UNIVERSITY OF TWENTE.

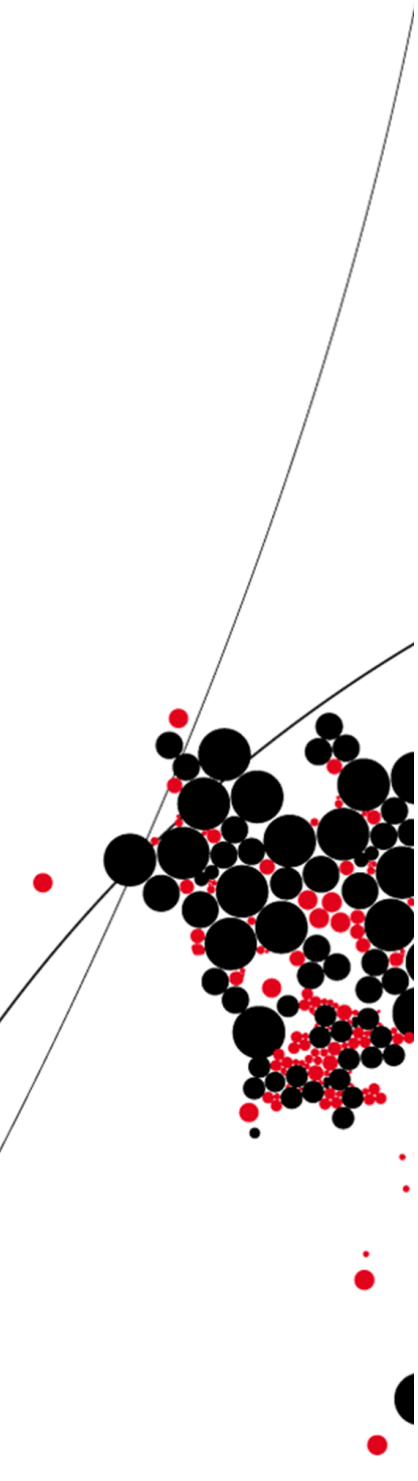
Faculty of Electrical Engineering, Mathematics & Computer Science

The evolution of DNS encryption and the effect on digital forensic research

Anja Scherjon - 434766

Bachelor Thesis - Forensic Research

June 2020



SAXION UNIVERSITY OF APPLIED SCIENCES

DACS Design and Analysis of Communication Systems

Supervisors:
Dr. ir. Roland van Rijswijk-Deij
Dr. ir. Richard Brinkman

Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
Drienerlolaan 5
7522 NB Enschede
The Netherlands

Preface

I found this research because I went looking for prof. Aiko Pras, chair of the Design and Analysis of Communication systems (DACS) group at the University of Twente, after I heard his presentation on ”how a cyberattack could bring down the internet” at the ”Universiteit van Nederland” podcast. When I contacted him, I was welcomed with open arms. An effort was made to find an appropriate topic for my graduating thesis and I am very happy with the research and everything I have learned from the DACS group at the University of Twente. Even though the COVID-19 pandemic hit after the first month of this research, I am proud of the accomplished results. Firstly, I want to thank my supervisor Roland van Rijswijk-Deij for the constant help and support during this process. I also want to thank Etienne Khan for helping me with all my questions and getting the experiment working and measured. Lastly, I want to emphasise the flexibility shown by the experts, and their willingness to do the interviews digitally. I hope all the hard work shines through.

Summary

The Domain Name System (DNS) is responsible for connecting domain names to IP addresses. Your computer sends the (typically) unencrypted DNS request (query) to a resolver. This resolver will ask different servers for the allocated IP address and the response will be sent back to your computer. Recent efforts have worked towards including encryption in the DNS, with the latest protocols being announced in 2016: DNS-over-TLS (DoT) and in 2018: DNS-over-HTTPS (DoH). These protocols use TLS and HTTPS to encrypt the DNS data between the computer and resolver. This could have a major impact on the visible DNS information for incident researchers and law enforcement. Therefore, the research question is: "What are the main restrictions encountered by incident researchers from the encryption of DNS and what are alternatives to gain the wanted information?" This will be answered by researching: current leading types of DNS encryption, currently available investigative options and use of DNS information, experts experiences and advice when encountering these forms of DNS encryption and the potential misuses of encrypted DNS. Results are gathered through qualitative interviews and a quantitative experiment. The main results from the interviews show that most experts were not fully aware of the possible impact of the new encryption protocols. Different ways to circumvent the encryption were stated. However, most have not chosen to implement these yet. It can be concluded that companies that use data gathered between client and resolver will experience the most effect from these encryptions of the DNS. Real-time detection and trace-backs to the original client will become increasingly difficult. The best alternatives as stated by the experts are: changing the location of detection to endpoints and the resolver itself, and offering DoT and DoH within the local network. The main result from the experiment is that it is possible to set up a DNS tunnel over DoH in a day. Therefore, it can be concluded that it is trivial to use the new encryption for criminal activity like the in- and exfiltration of data or control of a botnet by using a DNS tunnel. Further research should look into the effect DoT and DoH will have on the available threat intelligence and other passive DNS sources; the effect of other security aspects after the encryption that have not specifically been looked into in this research. Finally, more possibilities to limit DNS tunneling, normally and over encryption and more software for local DoT and DoH resolvers should be developed and implemented.

Samenvatting

Het Domain Name System (DNS) is verantwoordelijk voor het zoeken van IP-adressen bij domeinnamen. Het werkt door onversleutelde verzoeken (query's) te versturen naar een resolver. Deze resolver zal bij verschillende servers het bijbehorende IP-adres opvragen en het antwoord naar de computer terugsturen. Twee grote nieuwe protocollen voor DNS-encryptie zijn gepresenteerd in 2016: DNS-over-TLS (DoT) en in 2018: DNS-over-HTTPS (DoH). Deze protocollen gebruiken TLS en HTTPS om data tussen de computer en resolver te versleutelen. Dit kan een grote impact hebben op incidentonderzoekers, daarom is de hoofdvraag van dit onderzoek: "Wat zijn de voornaamste beperkingen die incidentonderzoekers tegenkomen na de encryptie van het DNS, en wat zijn alternatieven om deze informatie toch te verkrijgen?" Deze vraag zal beantwoord worden door te kijken naar de laatste vormen van DNS-encryptie, het verkrijgen en gebruik van DNS-informatie, ervaringen en advies van experts over de omgang met de nieuwe encryptievormen en misbruik van versleuteld DNS. De resultaten zijn verzameld aan de hand van kwalitatieve interviews en een kwantitatief experiment. De voornaamste resultaten van de interviews laten zien dat de meeste experts zich niet volledig bewust waren van de mogelijke impact van de nieuwe versleutelingen. Verschillende manieren werden genoemd om de informatie alsnog te verzamelen. Deze zijn echter veelal nog niet geïmplementeerd. Er kan worden geconcludeerd dat bedrijven die informatie verzamelen tussen de client en resolver het meeste effect van de DNS-encryptie zullen ervaren. Realtime detectie en het achterhalen van de originele client zullen moeilijker worden. De beste alternatieven genoemd door de experts zijn: het verplaatsen van de sensoren naar de endpoints of resolvers toe, en het aanbieden van DoT en DoH op lokale resolvers. Het voornaamste resultaat van het experiment is dat het mogelijk is om een DNS tunnel over DoH op te zetten in een dag. Er kan worden geconcludeerd dat het gemakkelijk is om de encryptie voor criminele activiteiten in te zetten zoals het in- en exfiltreren van data via een DNS tunnel. Vervolgonderzoek kan kijken naar: het effect van DoT en DoH op de beschikbare threat intelligence en andere Passive-DNS-bronnen; het effect op andere beveiligingsaspecten na de encryptie waarnaar niet is gekeken in dit onderzoek. Tot slot moeten er meer mogelijkheden komen om DNS tunneling te beperken en moet er meer software voor locale DoT en DoH resolvers ontwikkeld en geïmplementeerd worden.

Contents

Preface iii

Summary v

Samenvatting vii

List of acronyms xiii

Glossary xv

1 Introduction 1
- 1.1 Context . . . 1
- 1.2 Problem statement . . . 2
- 1.3 Goal . . . 3
- 1.4 Research questions . . . 3
- 1.5 Research setup . . . 4

2 Background 5
- 2.1 What Is DNS? . . . 5
- 2.2 The encryption of DNS . . . 7
  - 2.2.1 DNS-over-TLS . . . 7
  - 2.2.2 DNS-over-HTTPS . . . 8
- 2.3 DNS Tunneling . . . 9

3 Expert interviews 11
- 3.1 Methodology . . . 11
- 3.2 Limitations . . . 12
- 3.3 Results . . . 13
  - 3.3.1 General detection system . . . 14
  - 3.3.2 Data collection . . . 14
  - 3.3.3 Data processing . . . 15
  - 3.3.4 Importance and effect . . . 15

3.3.5 Encryption 16
3.3.6 Criminality and alternatives 17
3.3.7 General conclusion 17

4 Practical experiment 19
4.1 Experiment design 19
4.2 Methodology 19
4.2.1 Stage 1: Setting up the tunnel 20
4.2.2 Stage 2: Setting up the DoH stub resolver 22
4.2.3 Combining the tunnel and the proxy 23
4.3 Results 25
4.3.1 The rst stage 25
4.3.2 The second stage 26
4.3.3 General conclusion 27

5 Discussion 29

6 Conclusion 31

7 Recommendations 33

References 35

Appendices

A Interview questions 39
A.1 Conducting the interview 40

B General coding overview 41

C Speci c coding overviews 43
C.1 The general detection system 43
C.2 Data collection 44
C.3 Data processing 44
C.4 Importance and effect 45
C.5 Encryption 46
C.6 Criminality and alternatives 47

D setup and settings on experiment software 49
D.1 Iodine on the client 49
D.2 Iodined on the server 50
D.3 Dnscrypt-proxy on the client 51

E Located IP addresses 53
E.1 Authoritative name server 53
E.2 University of Twente 54
E.3 Cloud are incoming 54
E.4 Cloud are outgoing 55

F Interview Transcripts 57

List of acronyms

CERT	Computer Emergency Response Teams
CSIRT	Computer Security Incident Response Teams
DNS	Domain Name System
DoH	DNS-over-HTTPS
DoT	DNS-over-TLS
HTTP	HyperText Transfer Protocol
HTTPS	HyperText Transfer Protocol Secure
IETF	Internet Engineering Task Force
ISP	Internet Service Provider
LEA	Law Enforcement Agencies
NCSC	National Cyber Security Centre
NSA	National Security Agency
RFC	Request for Comments
THTC	Team High Tech Crime
TLS	Transport Layer Security
US	United States

Glossary

Threat intelligence
Information an organization uses to understand the threats that have, will or are currently targeting the organization. Sources can be open source intelligence, social media intelligence, intelligence from the deep and dark web and private or commercial sources.

Passive DNS
Records that contain DNS resolution data for a given location, record, and time period. This historical resolution data set shows domains resolved to an IP address and vice versa. This data set allows for time-based correlation based on domain or IP overlap.

Cobalt Strike
Software for security assessments that replicate the tactics and techniques of an advanced adversary in a network.

(Threat) Actors
A person, group, or entity that attempts to or successfully conducts malicious activities against enterprises.

Encryption
The process of encoding information known as plaintext, into an alternative form known as ciphertext. Only authorized parties can decipher a ciphertext back to plaintext and access the original information.

Chapter 1

Introduction

1.1 Context

In 2013 Edward Snowden shocked the world. In a series of releases he exposed multiple secret projects the United States (US) government was running to spy on people around the world [1]. One of the biggest revelations was the existence of multiple National Security Agency (NSA) surveillance programs. These consisted of two categories: wiretaps that pull data directly from the undersea telecommunication cables, and programs such as PRISM that gather information directly from US service providers such as Google, Facebook and Microsoft [2]. Since these revelations and the more vital role the Internet is playing in our society, the aspects of security and privacy have become more and more of a priority.
In the early days of the Internet, security and privacy were not much of a concern. However, the more the Internet grew, the more people became concerned about their personal information. In 1999 concerns about the safety of personal information and privacy of communications were the main reasons many consumers stayed off the Internet [3]. The main concern was that unauthorised people would nd and use the information without consent. At this point the concerns were mainly focused on the companies owning the websites and malicious intruders [3]. This changed in 2013 with the Snowden revelations [1]. These showed that intelligence agencies were gathering large amounts of data from civilians. This sparked a debate on the state of legislation concerning Internet privacy and security [1]. The technicians, however, went back to an old principle: ”The best solution is privacy through technology, not through legislation. The objective must be to bring privacy to the Internet, and bring the Internet to everyday practices [3]”. Over the years, countless efforts have been made to protect data in transit. One of the biggest efforts was the encryption of connections through Transport Layer Security (TLS) [4].
More recently, a vital Internet infrastructure, the Domain Name System (DNS), has also been the focus of privacy improvements. The DNS is responsible for looking up at

which IP address the domain name you are trying to find is, so your computer can connect to that IP [5]. Data can give great insight into the Internet behaviour of individuals such as which social media, banks and other websites are visited. The introduction of encryption through DNS-over-TLS (DoT) in 2016 [6] and DNS-over-HTTPS (DoH) in 2018 [7] by the Internet Engineering Task Force (IETF) have made huge changes to how much, and to whom, DNS data is visible. The implementation of the encryption of the DNS could have a major impact on the research possibilities of incident researchers and law enforcement [8]. This will be further examined in this research.

1.2 Problem statement

The encryption of the DNS has made it so secure it can not be “tracked, spoofed, or blocked [9]”. DNS encryption can take the power away from the Internet Service Provider (ISP) that could exploit your data and complicates censoring by state-run ISPs [9]. However, It has also received a lot of criticism. Although the encryption eradicates snooping on the wire, the information is still visible on the used resolver, such as the one from the state-run ISP. Therefore, mainly in the use of DoH, centralised public resolvers such as Cloud are and Google are being used [10], [11]. The data on the wire is now encrypted and is sometimes no longer stored within the national borders. This means that other laws and jurisdictions come into play [8], [11].
The value of DNS information in current incident investigations is hard to determine. In an overview from 2012, Wright stated: ”DNS logs, where captured, have value in confirming browsing behaviour versus malware behaviour, identifying system configuration, as well as providing time-line data for investigations” [12]. However, this was before the Internet started becoming more secure with, for example, more of a widespread implementation of TLS for the HyperText Transfer Protocol (HTTP) [13]. This meant that nowadays, less information can be gathered from the HTTP protocol and the value of DNS data has increased.
The possible influence of the encryption of DNS on law enforcement is stated in a recent overview by Europol in 2019 [8]. They state: “The DoH protocol can affect the judicial use of DNS query history in relation to malware investigation, lawful interception, and blocking of IP addresses linked to malware or child sexual exploitation material. It should be noted that if DNS resolution continues to be local and encrypted; Law Enforcement Agencies (LEA)s will continue to access the data through appropriate judiciary requests to the ISP. However, if there is a case of a remote resolver being used for DNS resolution, the data will not be accessible to national authorities”. However, they do not go in depth about the specifics of how DNS data is currently used in investigations or alternative options for still gathering the wanted data after DNS encryption.

Moreover, to the best of our knowledge, no complete and recent information on the use of DNS information in current incident research can be found. Therefore it is also hard to nd how incident researchers are adapting to the new implementations of DNS encryption and the availability of data. This will be the main focus of this research.

1.3 Goal

The goal of this research is to create an overview of the use of DNS information by digital incident responders and researchers within the Netherlands. This overview will cover how they gather DNS information, what they use the information for and exploratory questions on encounters with the encryption and expected new criminal techniques. Next, the effects of DNS encryption will be examined by comparing how the different ways of encryption can affect the techniques and information currently gathered. Lastly, this gathered information will be used to look into alternative methods for retrieving the wanted information and possible misuses of the new encryption form.

1.4 Research questions

Main research question

What are the main restrictions encountered by incident researchers due to the encryption of DNS and what are alternatives to gain the wanted information?

Sub research questions

1. What are currently the leading types of DNS encryption in development and being implemented?
2. What are the available investigative options for incident researchers at this moment to gain (unencrypted) DNS information and what is it used for?
3. What are experts experiences and advice when encountering encrypted DNS at the moment and in the future?
4. What potential misuse of encrypted DNS can be expected?

1.5 Research setup

The research is conducted in three stages. The rst stage consists of a literature study into the different types of DNS encryption currently being implemented. This information can be found in Chapter 2, together with background information on the DNS. In the second stage, Chapter 3 and 4, this information is used to gather the real-life implications of these encryption types. The second stage consists of interviews with digital incident responders and researchers to gain an overview of how they gather and use DNS information. The interviews also investigate the effect the encryption can have on the current possibilities in digital forensic research, developments in criminality and alternatives after encryption has been implemented fully. This is investigated for law enforcement agencies as well as private companies. The third stage consists of an experiment and investigates possible misuse of the encryption. By testing how easy it is to set up a DNS tunnel for data in- and ex ltration over DoH.

Chapter 2

Background

In this chapter the most important topics of this research are explained. It will start off with a more detailed explanation of the DNS system. Next, a general explanation of encryption in the DNS and two speci c types of DNS encryption: DoT and DoH, are explained. Lastly, a speci c use of the DNS, DNS tunneling is explained.

2.1 What Is DNS?

Every protocol on the Internet is speci ed in a so-called Request for Comments (RFC), which are documents maintained by the IETF. In November of 1987 the two main RFCs for the DNS were introduced [5], [14]. In these documents, the concepts, implementations and speci cations of the protocol are explained. A simpli ed version is given below.

The DNS is generally described as the “phone book” of the Internet. This entails that every time you want to connect to a page or service there is a “lookup” that happens rst to see where you exactly want to connect to. This lookup, through recursive DNS, consists of a couple of steps and is also visible in Figure 2.1.

The rst step is that you, the client, type in an domain in your browser you want to connect to, like www.example.com, this is your DNS request. Your browser then connects to the stub resolver built into your client, this is a process that is part of your operating system designed to handle DNS requests. If you have not searched for it recently, it will send the request to a local resolver (1).

The next step is that the resolver, which likely is supplied by your ISP, will try to resolve your request by asking authoritative name servers if they have the IP address connected to your requested domain. This lookup will be from the root down.

Figure 2.1: DNS recursion overview

If you look at the requested domain www.example.com in reality there is an extra (invisible) dot at the end of it (www.example.com.). This dot at the end represents the root. This is the top of the DNS hierarchy and signifies a starting point. When a lookup is done the resolver will first request the root server for the information (2). Most likely, the root server will not have the IP address for your specific request, but will know the top-level domain you requested and will send the location for its authoritative name server back to your resolver (3). In this case it would be for the .com domain. Again, this server will not have the specific IP address for example.com (4) but will know its authoritative name server (5). Eventually the request will be send to the example.com authoritative name server (6) that will know the specific IP address related to www.example.com. This will be sent to the resolver (7) who will inform the stub resolver and therefore the browser on your computer and the connection can be established (8). The resolver and stub resolver typically keep a database of the most resent requests (cache) so the entire process does not have to be executed for every request every time.

2.2 The encryption of DNS

Since the DNS system was first specified in 1987 it has not changed much. The main design goals consisted of consistency, dealing with size, costs of acquiring data, availability for different protocols, independence and host capabilities [5]. This shows that in that time, privacy was not an aspect considered. In 2015, the IETF made an RFC containing an analysis of the privacy issues associated with the use of DNS by Internet users [15]. This overview showed the best place to “eavesdrop” on information is between the stub resolvers and the recursive resolvers. This is because traffic at this point is not limited by DNS caching and everything is in plain text. To try and change this ease of eavesdropping, two major ways of encrypting DNS have been implemented in recent years. More ways of encrypting DNS are available, but in a comparison of these protocols “DoT and DoH are two leading and mature protocols to secure traditional DNS communications. On top of well-supported and standard protocols, they are both standardised by the IETF and extensively implemented by various DNS software and public resolvers [16]”. Therefore, they are most relevant and the focus of this research.

2.2.1 DNS-over-TLS

The first approach to encrypting part of the DNS data was DNS-over-TLS (DoT) [6]. The concept for this protocol is that first a TLS connection will be established and multiple DNS requests can be sent over this encrypted, established connection. For this connection, only port 853 will be used [6]. The advantages of DoT are that clients and resolvers can exchange encrypted DNS requests and resolvers can be authenticated by verifying TLS certificates. By padding the message no information can be deduced through data analysis [17]. However, by using a designated port, the DNS traffic is still distinguishable from other traffic [16]. Also it has been found that “25% of the DNS-over-TLS service providers use invalid SSL certificates [16]”. This also influences the reliability of this encryption protocol. Most mainstream DNS software supports DoT, which allows its deployment to grow. While a major mobile operating system supports DoT on the client side (Android), the two major operating systems (Windows and MacOS) still do not, which limits the use and growth on the client side [16].

2.2.2 DNS-over-HTTPS

The second and still upcoming way of encrypting DNS is DoH [7]. This, like DoT, uses the TLS protocol but makes the connection through the HyperText Transfer Protocol Secure (HTTPS) application protocol. Therefore, it is allowed to send the requests through port 443, which makes it harder to distinguish from other HTTPS traf c [7]. DoH creates an encrypted connection between client and resolver and stops network operators from blocking DNS encryption by restricting the use of certain ports (like 853 for DoT) [18]. Currently, the blocking of DoH is often still possible because services are offered on a separate server. For example, Google is currently at https://dns.google.-com/dns-query, this is a separate server so still blockable. However, if Google changed this to https://google.com/dns-doh it would become impossible to block. To block DoH traf c, the only indicator visible is the connection to the google.com server, and by blocking this all other Google services would also be blocked.
Previously, with classic DNS and DoT, applications used the operating system to perform DNS. DoH enables applications to make DNS requests to any chosen server that supports DoH without use of the operating system. This means that application providers can choose the server and its deployment and policy choices [8]. Applications often choose to use public DNS resolvers offering DoH such as Google and Cloud are. The request sent to these big central resolvers, which are at the moment mostly located in the United States, might result in DNS requests not being resolved within a country's borders. This can bring dif culties for Law Enforcement Agencies (LEAs) of other countries to of cially request data or get access through the judicial system [8]. The DoH protocol also has some other dif culties, the main one being: “It can affect the judicial use of DNS query history in relation to malware investigation, lawful interception, and blocking of IP addresses linked to malware or child sexual exploitation material [8]”. Though there are large amounts of criticism given on the protocol [8], [10] the implementation is still growing and expected to keep growing for the foreseeable future [16].

2.3 DNS Tunneling

DNS Tunneling is a form of a covert network channel. These covert channels can be used to bypass rewalls, hide data for con dentiality, anonymity or to counter censorship [19].
DNS Tunneling can be done by using software to set up a private network between a client and a server. To make this possible, the receiving side needs to have a domain name server functioning as an authoritative name server. This server also needs to facilitate server side tunneling and decoding programs [20]. In this research the focus will be on tunnels for the use of transferring data inside hostnames (data ex ltration and in ltration). This is most often used in malware and in communication with botnets [21]. By encoding the wanted data with base32 [22], and putting it in front of the ”normal” query such as base32(data).example.com, data can be ex ltrated without being detected. In the response of a DNS query, data can also be encoded with base64 for the in ltration of data. In the query response, commands from the command and control server can also be sent to a botnet.

Chapter 3

Expert interviews

In this chapter the full process of the interviews with experts is shown. First, the methodology of the interviews is shown. Next, it shows how the gathered interview data was processed and the experienced limitations. Finally, the results given are divided into subsections to create an overview of the answers sorted by topic.

3.1 Methodology

The approach to determining the experts views on the use of DNS information, the expected effect of the DNS encryption, developments in criminality due to the encryption and alternative ways to gather information after encryption are qualitative interviews in a semi-structured way. Ahead of the interviews, an interviewing procedure and a list of 17 questions were made. These can be found in Appendix A. During the interview, the interviewer went along with the answers given and asked extra questions to go more in depth on interesting and relevant topics. [23].
All interviews were held digitally[1] using different online platforms. When agreed to, the interviews were recorded and in all cases the interviewer wrote along to capture the most important aspects.
Afterwards, the interviews were transcribed in summary and anonymized. These transcriptions can be found in Appendix F. Thereafter, the answers were coded into general categories related to the sub research question (Appendix B) whereupon the results were reformatted into tables that gave a clear overview of the answers given (Appendix C).

[1] Due to the Corona virus in person interviews were not possible

To get a broad view of the field of incident researchers that use the DNS system for different purposes, three different groups were formed. From each group, two to three companies were interviewed. The three groups are:

Law enforcement and intelligence services , these are researchers that work for one of the Dutch ministries and work from a governmental perspective. The interviews were conducted with members of Team High Tech Crime (THTC) of the Dutch national police and the Dutch National Cyber Security Centre (NCSC).

In-house security teams , these are researchers for Computer Emergency Response Teams (CERT) or Computer Security Incident Response Teams (CSIRT) from universities, schools or companies and they keep an eye on the internal safety and incident response. The interviews were conducted with members of the security team of the University of Twente and the network maintenance of Saxion University of Applied Sciences.

Companies facilitating a service , these are researchers working for a company providing security services and consultancy. They are selected by the more well known companies with a focus on Internet safety and incident response. They have also been screened for having a forensic investigative service within the company. The interviews were conducted with staff members of Tesorion, Fox-IT and Northwave.

3.2 Limitations

This research has potential limitations. The most important are stated below:

Limited sample size
Because of a limited time of 5 months and the start of the Corona virus pandemic, there was a limit to the availability and amount of experts to be interviewed. This restricts a strong statistical analysis of the results. However, by choosing a broad spectrum of experts, the results will still give a general overview representative to the field.

Limited comparability of results
Because the chosen experts and way of interviewing, experts gave a broad set of possible answers to the asked questions. This is a common limitation with semi-structured interviews. Their experience and way of thinking resulted in less comparable answers. However, this gave a broader view of the field and different approaches to the DNS data.

3.3 Results

The questions were made in a way, that from the transcriptions the coded answers could be divided into six main categories. These six categories are:

1. The general detection system
   In this category a general overview of the experts systems will be given. This, to show how it is set up and to give a general view of the situation at that company. This will give more insight into the answers given in the other categories.

2. Data collection
   In this category the types of DNS data and the way the DNS data is collected will be given. This, to show what the sources are, and which sources might experience impact from the new encryption.

3. Data processing
   In this category the way the DNS data is used or processed will be given. This, to show how the system uses the information and responds to it.

4. Importance and effect
   In this category the speci c usefulness and possible effects of the loss of DNS data is shown.

5. Encryption
   In this category everything about the new forms of DNS encryption is shown. It will look at the effect, how much it is observed in the eld and the implementation and development expected by the experts.

6. Criminality and alternatives
   In this last category the possible new forms of criminality developed because of the DNS encryption will be shown. It will also show alternatives experts could use to still be able to gain the wanted DNS information after the encryption has been implemented.

3.3.1 General detection system

The general buildup of the detection system can be divided in three main topics: the location in the recursion process, whether the organisation has control over the resolver within the network and if the experts are structural and preventative viewers of a system, or incidental viewers because of a incident or crime.
Figure 3.1 shows the possible locations for data collection and observation in the recursion process. Most experts stated that the space between the local resolver and name servers (4) is chosen. Some experts say they choose the local resolver (2) or between the client and the resolver (1). A single expert stated to want to be as close as possible to the client (1).
When interviewed about the availability of local resolvers within a network, most experts stated that the companies they work with (or company they work for in case of the UT and Saxion) do have their own local resolver. A single expert stated they do a takeover of a resolver if they want information from it.
Lastly it is good to know that almost all experts are structural and preventative viewers of data in the system. They are hired to keep an eye on the system and help protect it. Only the THTC is most often not a structural viewer in a system. They are most often present in a system to research an incident or a crime. This will also have an effect on the rest of the results given by this expert.

Figure 3.1: Observation locations

3.3.2 Data collection

The data collection is divided into two sections: Which data is being collected? and How and where is data collected?
When interviewed about which data is being collected by experts, almost all experts stated to collect domain names, IP addresses, ports and indicators. These indicators are often collected through threat intelligence. Most experts also stated to collect log les. Some experts collect metadata and owdata and network traf c from and to the Internet. When interviewed about how and where the data is collected by experts, all use threat intelligence. This is gathered by the experts from different sources: passive

DNS, Cyber threat intelligence, Virustotal, blacklists and gathering own information from, for example, e-mails. Next to threat intelligence, most experts have a device on the edge of a network, at the rewall or at the core switch (3). Most experts also gather data from DNS servers or resolvers (2). A few use a (DNS)tap to gather data.

3.3.3 Data processing

The data processing category is divided into two questions: How is the acquired knowledge used and processed? and What is the response of the system to the acquired knowledge?
When interviewed about the use and processing of the acquired knowledge, one answer was given by all experts: detection. All experts stated they use the knowledge and system for detection, some even for automated real time detection. This detection is used to detect malware, actors, anomalies, domain name generating algorithms and phishing. Next to detection, the acquired knowledge is also used by all experts to correlate marked cases with threat intelligence. Most experts stated they use the information for behavioural analysis and monitoring. Some experts stated they also use the information for research. A few stated they use the information for the tracking of servers. A single expert stated they use the information for dealing with complaints, another stated they use it for blocking requests.
When interviewed about the response of the system to the acquired knowledge, most experts stated hits or alarms are generated at possible suspicious behaviour. Most of them also stated the system tries to identify the original client that generated the hit or alarm. When identi ed, some experts stated they place the identi ed client in isolation.
A few experts stated they also use the information to check saved data retroactively. Responses of the system to the knowledge stated by a single expert are: proof of malicious activity for prosecution, checking indicators by using log les and the blocking of queries and IP addresses.

3.3.4 Importance and effect

When looking at the importance and effect, the two questions being answered are: How important is DNS information to you? and What if the DNS information was not visible anymore?
When interviewed about the importance of DNS information, some experts stated it is important for threat detection. A few experts stated the information is also available somewhere else, and it is mostly used as an indicator. A single expert stated it is very important with penetration testing software Cobalt Strike [24], and that the importance depends on the suspect and the goal.

When interviewed about what it would mean if DNS information was not visible anymore, most experts stated alternatives are needed to still get the necessary data. Some experts stated threat detection will become harder. A single expert stated that what is not there, can not be found and you need to make sure you are not dependent on the information. Lastly a single expert stated that functionalities of protection will not work anymore.

3.3.5 Encryption

When looking at the new types of encryption being implemented (DoT and DoH), three questions were asked: What will be the effect of the encryption? Have you already seen it in the eld? and What will be the implementation and development of the encryption techniques?

Figure 3.2 shows which observation locations will no longer be visible after full implementation of the encryption. When interviewed about the predicted effect of the encryption by experts, it was stated that it depends if applications will chose local encrypted resolvers over open resolvers, or by default will use open resolvers. However, most stated that more effort has to be put in to still making the wanted data visible. A few experts stated that real-time monitoring will become harder. A single expert stated that data will still leak. Another single expert stated that the endpoint will remain visible. Also, the encryption will create more single points of failure and even though you can block DoT at the edge, for DoH and cloud resolving, it is harder.

Figure 3.2: Observation locations after encryption

When looking how often it has been seen so far, some experts stated they have rarely seen DoT or DoH being used or have not been paying attention to it yet. Other responses to if it has been seen so far by a single expert state they use the encryption themselves, done research into the use of DoH by malware or blocks all DoT and DoH traf c at the rewall.

Lastly, when interviewed about the implementation and development, some experts stated it will be similar to the implementation of TLS for HTTP. A few experts stated that DoT and DoH will never take over all DNS traf c, and that getting DoT and DoH

supported and working correctly can be dif cult. A single expert stated that DoT and DoH will be blocked on networks and jurisdictional adjustments are probably necessary to work properly with DoH. A single expert also stated that less data will be available from the TLS handshake and that network monitoring will only be used for context.

3.3.6 Criminality and alternatives

This last category will look into new kinds of criminal behaviour that experts predict will arise with the new forms of encryption, and alternatives the experts can use to still gather the wanted information.
When interviewed about the new kinds of criminal behaviour that can arise with these forms of encryption, most experts stated that DoH can be used as an in- and ex ltration channel and therefore also used for command and control. A single expert stated criminals can use encryption because it is set as the default mode, or build their own system for protection and security. A single expert also stated that it can be made easier to use by generic tooling and that detecting attacks from your own network will become harder. Lastly, a single expert stated that it can facilitate spoo ng attacks by falling back to the NetBios, and cache poisoning and the validation of query and response will become harder to detect.
When interviewed about the alternatives for collecting data, most experts stated that they would change the location for the sensor to endpoints (clients) and the local DNS resolver. At these locations the unencrypted DNS queries will still be available for monitoring. Most experts also stated that some data will leak in the TLS handshake (server name and IP) and in metadata and owdata. Some experts stated that TLS interception could also be done, but all also stated they prefer not to do this. A few experts stated that offering your own DoT and DoH resolver would also be an alternative. A single expert stated that mapping and predicting criminal behaviour or demanding information at ISPs and open resolvers could also be an alternative for gaining information.

3.3.7 General conclusion

The main results of the interviews is that most experts were not fully aware of the possible impact of the new encryption systems. Most will lose important information for detection from the implementation of the new encryption systems. However, over the interviews the attitude changed and by the end they stated different ways they can circumvent these measures but most have not chosen to implement these yet. Next to that, the impact the measures will have on the available threat intelligence and passive DNS sources is still unknown and could also greatly impact the experts in their work.

Chapter 4

Practical experiment

The goal of the experiment is to show how easy it is to set up a DNS tunnel for data in- and ex ltration over DoH. In this chapter the full process of the experiment is shown. First, a more in depth look at the design is given because this is based upon results from the expert interviews stated in Chapter 3. Next the methodology is shown in which the materials used and the experimental setup are given. Lastly, the gathered results are shown.

4.1 Experiment design

The design of the experiment is based upon the answers given by experts in Chapter 3. They stated that it is highly expected that DoH will be used in the future for in- and ex ltration of data and command and control. However, they also stated criminals most often take the path of least resistance. Therefore, this experiment was designed with existing software and tooling. The goal of the experiment is to test if ex ltration by using a DNS tunnel is possible when using DoH.

4.2 Methodology

The experiment was done in two stages, the rst stage was setting up the DNS tunnel and the second stage was setting up a DoH stub resolver by using proxy software. After both stages are set up, the systems can be used together to run a DNS tunnel over DoH.

4.2.1 Stage 1: Setting up the tunnel

The first stage consisted of setting up the DNS tunnel. This needed a client and a server to work. Both the client computer and server were set up in a virtual machine and both were to run on Linux (Ubuntu). Unfortunately, on the client Ubuntu gave some difficulties with running the proxy software. Therefore, the switch to Debian was made. After some research into different software available for DNS tunnelling [25], and searching for others that have used a DNS tunnel over DoH [26], the software Iodine [27] was chosen. This software consists of a client (iodine) and server (iodined) version. For both, version 0.7.0 was used, as can be seen in Appendix D.
The setup of the server, the domain and the software was done by following the accompanied manual on github [27], and the experiences of software developer and blogger David Hamann [28].
The server needed to be an authoritative name server for a domain, which allowed for the client to query the domain and connect to the server. This server was set up to be the authoritative name server for: anja.dnsjedi.org on the IP 192.87.172.251. Next, a subdomain was created that routes all the information to the server. This domain is often shorter to allow more data to be trafficked. However, this was not the case in this experiment. In this experiment the domain: iodine.dnsjedi.org was used. The specific DNS zone file can be found in Appendix D. To start Iodined on the server, the software was installed and started by entering a password (12345), the new private IP address for the server inside the tunnel (10.10.10.1) and the subdomain to tunnel to (iodine.dnsjedi.org). This can be seen in Figure 4.1. The server side was now up and running and a tunnel could be made from the client side.

Figure 4.1: Iodined active and running on the server

Next, the client side was set up. This consisted of installing iodine and starting up the software. To start up, the same password and subdomain for the server had to be given. This can be seen in Figure 4.2. Because no resolver was given, iodine automatically uses the local resolver (130.89.0.128), and the client is given a private IP inside the tunnel (10.10.10.2).