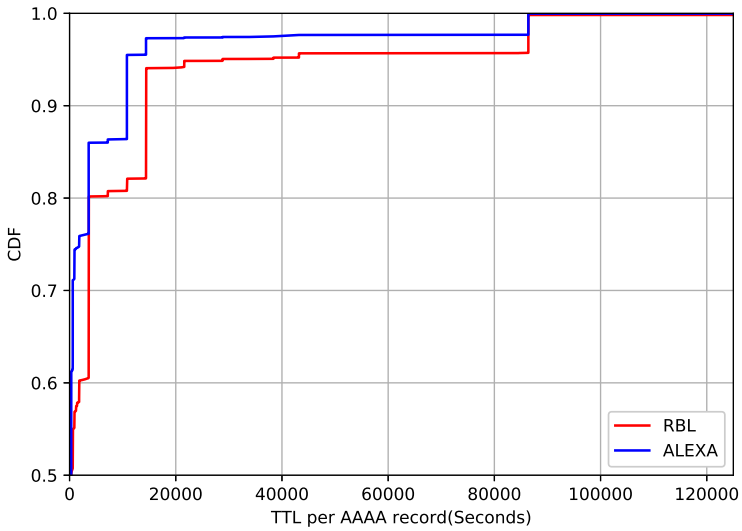
The TTL of the MX and TXT records

Usually, the TTL of MX, TXT, and NS records are kept high due to the settings not changing that frequently. Figure 4.4a shows a CDF of the MX, TXT, and NS record TTL for malicious domains vs benign domains. This analysis in Figure 4.7a indicates that at the 92th percentile for the MX records, the TTL distribution of malicious domains has, on average 67500 seconds lower TTL value then benign domains. Similarly, in Figure 4.4b at the 70th percentile of TXT records, TTL distribution malicious domains have an average of 10000 seconds higher TTL. Furthermore, in

Figure 4.4C, the TTL of the NS records indicates an average deviation of 100000 seconds at 93th percentile. These results show that the TTL values of TXT, MX, and NS records for malicious domains deviate from that of benign domains, indicating this could be a potential feature.

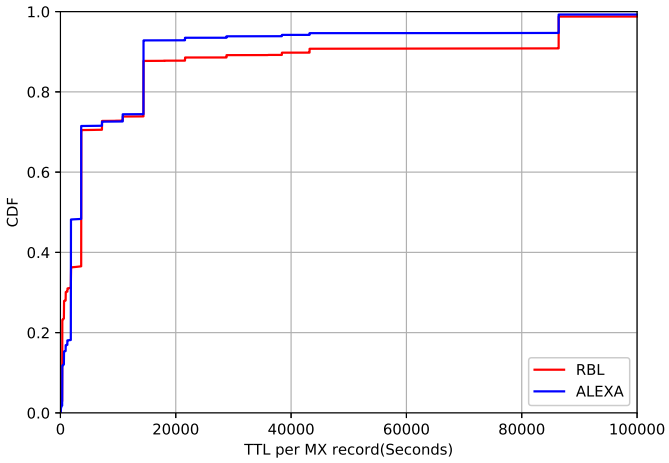


(a)

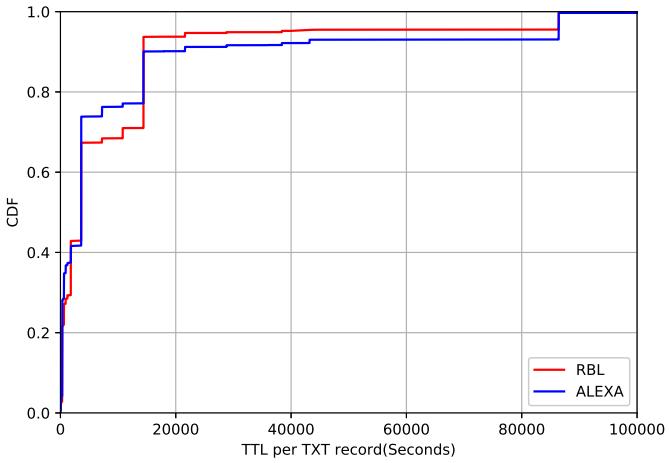


(b)

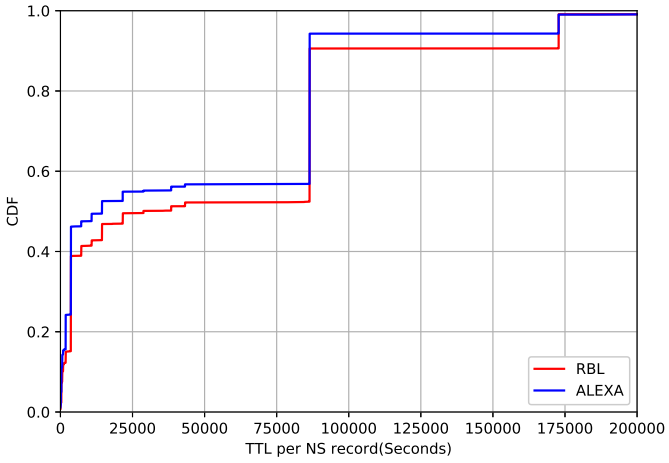
Figure 4.3: Cumulative distribution of TTL records for malicious and benign domains using: *(a)* TTL of A records (b) TTL of AAAA records



(a)



(b)



(c)

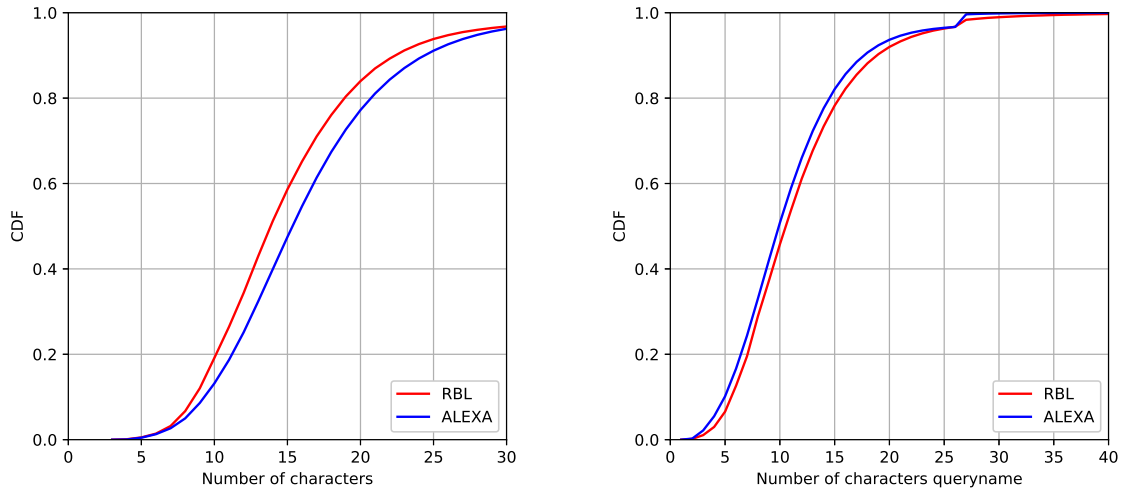
Figure 4.4: Cumulative distribution of TTL records for malicious and benign domains using: *(a)* TTL of MX records *(b)* TTL of TXT records *(c)* TTL of NS records

Domain name based

As mentioned before, in Section 2.1, the purpose of DNS is to translate human-readable addresses into IP addresses. When creating a domain name, benign services try to choose names that are easy to remember for their clients. The attackers capitalize on this by making very similar-looking domain names. This is utilized by setting up phishing campaigns, or by utilizing domain generating algorithms that produce short-lived domains. Hence including simple lexical features could give us some inside on the behavior of blacklisted domains in comparison to the benign domains.

Query and domain name length

The RBL consists of second-level domains. A simple way of making a distinction between malicious and benign domains is by comparing the length of their query name. Figure 4.5 shows the CDF of the length of the Second level + top-level domain and length of the second-level domain for malicious and benign domains. The analysis indicates that the Second level + top-level domain do not show a significant difference between them. When analyzing the SLD on its own, the results yield the same. This analysis showed no significant deviation between the length of malicious and benign domains.



(a) *(b)*

Figure 4.5: Cumulative distribution of Domain name based features for malicious and benign domains using :*(a)* SLD + TLD length *(b)* SLD length

Number of Words in a domain name

Another way of distinguishing malicious from benign domains is by the number of words in a domain name (e.g a domain with one word would be ”google” and a domain with two words would be ”name mesh” consisting of ”name” and ”mesh”). Figure 4.6 shows the number of English words that each malicious and benign domain contains. The analysis indicates no significant deviation between the number of words in a malicious domain in comparison to a benign domain. One reason for this might be that the filter used could not recognize a lot of the words due to the use of an English dictionary.

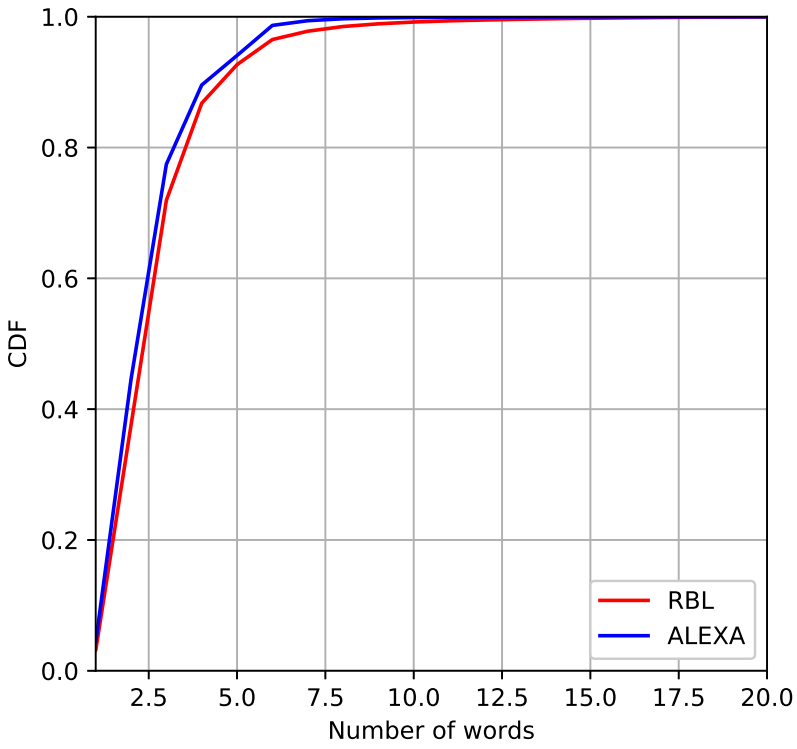


Figure 4.6: Cumulative distribution of Domain name based features for malicious and benign domains using: Words per domain name

4.5 Summary

In order to analyze the statistical difference between the active DNS data features on RBL and ALEXA, 27 features were extracted from the available data set. Analyzing these features [1-19,25-27] indicated no significant deviation for the used data set and features. It can be concluded from the results that although these blacklists contain criteria for detecting malicious domains, there is a strong indication that these values get averaged out when combining them, resulting in no deviation between the CDF plots. In order to make sure this is not a result of coming from the same distribution, the KS test was performed on the data sets, which can be seen in Table 4.1. In the case of TTL features [20-24], there was a deviation between ALEXA an RBL, which indicates that these could possibly be used as a feature for the detection of malicious domains. Though TTL based features indicated deviation between ALEXA and RBL, a useful signature cannot be build using 1 type of feature.

Chapter 5

Concept of bad neighborhoods in active DNS data

The DNS features analyzed in Chapter 4 did not result in any relevant features to make a useful signature. This resulted in the investigation of the clustering of malicious behavior at the IP level. In order to analyze the malicious behavior at IP level, the concept of internet bad neighborhoods was applied to our RBL data set. In this chapter, we will analyze the presence of bad-neighborhoods in the RBL DNS data and as a result of this answering RQ2. In order to detect these neighborhoods, different approaches were considered. The first approach will be to plot and visualize these bad neighborhoods along the Hilbert curve in Section 5.2. The second approach we observed the size of each /24 subnet by measuring the number of hosts, as seen in Section 5.3. The last approach is based on the adapted version of the bad neighborhood concept. In section 5.4, the thresholds for the model are set, and therefore the bad neighborhoods can be identified. This chapter will conclude with a summary of the results in section 5.5.

5.1 Internet bad neighborhoods

Internet bad neighborhoods are set of IP addresses clustered according to an aggregation criterion in which many IP addresses perform a particular malicious activity over a specified period of time [1]. Figure 5.1 shows how internet attack traces are passively logged and aggregated into a Bad Hood Blacklist. First internet traces are filtered and placed into a /32 blacklist, and then these are aggregated into /24 subnets blacklist. These /24 subnets are then considered the bad neighborhoods in which any incoming IP addresses that are associated with these neighborhoods have a higher likely hood of also being associated with malicious activity.

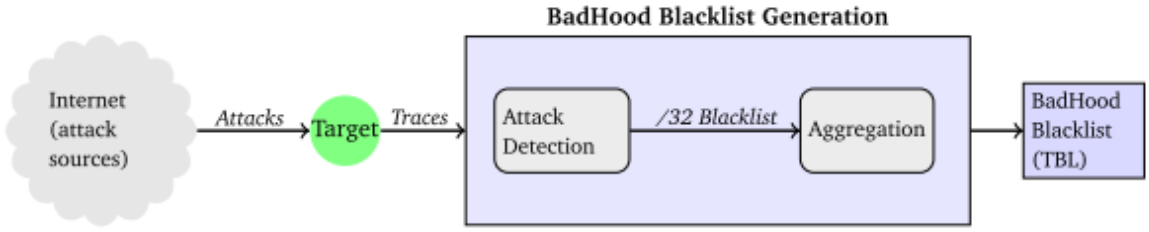


Figure 5.1: Approach to Find Internet Bad Neighborhoods [1]

5.2 Hilbert Curve

The IP space is vast and consists of precisely 4.2 billion IP addresses making it hard to visualize all those in a small space. The problem with displaying IP Addresses is that they are single-dimensional, meaning they only move up and down however, humans are not good at looking at a long list of single-dimensional points. One solution could be the use of the Hilbert curve, which is a fractal space-filling curve. This curve is useful due to its ability to map between 1D and 2D spaces meanwhile persevering locality. This makes it especially useful for plotting IP addresses by mapping 1D to 2D. Due to its properties, the IP addresses that are numerically close to another end up close to each other on the map. In the figure 5.2 below, we see the entire IP space sorted along a Hilbert curve. Each block represents a CIDR block maintained by registrar, and each pixel represents a /24. The different color gradient in each CIDR block represent the number of systems that are alive within the /24. The pixels that are not allocated are represented by a black pixel, which means there is no activity(seen in Figure 5.2). The black pixels that do not have activity could change over time due to the space being either reserved or not allocated by registrars to any hosters or users. This plot can be very useful in visualizing the bad

neighborhoods due to its nature of preserving locality. This could prove use full in identifying the clustering of IP addresses.

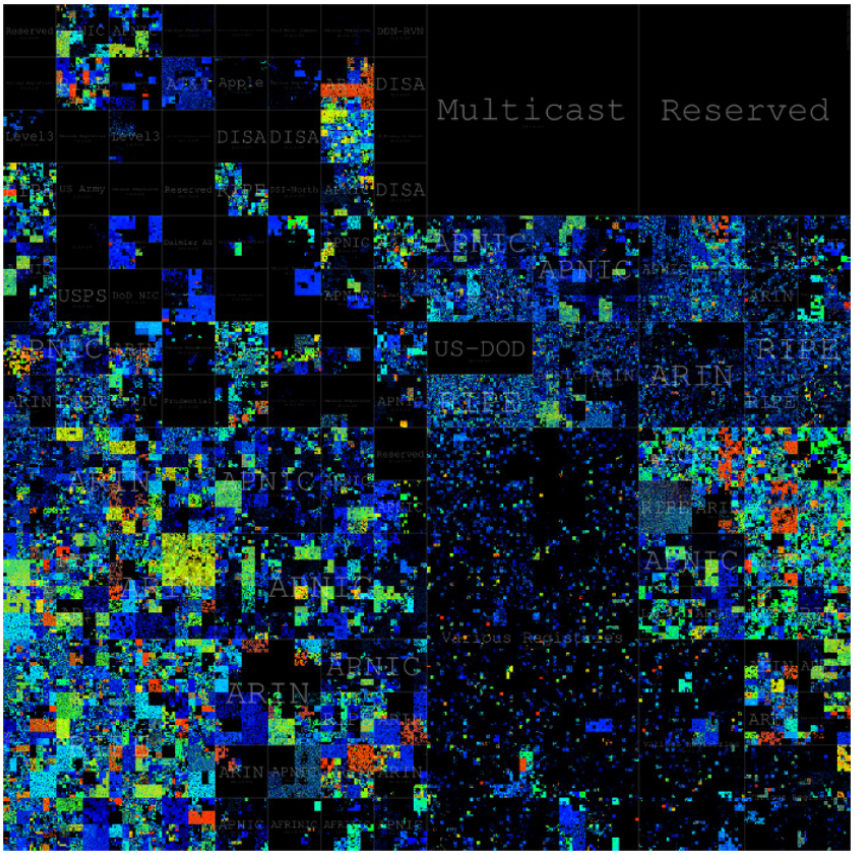
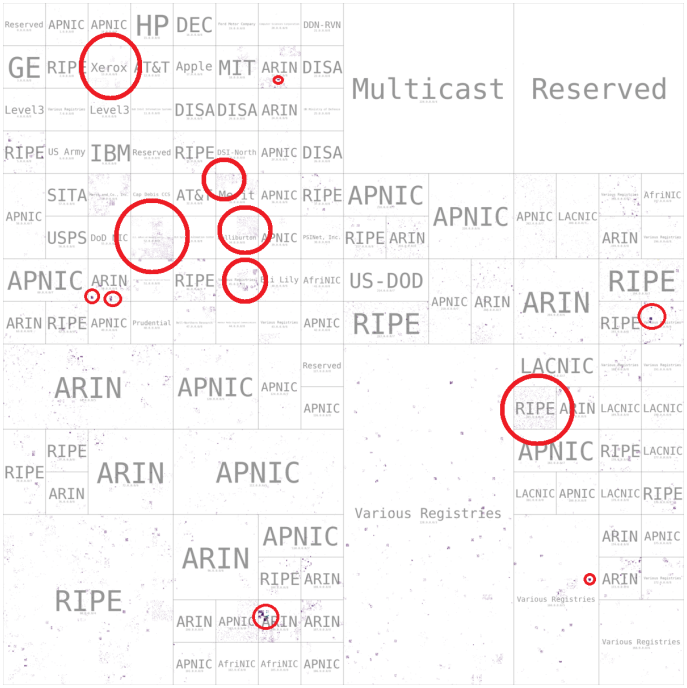


Figure 5.2: Hilbert curve of the entire IPv4 space

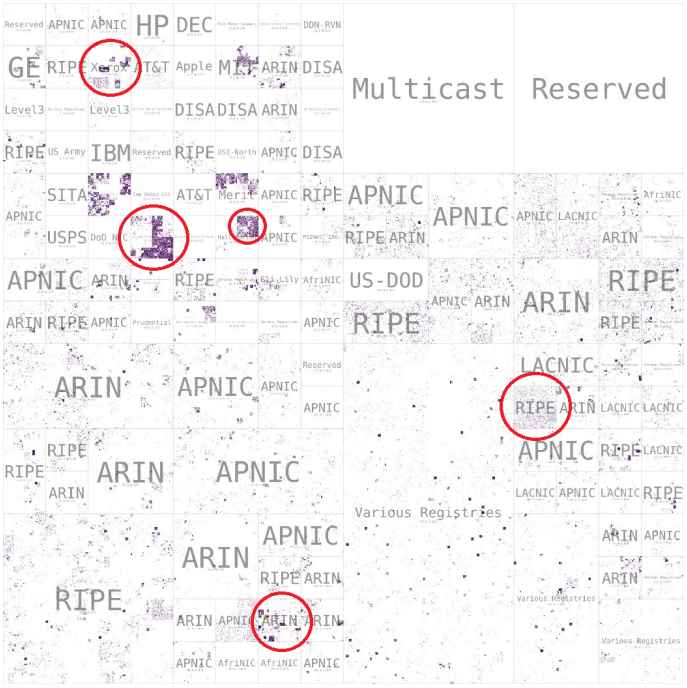
5.2.1 Hilbert curve ALEXA and RBL

In order to identify and visualize the clustering of both malicious and benign activity, the IPv4 space is plotted for ALEXA and RBL along the Hilbert curve. These plots can give an indication where different domains operated for both ground truths. The purple dots represent IP addresses within the IPv4 space, and the darker the purple gets, the higher the concentration of IP addresses. Figure 5.3 a, shows RBL plotted along the Hilbert curve, where the red circles highlight some of the clusters of IP addresses within the CIDR blocks. The analysis indicates clustering in almost all the blocks, with some showing more than others. These clusters can represent the presence of malicious infrastructure or hosters that neglect malicious behavior. Figure 5.3 b, shows the ALEXA plotted along the Hilbert curve, which shows the regions of the IP space the benign domains tend to reside. The highly dense regions are most likely occurring due to the presence of hosters. When comparing the two, it

seems that the malicious domains, for the most part, also reside amongst the benign space. Though this gives us some indication of which CIDR block the malicious and benign IP addresses reside, it does not quantify the maliciousness of the clusters, therefore not giving a proper assessment of the bad neighborhoods.



(a) RBL



(b) ALEXA

Figure 5.3: ALEXA and RBL plotted along Hilbert curve

5.3 Bad neighborhood concept in active DNS data

In Figure 5.3, the RBL IP space was visualized using the Hilbert curve this indicated grouping but fails to quantify the maliciousness of these clusters and therefore not able to do proper assessment of the bad neighborhoods. This section will elaborate on the scatter plot model and how it has been derived.

5.3.1 Utilization /24 Subnet

This approach identifies the presence of bad neighborhoods using the number of hosts within a /24 subnet size and is based on the traditional bad neighborhood concept [1]. The approach works by resolving blacklists and aggregating them into /24 subnets. After the aggregation, the level of maliciousness of the /24 subnet is measured based on the number of hosts within. Figure 5.4 shows an example of how the number of hosts in a neighborhood is counted. In the first column, there is

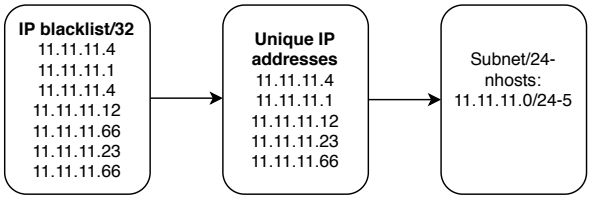


Figure 5.4: Bad neighborhoods based on number malicious host that resides in /24 subnet

a list of IP addresses that are resolved from domain blacklists. This column contains seven IP addresses, of which five of them are unique. In the column, in the middle, these five unique IP addresses are listed. These unique IP addresses all share the same three octets of the IP address (11.11.11.x), meaning they share the same prefix. Therefore these IP addresses can be aggregated into a single /24 subnet. This results in the 3rd column containing 11.11.11.x/24-5, which means this prefix contains five hosts. For this example, it would mean that this case, the /24 subnet 11.11.11. is utilized for 1.96%Ẇhen applying this to the RBL data set, it resulted in multiple /24 subnets containing a 100% utilization but when analyzing the number of domains hosted it led to a different conclusion. The results showed that even though a lot of hosts were malicious they hosted very little domains compared to other /24 subnets with slightly less hosts.

5.3.2 Host and domain count

Analyzing the utilization of /24 subnets indicates the malicious activity within a /24 subnet, but did not take into consideration the number of domains that were hosted. The main idea behind this approach is also to consider the number of domains hosted by the /24 subnets giving a better assessment of maliciousness. Figure 5.5 shows an example of how this aggregation approach works. The first column consists of a /32 blacklist, which is resolved from the domain name blacklist. This list contains 7 IP addresses that all share a common prefix (11.11.11.x), which is then split up into two lists. The first list counts the number of unique IP addresses, and the second list the number of IP addresses under that unique IP. This results in having five unique IP addresses and seven overall IP addresses. These are then aggregated together and result in the last column, which shows the prefix 11.11.11.x that contains five hosts that host seven domains. The approach then considers all domains associated with /24 subnets outside the blacklist as suspicious.

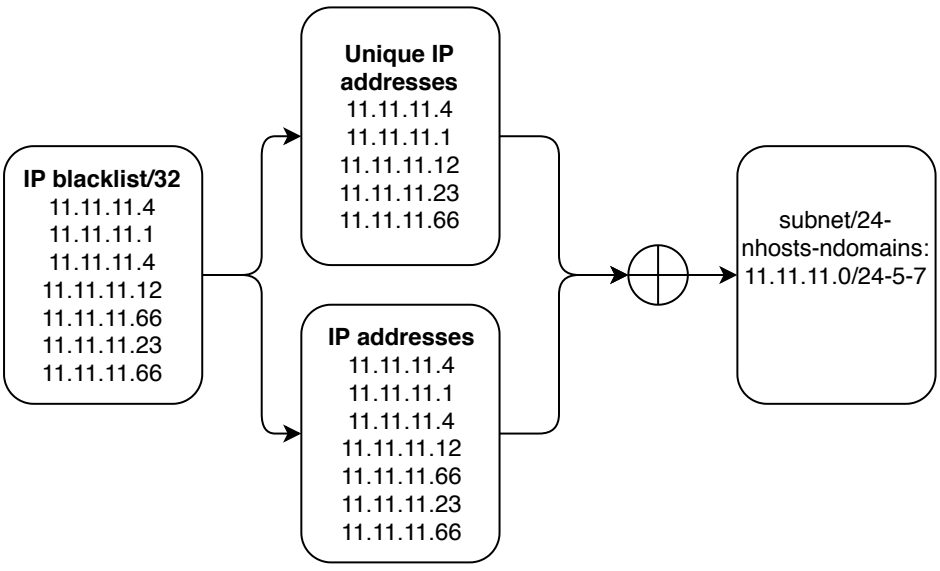
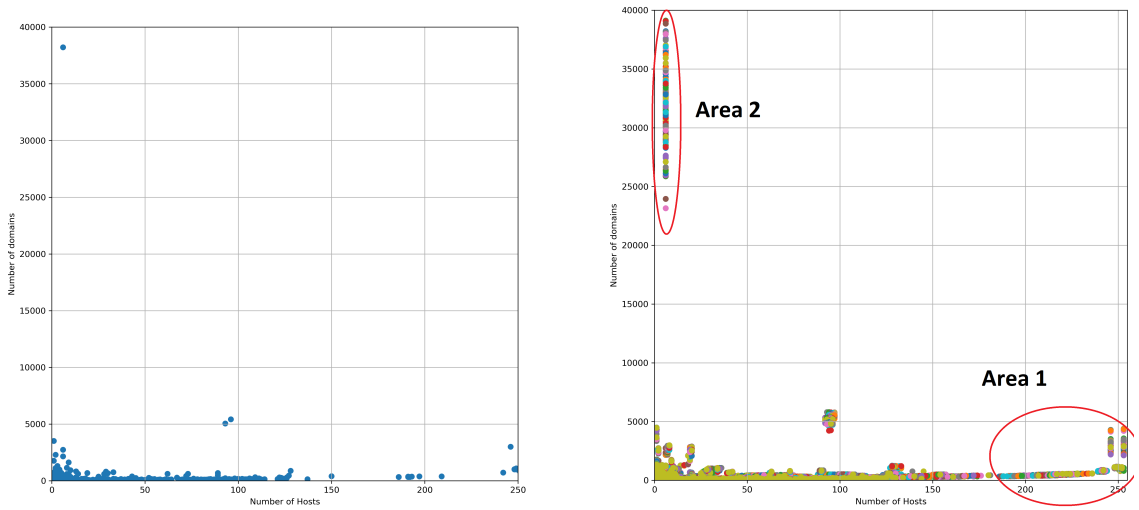


Figure 5.5: Detection of malicious /24 subnets using Host and domain count

5.4 Bad neighborhoods

In this section, the aggregated /24 subnets are placed in a scatter plot. In order to verify the location of these /24 subnets in relation to each other. Figure 5.6a plotted all /24 subnets generated for one day on the RBL. From this figure can be seen that there are outliers at the end of each axis. These outliers either have a very high domain or host count. These are the /24 subnets that are of interest due to them behaving outside the norm. In order to see if this behavior reoccurs, six months of data is plotted. In Figure 5.6b, this has been repeated for 180 days, which is represented by the colored dots. The different colors indicate the different days and /24 subnets. This period is assumed to indicate the behavior of the malicious subnets within the RBL. This results in two areas of interest from which thresholds can be derived. These thresholds are elaborated on in the following section.



(a) *(b)*

Figure 5.6: Bad neighborhood areas scatter plots: *(a)* /24 subnets RBL *(b)* /24 subnets RBL over a period of 6 months

5.4.1 Threshold 1

Area 1 in Figure 5.6b is identified by /24 subnets with an abnormally high number of hosts compared to a low number of domains being hosted. Figure 5.7 shows a scatter plot of area 1, the light blue dots represent the ALEXA list which has been aggregated into /24 subnets, and red highlighted dots represent /24 subnets on the RBL. When analyzing the clustered dots in the red highlighted area, there is a clear indication of clustering over six months period. The first cluster contains /24 subnets varying between [240-255] hosts and having a domain count between [790-1200]. The second cluster contains /24 subnets with a host count between [180-240] and domain count ranging from [200-600] domains. The third cluster has a host count between [235-245], and domain count between [600-800]. This results in set thresholds within the Geometrical method. These thresholds are set to filter out /24 subnets containing hosts count between [180-255] hosts.

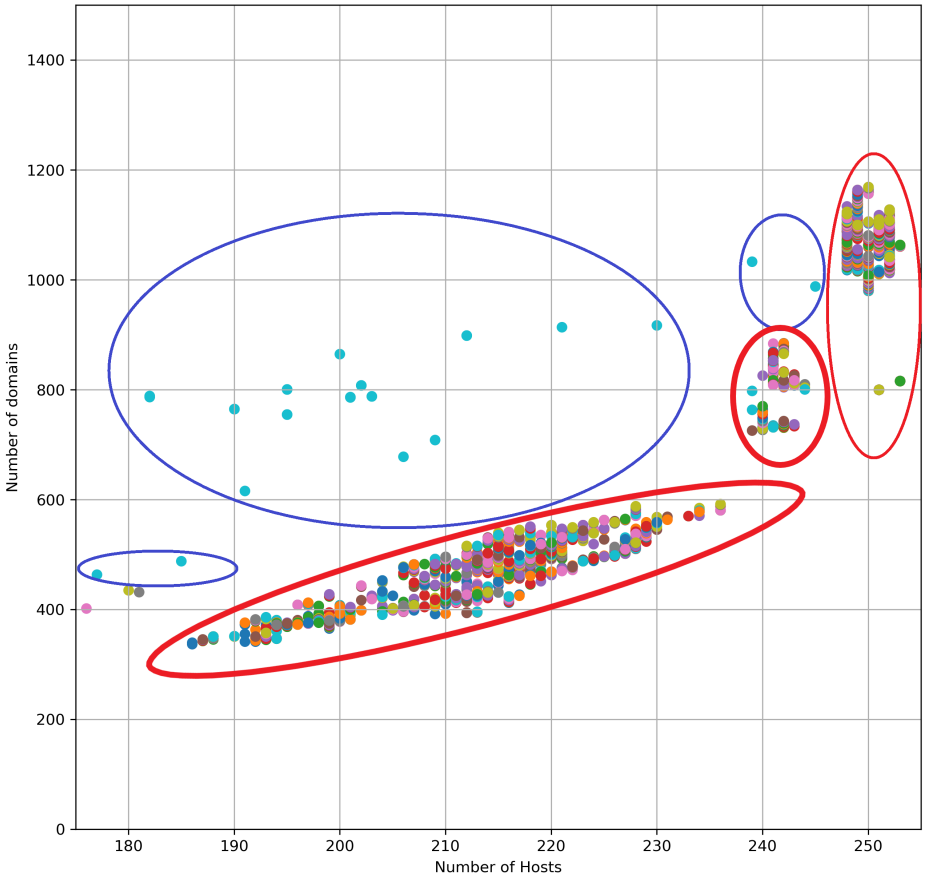
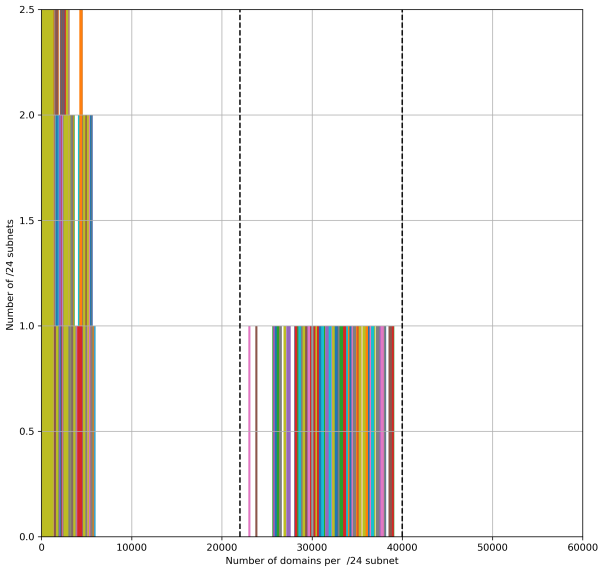


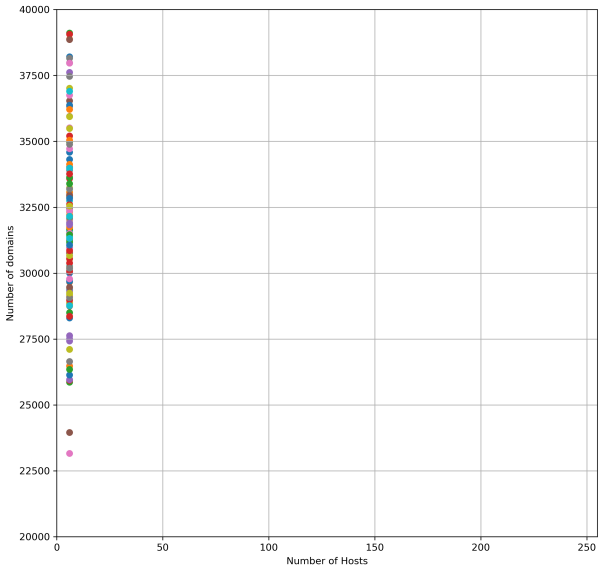
Figure 5.7: Thresholds set based on area 1

5.4.2 Threshold 2

Area 2 can be identified by /24 subnets with a high number of domains hosted by a low number of hosts. Figure 5.8a shows a histogram of the domain count for malicious domains over a six-month time span in area 2. This analysis indicates that there is a clustering of these /24 subnets between a domain count of [22000-40000] for six months of RBL data. Similarly, Fig 5.8b shows a scatter plot of area 2 over six months time period. It can be observed that these /24 subnets contain between [4-10] hosts. From the results of Figure 5.8, the derived thresholds are placed between [0-255] host counts, and [22000-40000] domain counts.



(a) RBL



(b) RBL6months

Figure 5.8: Thresholds set based on area 2

5.5 Summary

In this chapter, several approaches are analyzed to see if there any presence of bad neighborhoods in the RBL data. The result of analyzing these approaches answered research question 2. The first approach used the Hilbert curve to visualize the presence of bad neighborhoods by plotting the RBL and ALEXA along the Hilbert curve, but this did not provide a full assessment of the bad neighborhoods. Therefore in the second approach, this clustering was analyzed by measuring the occupation of /24 subnets. This is the traditional bad neighborhood concept which measures the number of malicious hosts that occupy an /24 subnet. However, the downside of this approach is that it failed to recognize how malicious a /24 subnet is due to its neglect of the number of hosted domains. The results indicated that the highest occupied 24/ subnets did not necessarily host the highest number of domains. This led us to our final approach, which did take into consideration the number of hosted domains. In order to see any form of clustering, these /24 subnets are placed into a scatter plot where the x-axis the number of hosts within a /24 subnet and the y-axis the number of domains belonging to those hosts. Using this scatter plot, we identified two areas that formed the region with bad neighborhoods, namely /24 subnets with a low number of hosts and high domain count and an area with low domain count and high host count. From these regions we have a derived 2 thresholds from which these bad neighborhoods can originate the first one is for /24 subnets residing between a host count of [180-255]. The second threshold are for /24 subnets containing a domain count between [20000-40000].

.

Chapter 6

Adapted bad neighborhood concept

In this chapter, we explain the approach of our method. Section 6.1 explains the main idea behind our method which is an adapted version of the bad neighborhood concept. Further on in Section 6.2 an overview is given of the different phases that are connected to the method in order to collect data, classify and validate suspicious domains. In Section 6.3, a brief explanation is given on the different components used with the model. Section 6.4 provides an overview of the first phase of the model, which explains the RBL database. Section 6.5 explains Phase 2, where the areas of the scatter plot are explained, and the model used to obtain them. Section 6.6 explains the components needed for the validation of the suspicious domains. The chapter concludes with the results of the validation period performed.

6.1 Approach

The adapted bad neighborhood or Geometrical method is designed to filter out the /24 subnets that appear within the highlighted region in Figure 6.1. The highlighted area is based on thresholds that have been derived in Chapter 5. The first area has /24 subnets containing a high number of domains and a relatively low number of hosts, and the second area contains a high number of hosts with a relatively low number of domains. These two areas are considered of interest within this Geometrical method due to the abnormal behavior they display compared to regular behavior of the malicious subnets. The /24 subnets within these areas are considered bad neighborhoods. These areas consist of two types of /24 subnets, entirely malicious and relatively malicious. The fully malicious /24 subnets have no benign domains when matched with ALEXA /24 subnet. The relatively malicious /24 subnets have a relative low amount of hosted domains. There are two main differences between bad neighborhoods concept presented in Section 5.1 and the adapted concept. The first difference is that the adapted concept uses domain name blacklist which are

aggregated to /24 subnets. These are then used to find unlisted domains that are associated to these /24 subnets. Secondly, the adapted model uses the number of hosts and domain count within /24 subnet instead of only measuring the number hosts that occupy the subnet.

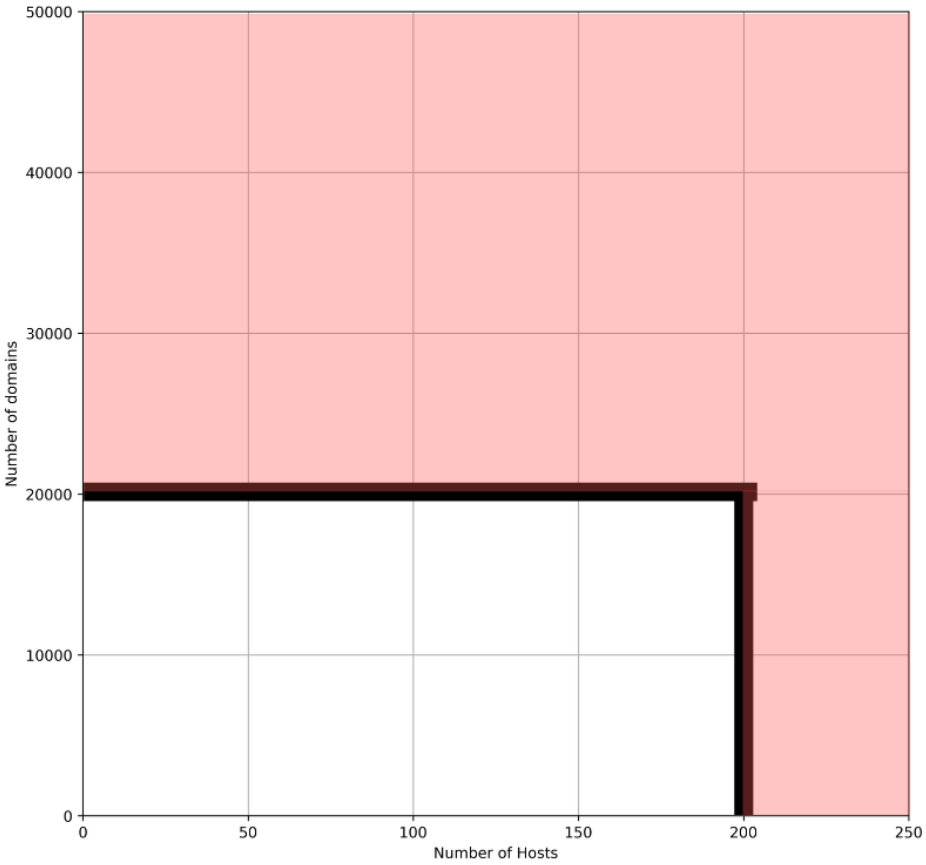


Figure 6.1: Geometrical method Area

6.2 Overview

In this part, a brief overview is given on the collection, training, and validation involved into making the Geometrical method work. This can be divided up into three phases: data collection and prepossessing, training the model, classification, and validation of the model, as seen in Figure 6.2. In the data collection phase, relevant data is extracted for the building of the model and pre-processed accordingly. In phase 2, the model is trained on the extracted RBL data, which provides the areas in Figure 6.1. In the third phase, the model classifies the suspicious domains, and these are then validated using ground truth.

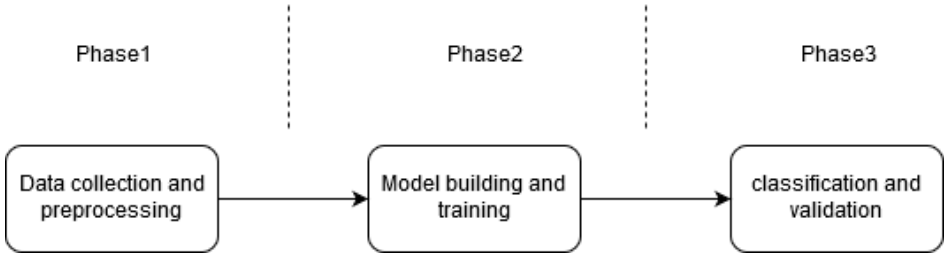


Figure 6.2: Overview of collection, training and validation phase

Figure 6.2 gives elaboration on each phase of the model, and it works as the following:

Phase 1 In the data-collection phase, data is extracted from the domain-blacklists on OpenINTEL. This is done by resolving the domain blacklists to IP addresses and aggregating them into /24 subnets which make up the potential bad neighborhoods.

Phase 2 In the second phase, the thresholds for the Geometrical method is set. The set thresholds identify the areas within a scatter plot where there the bad neighborhoods reside.

Phase 3 In the last phase, the classified domains originating from the resolved watch list are validated. This starts with the matching of the /24 subnet watch list in OpenINTEL, resulting in a list of domains that the model classifies as suspicious. These classified domains are then validated to measure the performance of the model.

Each phase consists of a set of components, as seen in Table 6.1 below.

In Figure 6.3 below, there is a block diagram with all the components belonging to each phase and how they are related. This is elaborated in the following sections.

Table 6.1: Overview of components used in different phases

ID	NAME	PHASE
A	DNS DATA RBL	1
B	Geometrical method	2
C	Watch list /24 subnets	2
D	ALEXA list/24 subnets	2
E	Suspicious domains	3
F	Validation	3

6.3 Components

This Section elaborates on the phases and components mentioned in Table 6.1.

A. **RBL data** This database contains a set of blacklists that are being actively queried by OpenINTEL once per day. The RBL data provides IP addresses that are resolved from the domain name blacklists.

B. **Geometrical method** The model domain count and host count to identity unlisted malicious domains. The thresholds used are based on areas that are defined in the scatter plot that was discussed in Section 5.3. The thresholds are derived using six months of blacklist data. This component is discussed in more detail in Section 6.5

C. **Watch list /24 subnets** This watch list contains the list /24 subnet list, which has been filtered based on the thresholds that are mentioned in section 5.3. This component is discussed in more detail in Section 6.4.

D. **ALEXA list/24 subnets** This watch list contains the aggregated /24 subnets list of the ALEXA Top 1M. This gives the assurance that benign activity within the malicious /24 subnets is as low as possible.

E. **Suspicious domains** This list contains the suspicious domains that were a result of matching the watch list with OpenINTEL.

F. **Validation** The classified domains are validated using the virus total. Virus total is an anti-virus website that provides 64 different anti-virus distributions.

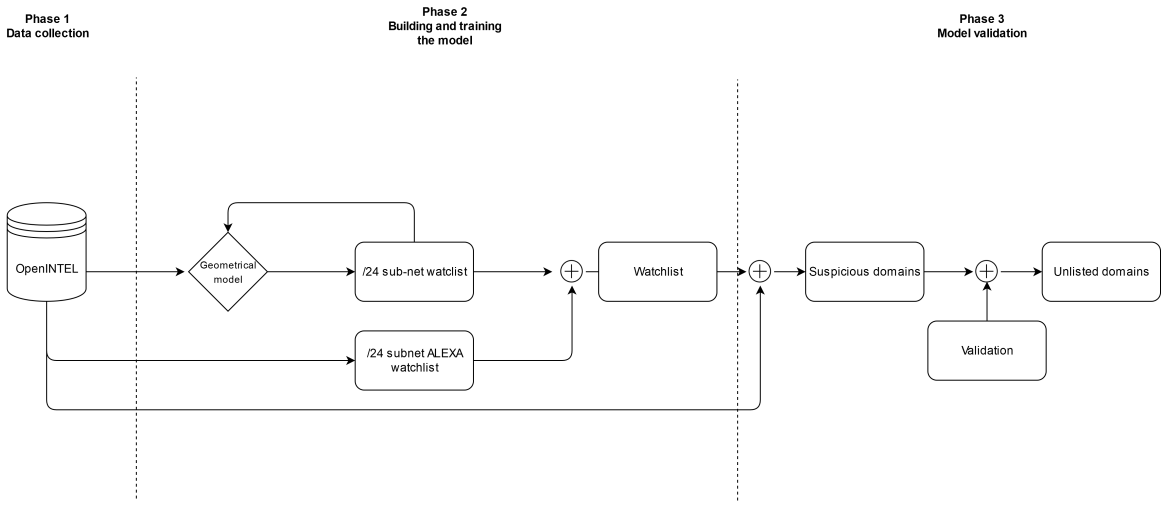


Figure 6.3: Overview of the relationship between all components in each phase

6.4 Phase 1

In this Section, the components relevant to data collection is explained. Within this part, the used sources and how they were processed are elaborated. In phase 1, the RBL database is discussed.

6.4.1 RBL data

The first and most crucial component of this section is the RBL database. This component consists of 22 publicly available blacklists, which are actively queried daily, as mentioned in section 2.1.2. OpenINTEL uses zone files as the basis of the measurement therefore, measurements beyond the second level cannot be performed. To unify the RBL measurements, domains on the list are truncated to second-level domains. The IP addresses of these second-level domains are then truncated under the assumption that it is /24 subnet. This is done by turning the last octet into a 0, as explained in section 5.3.

6.5 Phase 2

This section focuses on the main components of the system that are responsible for training the model. Which is made up out of the following components (B) Geometrical method, (C) /24 subnet ALEXA watch list, and (D) /24 subnet watch list.

6.5.1 Watch list /24

The watch list is created using the thresholds set within the Geometrical method. The Geometrical method filters out /24 subnets that fall within these thresholds. This results in a watch list containing the bad neighborhoods within the RBL data set.

6.5.2 ALEXA list/24

The ALEXA top 1M watch list presents the top 1 million benign domains. This domain watch list has been resolved into IP addresses and aggregated into /24 subnets. The result of this is an ALEXA /24 subnet watch list, which shows the malicious /24 subnets with a high level of benign domain counts.

6.5.3 Geometrical method

The Geometrical method identifies /24 subnets that appear within the set of thresholds. The thresholds are set within the scatter plot, which can be categorized in Area 1 and Area 2, which form the model. Area 1, the first threshold is set between 170 $\leq$ hosts $\leq$ 255 and domain count $>$ 0. Area 2 thresholds are set between 0 $\leq$ host count $\leq$ 255 and domain count between 20000 to 40000. Furthermore, the areas can be categorized into two types of /24 subnets. The first one filters subnets with relatively low benign domain count. The second type contains /24 subnets with no benign domain counts. These subnet types can be see in Table 6.2.

Prefix	Benign domains	Benign hosts	Malicious domains	Malicious hosts
Relatively no activity	22	10	1061	253
No benign activity	0	0	804	242

Table 6.2: 2 Types of /24 subnets

6.6 Phase 3

This section elaborates on the validation of the suspicious domains classified by the Geometrical method. First, the matching of domains from list/24 is discussed, and then how these domains are validated. The following components are relevant to the validation stage: (E)suspicious domains, (F)validation, and detected domains.

6.6.1 Suspicious domains

The suspicious domains are a result of matching /24 subnets watch list with Open-INTEL. This results in a list of domains that are classified by the Geometrical method as suspicious.

6.6.2 Validation

Validation of the suspicious domains is performed using two available databases. First, the suspicious domains are intersected with the currently available RBL. This results in a list of suspicious domains that are not yet listed on the RBL. These domains are then validated using 64 different antivirus distributions provided by virus total platform. The domains that get validated as malicious can then be listed.

6.7 Results

This section presents the validation period performed on the daily classified domains and their results. The validation period of the domains classified by the adapted bad neighborhood concept give insight into how well this model can function as a stand-alone detection method. The model will attempt to detect domains in real-time using publicly available blacklist. The validation is performed on domains originating out of /24 subnets residing within the thresholds on the Geometrical method. The validation has been designed to answer the following questions:

- How does the Geometrical method perform in the detection of malicious domains?
- How do the relatively malicious /24 subnet perform in comparison with a fully malicious /24 subnet?

The validation is first performed for the model, and then the individual results of the /24 subnets are analyzed. The /24 domains classified by /24 subnets in area 2 are not validated due to the limitation of the validations data set. In the following sections, the preparation of the validation is discussed. Further the setup of the measurements performed and the metrics used are discussed. The results of these questions answer Research question 3 and conclude the thesis.

6.7.1 The setup

The validation is performed on the domains originating from /24 subnets, which are identified as malicious by the adapted bad neighborhood concept. The first threshold contains a high number of hosts and a relatively low number of domains, and the second threshold contains a high number of domains hosted by a small number of unique IP addresses. These two thresholds form the model, which contains no benign domain count and subnets with relatively low benign domain count. After each test, the following metrics are calculated to determine the performance of the model:

True Positive(TP): The number of domains that classified as malicious and are malicious.

False Positive(FP): The number of domains that classified as malicious and are benign.

In this measurement, the False Negatives and True Negatives are not measured. True Negatives are benign domains that are not considered as malicious by our method. When analyzing True Negatives, which occur when a signature has been designed to detect a malicious domain, which are positive but can also recognize

benign domains as True Positive when detected. This cannot be calculated due to lack of data on benign domains because assuming domains that do not appear on RBL are benign is too uncertain, so we do not count True Negatives. False Negatives are malicious domains that are not detected by our method as malicious. The detection method is trained based on all the RBL data within the thresholds set. All the domains originating from these areas should be flagged as malicious by the model, but due to the filtering that we applied, the resulting number is not representative of the real False negatives, and therefore they were not measured.

6.7.2 Validation

In this section, the validation of domains classified by Geometrical method are analyzed. Figure 6.4 shows the malicious domains residing in the /24 subnets that are not listed in the RBL. The blue line represents the domains detected by our Geometrical method, which are not listed on the RBL. The number of domains daily classified is rather constant for the first part due too no new domains being detected. Another observation is the slight drop in the number of domains being matched, which could be a result of domains simply being taken down. Similarly, there is a rapid drop due to the disappearance of /24 subnets. This could be a result of /24 subnets been taken down or being part of fast flux networks. In this case, there is a spike in the number of domains detected could be a malicious hosting prefix which is taken down.

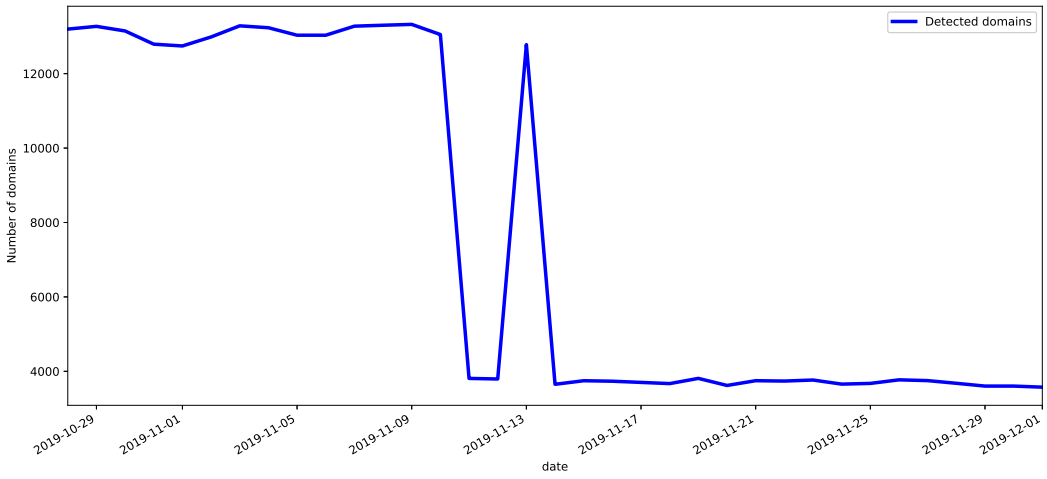


Figure 6.4: Daily detection Area 1

Figure 6.5 shows True Positive and False Positive for daily matched domains by the Geometrical method. The red line represents the daily True Positive rate, and the blue line the False Positive rate. The analysis indicates that the model has an average True Positive of 3% and average False Positive of 97%. The small bump represents the leaving of individual /24 subnets from the outside the set thresholds, which could not be validated, resulting in more True Positives and less False Positives. The result of the validation on the area shows that the domains originating from the model cannot be assumed to be malicious despite the fact that a high number of host resides in /24 subnets with a relatively low number of benign domains being hosted. In the next Section, the /24 subnets types are analyzed separately to measure their performance.

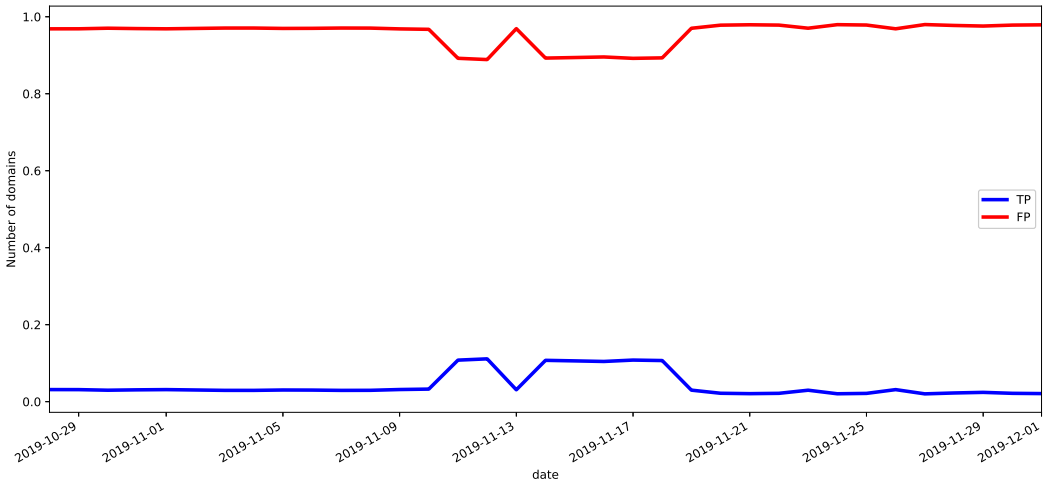


Figure 6.5: Daily True and False Positives for Area 1

Full malicious /24 subnets

This section will answer the question: how do entirely malicious /24 subnet perform in the RBL data set within thresholds of the model. These /24 subnets contain no sign of benign activity when intersected with ALEXA /24 subnets. Figure 6.6 shows the daily detection of subnets with no benign domains hosted. The blue line represents the number of domains that are detected by the method in fully malicious /24 subnets, which have not yet appeared on the RBL. The analysis of this Figure indicates that there initially no /24 Then we can observe the presence of fully malicious /24 subnets, which are slowly declining until they completely disappear and reappear again. This could be a result of these domains being taken down or no longer hosted under these /24 subnets or it could be part of BGP hijacking in which a prefix could be taken over for certain period. Due to the nature of /24 subnets, we

would assume that the domains found that are not yet listed in the RBL would also be malicious due to these /24 subnets not hosting any benign domains.

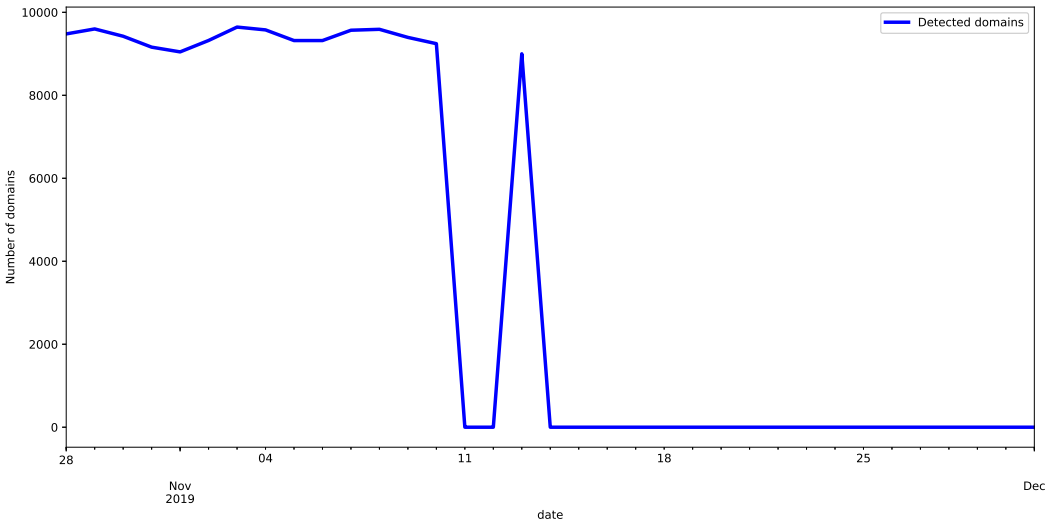


Figure 6.6: Daily detection of subnets/24 without benign activity in Area 1

Figure 6.7 shows the daily true and False Positive rate of fully malicious /24 subnets with no benign domain count. The red line represents the daily True Positives and the blue line the False Positives. The analysis of this Figure indicates there is no change in the daily True Positives that have a rate of 3.5%. Similarly, there is no change in the False Positives of 96.5%. The results show that even though the /24 subnets are occupied by only malicious domains, the number of validated domains is very low. This could also be due to not having sufficient data to validate these suspicious domains classified by our method.

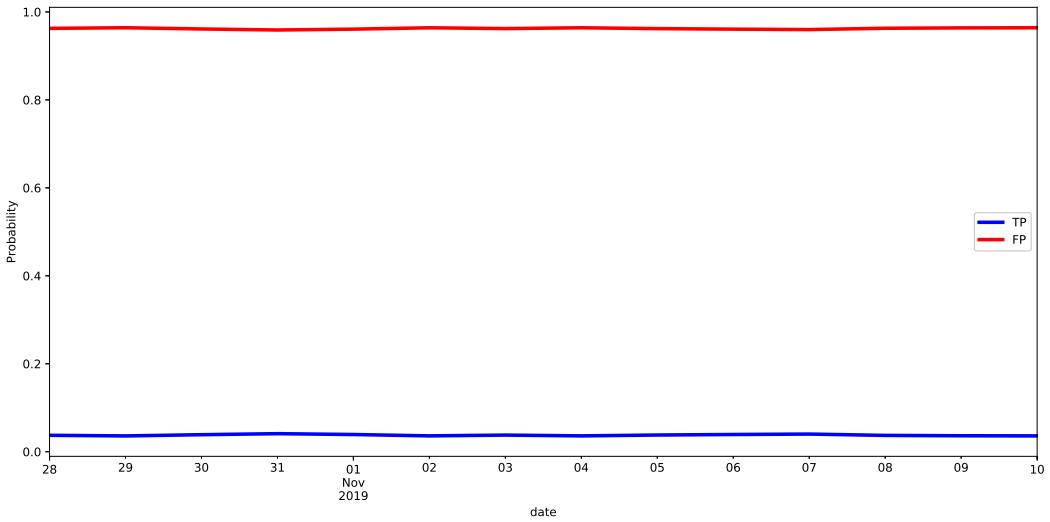


Figure 6.7: Daily validation for full malicious subnet/24 within Area 1

Relatively malicious /24 subnets

This section analyzes /24 subnets that contain a relatively low number of benign hosts. As a result of this, answering the question: How well do relatively malicious /24 subnets perform? Figure 6.8 shows the daily number of domains that are detected by the Geometrical method that originates from /24 subnets that contain a relatively low benign domain count. The blue line represents the number of domains that are not listed on RBL. This analysis indicates that the number of domains is reasonably constant, and there is no decline.

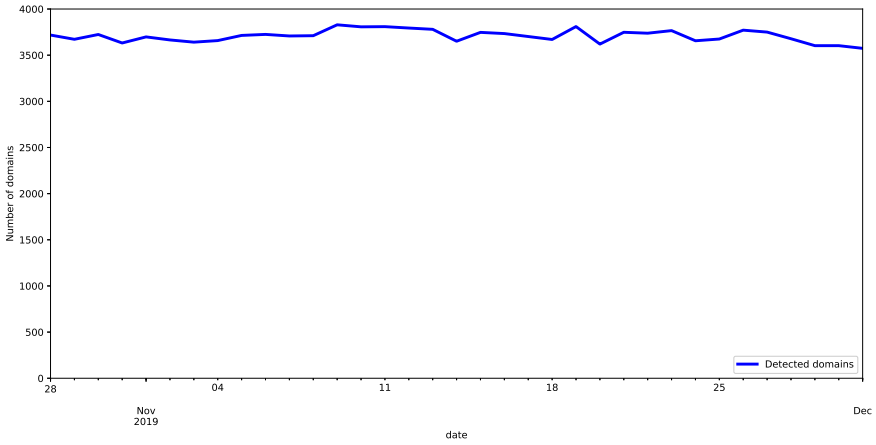


Figure 6.8: Daily detection of /24 subnets with a relative low number of benign hosts

Figure 6.9 shows the true and False Positive rate for relative malicious /24 subnets. The red line represents the False Positive rate and the blue the True Positive. This analysis indicates that the True Positive rate has a low daily value of 97.5%, and the number of False Positives is 2.5%. This result indicate these domains are not listed yet or they are part of collateral damage.

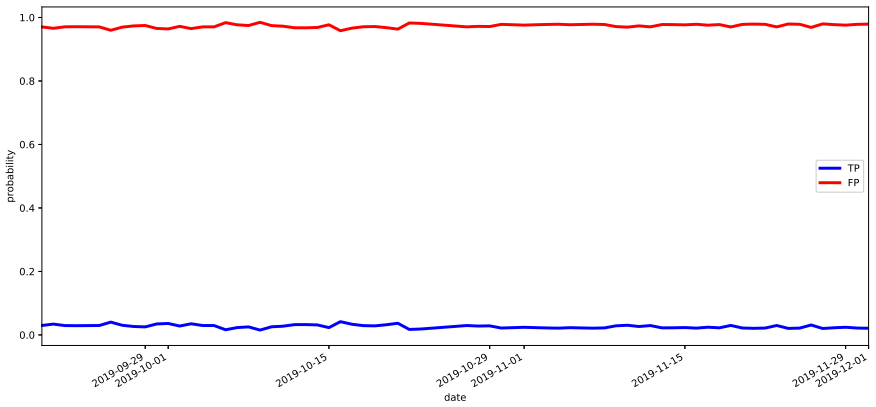


Figure 6.9: Daily True and False Positives of /24 subnets with a relative low number of benign hosts

6.7.3 Summary

To see if the adapted bad neigborhood concept can be used as a standalone detection model, validation was done on the domains originating from the /24 subnets. Summarising the evaluation results, we have concluded that the approach does not perform very well as a standalone model. It emerges from the evaluation that the validation rate of the model was low. Using the thresholds set, it has shown that we achieved a True Positive of 3% and a False Positive rate of 97%. In order to see how the different /24 subnets performed, we separated them into subnets that have no benign domains or a relatively low number of benign domains by matching the malicious /24 subnets with ALEXA. Similarly, these ended up with a low true and false-positive rate. The analysis indicated that the entirely malicious /24 subnets have a False Positive of 96.5% and True Positive of 3.5%. In the same way, relative malicious /24 subnets have a False Positive of 97.5% and True Positive 2.5%. The validation results indicate that aggregating IP addresses associated with domain blacklist are not the most efficient way of finding suspicious domains due to the results show that allot of benign domains being dragged along that cannot be identified as malicious or benign. The high number of detected domains that were not validated could be for various reasons. First of all, the domains detected could be hosted on ”malicious infrastructure” and could be collateral damage. Secondly, the domains detected are not yet listed as malicious by the validation databases, so the lack of blacklist the validate the classified domains.

Chapter 7

Conclusion

This chapter gives a conclusive summary of this thesis. It pinpoints the contribution of analyzing the active DNS data features on ALEXA and RBL to make for a useful signature. Further, it analyzes how the bad neighborhood can be adapted to domains and if it can function as a standalone detection method for malicious domains.

7.1 Summary

With the rapid expansion of the internet and daily creation of millions of domains, security is becoming more critical. This also brings an exponential increase in malicious activity, which according to a study done by Akamai, leverage DNS for 90 % [23]. These malicious activities consist of phishing domains, randomly generated domains, and DDoS attacks. To detect these domains, various methods have been introduced, such as web-based detection and DNS based detection. Though most DNS based detection methods use passive DNS data, which yields good results, another method would be the use of active DNS data because it provides a complete view of domain name service. It allows for preemptive detection of malicious domains. In this thesis, we have analyzed active DNS data to find a useful signature that can be used in the preemptive detection of malicious domains. To reach the goal of this thesis, three Research Questions (RQ) were defined:(1) How much statistical difference can be observed between the active DNS data features on Alexa and RBL? (2) Can the concept of the bad neighborhood be adapted for domains? If yes, can witness any form of bad neighborhoods forming inside RBL data (3) How effective is the use of domains originating from the bad neighborhood as a valid standalone method to detect future malicious domains.

In order to observe the difference between the active DNS data features on ALEXA and RBL, an analysis is performed on 27 features that were selected and analyzed using CDF plots, and the results validated using KS test. These features were separated into four categories, namely DNS record count, Network based, DNS record TTL, and Domain name based. When analyzing DNS records, Network based, and domain name based no significant deviation was detected between these features on ALEXA and RBL. The TTL based features did show a statistically significant deviation, but one feature is not sufficient to create a use full signature. We have concluded there is no statistical deviation found in the DNS record based, DNS answered based and domain name based features with the method used. This could be a result of values averaging each other out due to merging without consideration of the threat the blacklist was build for. We conclude that merging these blacklists is not useful for extracting features if the criteria of the blacklists are not taken into consideration. However, the blacklist are still useful for the validation of current or future malicious domains.

As a result of not finding any relevant deviation in the DNS features analyzed using CDF plots, it was not possible to extract a useful feature. This led to the investigation of malicious clustering on an IP level. To analyze this clustering. First, these subnets were plotted along the Hilbert curve. Although clustering could be observed, this did not satisfy the bad neighborhood concept. After that, the traditional bad neighborhood concept was applied by aggregating the IP addresses under /24 subnet and measured the number of Hosts occupying the subnet. The limitation of this approach on our data set is that it did not take into account the number of hosted domains originating from the /24 subnets. This resulted in /24 subnets that were very utilized by malicious hosts but relatively hosted a low number of malicious domains and a very high number of benign domains. The adapted bad neighborhood concept for domains measured the number of hosts utilizing a /24 subnet and the number of domains. The /24 subnets were then plotted along a scatter plot to visualize the clustering of /24 subnets. As a result of this, thresholds were placed based on six months of RBL data. These thresholds placed formed our adapted bad neighbourhood concept and pointed where the bad neighborhoods reside.

Now that the presence of bad neighborhoods is confirmed using thresholds set within the scatter plot, the next step was validating if the geometrical model can function as a standalone method for detecting malicious domains as a result of this answering RQ3. The model detects domains that malicious and not listed on the RBL. This method observes domains that are associated with the bad neighborhoods which are filtered by our method. These domains are considered suspicious due to being associated with the filtered /24 subnets. The validation is performed on the model under the following thresholds. The first threshold is set between

[22000,40000] domain counts and a host count from [0,255]. The second threshold is set on a host count between [180,255] and domains from zero and up. After the thresholds are set, the watch lists are generated. There generated watch lists are then compared with the /24 Alexa to reduce the number of benign domains that are hosted under the same /24 subnet. These watch lists were matched using OpenINTEL and generated a list of suspicious domains that were validated daily.

The validation for the geometrical model gave an average true positive of 3% and False positive of 97%. Furthermore, the different /24 subnets are analyzed based on them being entirely or partially malicious. For the completely malicious /24 subnets, the True positive is 3.5%, and the False positive is 96.5%. Similarly, for relatively malicious subnets, results in a true positive of 2.9% and false-positive of 97.1%. With these results, we concluded that the geometrical model could not function as a standalone detection method for malicious domains due to the very low True positive and very false positive rate. This could be a result of not being able to validate the suspicious domains due to lack of blacklist, or it could be that classified domains are not listed as malicious yet, therefore, not finding them on any of the blacklist used in the validation.

7.2 Recommendations

This thesis shows that the extracted active DNS data features do not show any significant deviations between ALEXA and RBL. Moreover, the features analyzed provide a sufficient indication of the relation of the RBL to Alexa. It should be noted that if this RBL is used for characterization of malicious domain behavior the criteria of the aggregated blacklists used should be taken into consideration. Furthermore, it is not recommended to use the Geometrical model as a standalone detection method for malicious domains due to its very high false-positive rate. The geometrical model gave a high number of false positives and a low number of true positives. Since these are hosted on ”malicious infrastructure, it would be interesting to see if these domains are collateral damage from a web hosters or just domains that are not detected by the RBL.

Bibliography

[1] G. C. Moura, R. Sadre, A. Sperotto, and A. Pras, “Internet bad neighborhoods aggregation,” in *2012 IEEE Network Operations and Management Symposium*. IEEE, 2012, pp. 343–350.

[2] R. van Rijswijk-Deij, M. Jonker, A. Sperotto, and A. Pras, “A high-performance, scalable infrastructure for large-scale active dns measurements,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 6, pp. 1877–1888, 2016.

[3] “Internet facts.” [Online]. Available: https://hostingfacts.com/internet-facts-stats/

[4] Y. Zhauniarovich, I. Khalil, T. Yu, and M. Dacier, “A survey on malicious domains detection through dns data analysis,” *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, p. 67, 2018.

[5] F. Weimer, “Passive dns replication,” in *FIRST conference on computer security incident*, 2005, p. 98.

[6] A. Kountouras, P. Kintis, C. Lever, Y. Chen, Y. Nadji, D. Dagon, M. Antonakakis, and R. Joffe, “Enabling network security through active dns datasets,” in *International Symposium on Research in Attacks, Intrusions, and Defenses*. Springer, 2016, pp. 188–208.

[7] P. Mockapetris, “Domain names - concepts and facilities,” Internet Requests for Comments, RFC Editor, STD 13, November 1987, http://www.rfc-editor.org/rfc/rfc1034.txt. [Online]. Available: http://www.rfc-editor.org/rfc/rfc1034.txt

[8] ——, “Domain names - implementation and specification,” Internet Requests for Comments, RFC Editor, STD 13, November 1987, http://www.rfc-editor.org/rfc/rfc1035.txt. [Online]. Available: http://www.rfc-editor.org/rfc/rfc1035.txt

[9] P. Mockapetris and K. J. Dunlap, *Development of the domain name system*. ACM, 1988, vol. 18, no. 4.

[10] “Domain name system white list.” [Online]. Available: https://www.dnswl.org/

[11] "Creating Customized Whitelist Domains from DNS Traffic." [Online]. Available: https://www.microfocus.com/media/white-paper/creating_customized_whitelist_domains_from_dns_traffic_wp.pdf

[12] "Blocking Spammerswith DNS Blacklists." [Online]. Available: https://cdn.ttgtmedia.com/searchDomino/downloads/3167Xc05.pdf

[13] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster, "Building a dynamic reputation system for dns." in *USENIX security symposium*, 2010, pp. 273–290.

[14] 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.

[15] E. Passerini, R. Paleari, L. Martignoni, and D. Bruschi, "Fluxor: Detecting and monitoring fast-flux service networks," in *International conference on detection of intrusions and malware, and vulnerability assessment*. Springer, 2008, pp. 186–206.

[16] T. Holz, C. Gorecki, K. Rieck, and F. C. Freiling, "Measuring and detecting fast-flux service networks." in *NDSS*, 2008.

[17] S. Hao, N. Feamster, and R. Pandrangi, "Monitoring the initial dns behavior of malicious domains," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 269–278.

[18] D. Chiba, T. Yagi, M. Akiyama, T. Shibahara, T. Yada, T. Mori, and S. Goto, "Domainprofiler: Discovering domain names abused in future," in *2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 2016, pp. 491–502.

[19] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Beyond blacklists: learning to detect malicious web sites from suspicious urls," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 1245–1254.

[20] W. Van Wanrooij and A. Pras, "Filtering spam from bad neighborhoods," *International Journal of Network Management*, vol. 20, no. 6, pp. 433–444, 2010.

[21] O. van der Toorn, R. van Rijswijk-Deij, B. Geesink, and A. Sperotto, "Melting the snow: Using active dns measurements to detect snowshoe spam domains," in *NOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium*. IEEE, 2018, pp. 1–9.

[22] J. Nazario and T. Holz, “As the net churns: Fast-flux botnet observations,” in *2008 3rd International Conference on Malicious and Unwanted Software (MALWARE)*. IEEE, 2008, pp. 24–31.

[23] Akamai, “Data Revelations.” [Online]. Available: https://www.akamai.com/us/en/multimedia/documents/report/carrier-security-report-fall-2016.pdf