A study on the use of 3rd party DNS resolvers for malware filtering or censorship circumvention

Andersen, Martin Fejrskov; Vasilomanolakis, Emmanouil; Pedersen, Jens Myrup

Published in:
ICT Systems Security and Privacy Protection

DOI (link to publication from Publisher):
10.1007/978-3-031-06975-8_7

Publication date:
2022

Document Version
Accepted author manuscript, peer reviewed version

A study on the use of 3rd party DNS resolvers for malware filtering or censorship circumvention *

Martin Fejrskov[1], Emmanouil Vasilomanolakis[2], and Jens Myrup Pedersen[2]

[1] Telenor A/S, Aalborg, Denmark
mfea@telenor.dk (corresponding)
[2] Cyber Security Group, Aalborg University, Copenhagen, Denmark
{emv,jens}@es.aau.dk

Abstract. DNS resolvers perform the essential role of translating domain names into IP addresses. The default DNS resolver offered by an Internet Service Provider (ISP) can be undesirable for a number of reasons such as censorship, lack of malware filtering options and low service quality. In this paper, we propose a novel method for estimating the amount of DNS traffic directed at non-ISP resolvers by using DNS and NetFlow data from an ISP. This method is extended to also estimate the amount of DNS traffic towards resolvers that offer malware filtering or parental control functionality. Finally, we propose a novel method for estimating the amount of DNS traffic at non-ISP resolvers that would have been censored by ISP resolvers. The results of applying these methods on an ISP dataset shows to which extent 3rd party resolvers are chosen by users for either malware filtering or censorship circumvention purposes.

Keywords: DNS · NetFlow · resolver · ISP · filtering · censorship.

1 Introduction

The DNS resolver service has traditionally been provided to customers by Internet Service Providers (ISPs). Recently, providers of public DNS resolver services, such as Google and Cloudflare, have gained popularity, and are estimated by Radu et al. to handle more than 50% of all DNS resolutions globally [17]. Although Radu et al. discuss the possible reasons users can have for choosing public DNS services, the authors remain at speculations on this topic.

Some equipment vendors (e.g. webcams) use 3rd party DNS resolvers as a default setting in products. Three main reasons for a user to *actively* choose a 3rd party DNS resolver are presented by web pages containing security advice:

 Funded by Telenor A/S and Innovation Fund Denmark.

- Service quality: Speed, reliability, and basic security features such as DNS-over-TLS (DoT), DNS-over-HTTPs (DoH) and DNSSEC validation.
- Privacy: Adherence to more strict privacy principles and no modification of the responses, for example to inject ads in NXDOMAIN responses [23].
- Filtering/censoring: The 3rd party provider does not follow government orders to censor responses. Conversely, the 3rd party provider may offer filtering of domains related to malware, porn, drugs, etc. as an add-on service.

As ISPs can deploy resolvers topologically closer to the end users than any 3rd party resolver, an ISP will always be able to offer a faster resolver service than any 3rd party resolver. As a fast DNS resolution can make an Internet connection appear faster, this represents a competitive advantage to an ISP. A competitive ISP can therefore be assumed to offer DNS resolvers with good service quality (although examples of ISPs not having this focus do exist [1]). European Union legislation forbids ISPs to collect personal information, and forbids ISPs to modify DNS responses for ad injection. Therefore a rational customer at a competitive, European ISP should not be inclined to use service quality or privacy as the main reason for choosing a 3rd party DNS resolver.

Following the arguments presented above, and assuming a rational customer and a competitive, European ISP, only the third category, filtering/censoring, is relevant, which will therefore be the focus of this paper. We recognize that there can be a difference between perceived privacy and actual privacy, as well as a difference between perceived and actual service quality, however we consider this topic out of scope of our paper. The contribution of the paper is the methods and measurements needed to answer the following research questions:

- RQ1: To which extent are 3rd party resolvers used compared to the default ISP resolvers?
- RQ2: To which extent are 3rd party resolvers that offer malware filtering or parental control used?
- RQ3: To which extent are 3rd party resolvers used to circumvent censorship?

These methods and associated results can be relevant for ISPs to assess the business case for offering DNS based filtering services. The results can also be relevant to regulatory bodies to assess the effect of DNS based censorship.

Section 2 introduces related work and other background information. The three following sections (3, 4 and 5) each answer one of the research questions outlined above. Section 6 summarizes the answers and concludes the paper.

2 Background and related work

2.1 Data availability

The simplest way to examine how much and which DNS traffic is directed at 3rd party resolvers is to ask the operators of those services. The privacy policies of the five major public DNS resolver providers (according to Radu et al.) reveal that the providers store data that could answer the question in either

anonymized or non-anonymized form, however, they are generally not willing to share the data [4,5,11,16,25]. Another approach is to collect data by interacting with user equipment. One example is the use of apps as probes by the Open Observatory of Network Interference (OONI) project. A second example is the use of advertisement campaigns (or similar mechanisms) that trigger a resolver to query observer-controlled authoritative servers [3]. These approaches can measure which resolvers are used relative to other resolvers, but do not quantify the amount of traffic from each client towards each resolver, which is the purpose of our paper.

Although ISPs are not legally allowed to inspect the DNS traffic to 3rd party resolvers, Fejrskov et al. describe that DNS data from the ISPs own resolvers as well as sampled NetFlow data (that includes 3rd party resolver traffic) can be used in anonymized form even when considering European Union legislation [9]. In our paper the ISP approach is adopted, and data from Telenor Denmark, a national ISP in Europe with 1,5M mobile and 100k broadband subscriptions, is used. Their DNS resolvers adhere to the service quality and privacy criteria mentioned in the introduction, and provide no add-on block offerings.

2.2 Estimating DNS traffic based on NetFlow data

Konopa et al. suggest a method to detect DoH traffic based on NetFlow records [14]. However, the method relies on access to unsampled NetFlow records which is not available in our paper. Although some papers discuss using NetFlow to identify specific applications, we are not aware of any other papers that directly focus on estimating the amount of DNS traffic. An intermediate step is to use the NetFlow records to estimate the actual number of UDP or TCP flows, a technique often referred to as flow inversion. Several papers, most recently [2], estimate the flow size distribution using various sampling methods, different traffic models, and uses different information from the sampled packets, such as the presence of TCP SYN packets and sequence numbers. Duffield et al. describe and validate a simpler technique that estimates the actual amount of TCP flows as the multiplication of the sample rate and the observed number of flows for which the initial SYN packet was observed [7]. Neither paper present any methods that are applicable to this paper for estimating the amount of UDP flows.

2.3 DNS Response manipulation

Several studies characterize the use of response manipulation in resolvers [13, 15, 24], including both filtering, censoring, injection, etc. Most papers consider response manipulation as an undesired feature as opposed to something positive that the user has actively chosen to gain features such as malware protection. In all papers, the characterization of servers is based on whether or not the server actually performs response manipulation, independently of whether it is advertised or not. In our paper, we therefore find it interesting to characterize

resolvers based on whether they advertise themselves as filtering or not, in order to investigate to which extent such functionality is desirable by users.

2.4 Censorship and circumvention detection

The legislation in Denmark requires ISPs to perform DNS based blocking of certain domains in 7 different categories [21]. In our paper, all categories are included with no distinction between them, giving a total of approximately 800 domains that have a DNS A record. The legislation (and following public discussion) is about blocking web pages, and DNS is seen as the tool that can implement this [6].

Related work on censorship fall in four categories: Techniques for implementing censorship, detecting censorship, circumventing censorship, and measure circumvention attempts. Only the last category is relevant to this paper, and this seems to be the topic of only a few papers. Three of these focus on the use of specific tools or apps like TOR [19], an app for changing DNS resolver [10], and on the use of DNS servers owned by VPN providers [8]. Our focus is only on circumvention that involves the use of 3rd party resolvers, not on specific tools.

The Danish Rights Alliance, an organisation focusing on copyright and other conditions for content creators, measures the effect of DNS based blocking by analysing web site visits [22]. They concluded that the effect of blocking a specific site through DNS blocking reduces the number of visits to the specific site by up to 75% after 4-5 months. In our paper, it is not a requirement that the censored sites consent to embedding code in their web page that measure usage statistics, and the focus is not limited to copyright.

Callejo et al. conclude that 13% of the global DNS queries are resolved by 3rd party rather than by ISP-provided DNS resolvers [3]. They also conclude that the use of 3rd party providers is more frequent in countries with a high level of censorship (a poor rating by the Reporters Without Borders' (RWB) World Press Freedom Index). Their approach relies on serving ads through browsers, and for the reasons mentioned initially in this section, the approach is not applicable for our paper. However, they conclude that the use of 3rd party resolvers in countries rated as Good by RWB is around 7-11% of the total traffic, which is an interesting figure to compare to our results.

3 Prevalence of 3rd party resolvers

This section presents a method for estimating the number of DNS responses represented by a set of sampled NetFlow records towards 3rd party DNS resolvers. The method consists of three steps that are described in more detail in the following three subsections. The number of 3rd party DNS responses is compared to the number of responses served by Telenor Denmark's DNS resolvers to answer the first question (RQ1) posed in the introduction.

Four different DNS traffic types are considered in this section: DNS over UDP and TCP, DNS-over-TLS (DoT) and DNS-over-HTTPS (DoH). As DNS requests

can potentially be malformed, and as only requests that result in a response are relevant from a user perspective, this study will focus on the number of responses rather than the number of requests.

3.1 Identifying relevant Netflow records

The first step is to identify the NetFlow records that represent 3rd party DNS resolver traffic. In this paper, it is a precondition that the available NetFlow records represent a view of all flows crossing a well-defined network boundary. Users and the default DNS resolvers are defined to be on the internal of the network boundary, 3rd party resolvers and other servers are defined to be on the external side. The NetFlows are considered sampled with a rate of 1:Q.

For an external IP address to be considered a potential 3rd party DNS resolver, and to filter away irregular and irrelevant traffic such as that originating from DDoS attacks and port scanning, some TCP or UDP traffic must be observed on port 53 or 853 in both directions, that is both to and from the server. However, due to the use of sampled NetFlow, observing records that form a bidirectional flow is not required, as both directions of the same flow will rarely be sampled given a high sample rate. TCP and DoT records originating from the potential resolver IP must report a packet size of at least 54 bytes to ensure that the response is at least large enough to contain a valid IP, TCP and DNS header. Therefore, packets only containing, for example, a TCP Reset flag indicating that no service is available do not qualify. This packet size criterion is not necessary for UDP based flows, as a server with no UDP service will respond with an ICMP packet instead of a UDP packet.

TCP port 443 traffic towards the resolvers outlined above is considered DoH traffic. We recognize that operators could run both DoH and Web services on the same IP address, and therefore the amount of DoH traffic estimated using this method should be considered as an upper bound rather than an exact number.

Traffic towards authoritative servers also satisfies the aforementioned criteria for a potential resolver, and these flows must be disregarded. Any of the following criteria are used to identify authoritative server IPs:

- The server returns an error code when resolving a well-known domain name, but answers succesfully when resolving the domain name found in the server's reverse/pointer (PTR) record.
- The IP address of the server is identical to any IP address with which the default resolvers communicate.
- The PTR record of the server IP reveals that the server is a well-known authoritative server, such as the DNS root servers or the authoritative servers of major commercial DNS providers.

As a result of the selection process described above, N NetFlow records are considered to represent user-initiated traffic to/from 3rd party resolvers, and only these records are considered for further analysis.

3.2 Average number of flows per Netflow record

Having identified a number of NetFlow records that represent a number of observed flows towards 3rd party resolvers, the next step is to estimate the number of actual flows. This requires different approaches for TCP and UDP traffic.

As outlined in Section 2, the estimated number of actual TCP flows, $\hat{F}_{TCP}$, can be found by multiplying the NetFlow sample rate with the number of flows in which a SYN packet is observed, $\hat{F}_{TCP} = Q \cdot F_{SYN}$. The number of observed SYN flows, F_{SYN}, is determined by aggregating the observed response SYN records, N_{SYN}, by the 6-tuple of observed flow start time, source and destination IP address, source and destination port number and protocol. For a Q much larger than the expected number of packets in a TCP flow, it is only expected that each TCP flow is sampled once, and in that case $\hat{F}_{SYN} = N_{SYN}$, which is demonstrated as a valid practice in Section 3.5.

To estimate the number of actual UDP flows, we use the property that a DNS request or response is always contained within a single UDP packet, and the property that a new UDP flow is made for each request due to the prevalence of source port randomization [12]. In other words, one UDP NetFlow record represents one flow and one DNS response. Therefore, the estimated number of UDP flows, $\hat{F}_{UDP}$, is given by the number of observed UDP response records multiplied by the NetFlow sample rate, $\hat{F}_{UDP} = Q \cdot N_{UDP}$. Note that although a response is always contained within a single UDP packet, this packet may be split into several IP packets due to fragmentation. In this case, only the first IP packet will contain UDP headers, and therefore only the first packet will be considered a UDP packet by the NetFlow emitting router. Therefore, the assumption of a one-to-one relation between DNS responses and UDP packets should be considered valid when using NetFlow as measurement method.

3.3 Average number of DNS responses per flow

Having estimated the number of actual TCP/UDP flows represented by NetFlow records, the next step is to identify the number of DNS responses per flow. For this purpose, it is assumed that the average number of responses per TCP flow for 3rd party resolvers and for the default resolvers are similar, that the average number of responses per DoT flow for 3rd party resolvers and for the default resolvers are similar, and that these numbers can be calculated from the collected data from the default resolvers. Different collection methods will allow for different methods for calculating the numbers, and the method described below reflects an approach applicable to our data set.

To estimate the average number of responses per TCP/DoT/DoH session, DNS response data from the default resolvers that include the ports of the response is used. The minimum time between flow closure and the allowed reuse of the related source port from the same request source IP address is denoted $t_{graceperiod}$. The longest allowed time for a TCP session to be open is denoted $t_{maxsessionlength}$, and therefore should be true that $t_{maxsessionlength} >$

$t_{graceperiod}$. A response, c is considered belonging to the same flow as another response b, if the two responses are less than $t_{graceperiod}$ apart ($t_b + t_{graceperiod} > t_c$), and if the response c and the first response in the flow, a, are less than $t_{maxsessionlength}$ apart ($t_a + t_{maxsessionlength} > t_c$).

It should be noted that the specific values of both $t_{maxsessionlength}$ and $t_{graceperiod}$ can differ among clients and servers, as such settings can be either operating system, application or deployment specific. The choice of values for these will therefore depend on the specific DNS server software settings.

Using this method to estimate which DNS responses belong to the same flow makes it possible to calculate an estimated, average number of responses per TCP flow, $\hat{R}_{TCP}$, and an estimated, average number of responses per DoT flow, $\hat{R}_{DoT}$. Notice that the similar number for UDP flows, $\hat{R}_{UDP}$, is always 1 for the reasons outlined in Section 3.2.

3.4 Method summary

The number of DNS responses from 3rd party DNS resolvers, $\hat{D}$, is estimated using NetFlow records as

$$\begin{aligned}\hat{D} =& \hat{D}_{UDP} + \hat{D}_{TCP} + \hat{D}_{DoT} + \hat{D}_{DoH} \\ =& \hat{F}_{UDP} \cdot \hat{R}_{UDP} + \hat{F}_{TCP} \cdot \hat{R}_{TCP} + \hat{F}_{DoT} \cdot \hat{R}_{DoT} + \hat{F}_{DoH} \cdot \hat{R}_{DoH} \\ =& Q(N_{UDP} + N_{TCP,SYN} \cdot \hat{R}_{TCP} + N_{DoT,SYN} \cdot \hat{R}_{DoT} + N_{DoH,SYN} \cdot \hat{R}_{DoH})\end{aligned}$$

for a large NetFlow sample rate Q, the number of relevant UDP NetFlow records, N_{UDP}, the number of relevant NetFlow records observing a SYN packet, $N_{TCP,SYN}$, $N_{DoT,SYN}$ and $N_{DoH,SYN}$, and the estimated, average number of DNS responses per TCP/DoT/DoH flow, $\hat{R}_{TCP}$, $\hat{R}_{DoT}$ and $\hat{R}_{DoH}$.

3.5 Measurements and discussion

Anonymized DNS and NetFlow data collected over a period of 4 days (covering both weekdays and weekend) from 2021-08-08 to 2021-08-11 from Telenor Denmark's network is used to demonstrate the use of the estimation method elaborated in the previous section. The DNS data is derived from the response packets for all DNS queries towards the default DNS resolvers. The NetFlow data is derived from traffic passing the BGP AS border with sample rate $Q = 512$. Metrics are summarized in Table 1. Although the data set only contains 4 days of data, we consider it to be representative, as DNS services are used on a daily basis, and as the amount of users is large (1,6M). The internal IP addresses in the data are anonymized by truncation to a /24 prefix, and the AM/PM information of the timestamps is truncated as suggested by Fejrskov et al. [9].

The NetFlow sample rate, Q=512, is higher than the expected number of packets in a DNS TCP flow. Therefore the number of observed flows is almost identical to the number of NetFlow records ($\hat{F}_{TCP,SYN} \approx N_{TCP,SYN}$ and $\hat{F}_{DoT,SYN} \approx N_{DoT,SYN}$) as anticipated in Section 3.2.

Table 1. *Metrics for 3rd party DNS resolver traffic estimation.*

Metric	**Symbol**	**Count**
Total NetFlow records	n	$2,75 \cdot 10^9$
Relevant NetFlow records	N	$3,32 \cdot 10^6$
NetFlow UDP records	N_{UDP}	$2,85 \cdot 10^6$
NetFlow TCP SYN records	$N_{TCP,SYN}$	$98,9 \cdot 10^3$
NetFlow TCP SYN flow	$\hat{F}_{TCP,SYN}$	$98,5 \cdot 10^3$
NetFlow DoT SYN records	$N_{DoT,SYN}$	$12,6 \cdot 10^3$
NetFlow DoT SYN flow	$\hat{F}_{DoT,SYN}$	$12,6 \cdot 10^3$
NetFlow DoH SYN records	$N_{DoH,SYN}$	$15,9 \cdot 10^3$
NetFlow DoH SYN flow	$\hat{F}_{DoH,SYN}$	$15,9 \cdot 10^3$
Max TCP session length	$t_{maxsessionlength}$	100 s
TCP source port grace period	$t_{graceperiod}$	30 s
DNS responses per TCP flow	$\hat{R}_{TCP}$	1,19
DNS responses per DoT flow	$\hat{R}_{DoT}$	11,3

232 NetFlow records relating to UDP traffic on port 853 were observed. This could represent DNS-over-DTLS (DNSoD) traffic [18]. Due to the small amount and the experimental status of the DNSoD standard, we disregard these records.

Moreover, $43,2 \cdot 10^3$ NetFlow records relating to UDP traffic (from port 53) report more than one packet per flow, which seems to contradict the assumption of one UDP packet per flow made in Section 3.2. Although an experimental IETF RFC from 2016 [20] describes the use of multiple UDP packets for responses, it seems unlikely that this should be implemented in several 3rd party resolvers. We therefore believe that a more plausible explanation is that this is caused by re-transmission of requests and responses. As re-transmissions are of no interest to this paper, a UDP NetFlow record (from port 53) reporting more than one packet will only be counted as one packet, and therefore as one request or response.

The value of $t_{graceperiod}$=30 seconds is chosen to match the default tcp-idle-timeout value of the Bind software running on the default DNS resolvers. The value of $t_{maxsessionlength}$=100 seconds is chosen arbitrarily to a value larger than $t_{graceperiod}$. Experiments show that choosing a significantly higher value, $t_{maxsessionlength}$=1000 seconds, does not change the estimated average number of requests per flow significantly.

Table 2. *Number of responses observed on the default resolvers and estimated from 3rd party resolvers. Notice that the DoH number should be considered an upper bound.*

	UDP	**TCP**	**DoT**	**DoH**	**Sum**
Default	$15,2 \cdot 10^9$	$10,9 \cdot 10^6$	$446 \cdot 10^6$	0	
	87,67%	0,06%	2,57%	0%	90,31%
3rd party	$1,46 \cdot 10^9$	$60,3 \cdot 10^6$	$73,2 \cdot 10^6$	$92,3 \cdot 10^6$	
	8,39%	0,35%	0,42%	0,53%	9,69%

The estimated 3rd party DNS resolver traffic is summarized in Table 2 in comparison to the amount of traffic at Telenor Denmark's default DNS resolvers. As Telenor Denmark's default DNS resolvers do not offer DoH service, the 3rd party DoH number is calculated by assuming that $\hat{R}_{DoH} = \hat{R}_{DoT}$.

Note that the estimated number of DNS responses from 3rd party resolvers listed in Table 2 also include responses for servers that could not be explicitly identified as either authoritative or resolving. This is applicable to approximately 0,79% of the listed responses from 3rd party resolvers.

Some customers use VPN services for connecting to their employer's VPN gateway or for keeping the traffic private. We consider it most likely that such traffic will use the 3rd party resolvers operated by the VPN gateway operator, that this operator is located outside Telenor Denmark's network, and that the DNS traffic is therefore not visible in the data set used for this study. Although a study of how widespread the use of VPN services is could be interesting, we consider it complementary to the scope of this paper.

The first question posed in the introduction (RQ1) asks to which extent the DNS traffic is directed at 3rd party resolvers. In Table 2 it can be seen that the fraction of the total DNS traffic that is directed at 3rd party resolvers is estimated to be between 9,69-0,79=8,90% and 9,69%. These results are in line with the 7-11% measured by Callejo et al. [3].

4 Prevalence of filtering 3rd party resolvers

The second research question (RQ2) asks to which extent 3rd party resolvers that offer desirable filtering services (such as malware filtering or parental control features) are used. In this section, the data presented in Section 3.5 is further enriched by adding information about which organisation runs the resolver, whether the resolver is public or private, and whether or not the resolvers are advertised by the owners as filtering.

4.1 Method

To identify if a 3rd party resolver is private or public, two methods are used:

- The resolver is queried with a popular domain name. If this query returns the correct result, the resolver is considered public. If no response is received, the server is considered private.
- If the owner of the resolver is known to only run private resolvers, the resolver is marked as private. These include the resolvers of other ISPs, some VPN services, as well as commercial DNS resolver companies known for only providing private services.

To identify the owner of a resolver, simple methods such as resolving the PTR record of the server, performing a Google or Whois search, are used.The owner's web page is then used to determine if the resolver offers filtering functionality.

Some DNS resolvers exist with the purpose of enabling the user to circumvent some restrictions put in place by web site owners, such as enabling the user to view TV shows that are only broadcasted in some countries due to copyright restrictions. Some, but not all, of these resolvers are associated with VPN services. For the purpose of this paper, we consider these as non-filtering resolvers, as actively choosing these resolvers is conceptually more similar to trying to avoid censorship, than to desire additional filtering.

Another category of resolvers are those that are associated with DNS hijacking malware that changes the DNS resolver settings on a device to point to a resolver under control of a malicious party. This resolver will then most likely manipulate the DNS response to achieve the purpose of the malicious actor. For the purpose of this paper, we consider these resolvers non-filtering, as they are unlikely to perform any kind of filtering that is considered desirable by the user.

4.2 Measurements and discussion

The result of identifying server owner, advertised filtering features and private/public category is summarized in Table 3. Unknown filtering status represents that we were not able to identify the owner/operator of the resolver. Unknown public/private status is typically caused by the server sending back a wrong answer or an error, such as REFUSED, NXDOMAIN or SERVFAIL.

Table 3. *Categorization of 3rd party DNS responses.*

	Public	**Private**	**Unknown**	**Sum**
Filtering	$202 \cdot 10^6$	$6,41 \cdot 10^6$	$101 \cdot 10^3$	
	12,02%	0,38%	0,01%	12,41%
Non-filtering	$1{,}37 \cdot 10^9$	$53,5 \cdot 10^6$	$204 \cdot 10^3$	
	81,11%	3,18%	0,01%	84,30%
Unknown	$16,0 \cdot 10^6$	$26,4 \cdot 10^6$	$12,9 \cdot 10^6$	
	0,95%	1,57%	0,77%	3,29%

A key finding is that between 12,41% and 12,41+3,29 =15,70% of traffic for 3rd party resolvers is for filtering resolvers. This suggests that malware filtering, etc., is not likely to be the primary motivation for using 3rd party resolvers.

In Section 3.5, it was concluded that the amount of 3rd party resolver responses is between 8,90% and 9,69% of all responses. In other words, the total fraction of responses that originate from filtering DNS resolvers is between $8,90\% \cdot 12,41\% = 1,10\%$ and $9,69\% \cdot 15,70\% = 1,52\%$, which answers the second research question. This shows that the use of filtering resolvers is not prevalent among Telenor Denmark's customers.

5 Censorship avoidance detection

The third question posed in the introduction (RQ3) asks if 3rd party resolvers are used to circumvent censorship. It is a prerequisite that the ISP's default DNS servers censor some domains based on national legal requirements, and that these are not censored by 3rd party resolvers. This section presents a method that uses ISP data to estimate how many DNS responses for censored domains are sent by 3rd party resolvers, and the results obtained by applying the method.

5.1 Method

As elaborated in Section 2, the censorship focuses on web domain names, and in contrast to the two previous sections that considered flows related to DNS servers, this section focuses on flows related to web servers only.

The core idea of the estimation method is to categorize the web flows seen in NetFlow records, use this categorization to estimate the fraction of the web flows that are towards censored sites, and then use this number of web flows to estimate the number of related DNS queries at 3rd party resolvers for censored domains. The categorization of flows is illustrated in Table 4 and elaborated in the following paragraphs. The lowercase w_1 to w_{12} represent the count of the flows within each category, and the uppercase W_1 to W_{12} represents the sets of flows within each category.

Table 4. *Categorization of the set of all web flows, W.*

				Default	**3rd par.**	**None**
W	**T**ainted	**S**hared	Censored dom.	$W_1 = \emptyset$	W_5	$W_9 = \emptyset$
			Non-censored dom.	W_2	W_6	W_{10}
		Non-**S**hared	Censored dom.	$W_3 = \emptyset$	W_7	$W_{11} = \emptyset$
	Non-**T**ainted		(Non-censored dom.)	W_4	W_8	W_{12}

The (uncensored) A records of all the censored domains contain a number of IP addresses, which will be referred to as tainted IP addresses. Some of the tainted IP addresses are assigned to servers that serve both censored and non-censored domains, and these addresses will be referred to as shared IP addresses. Flows relating to these servers are in categories W_1,W_2,W_5,W_6,W_9 and W_{10}). Conversely, some servers with tainted IP addresses only serve censored domains (no non-censored domains), and the IP addresses of these servers are referred to as non-shared IPs. Flows relating to these servers are in categories W_3,W_7 and W_{11}. Finally, the web flows that do not relate to any server IP found in the A record of any censored domain are referred to as non-tainted (categories W_4,W_8 and W_{18}). Some web flows are created following a DNS lookup at the default resolver (categories W_1 to W_4 in Table 4), some web flows are created following a DNS lookup at a 3rd party resolver (W_5 to W_8), and some web flows are created without any preceding DNS lookup (W_9 to W_{12}).

As queries for censored domains towards the default DNS server result in a censored response, such queries will not cause a subsequent flow to be created to the web server, therefore by definition $W_1 = \emptyset$ and $W_3 = \emptyset$. As the censoring is based on domain names only, we find it reasonable to assume that flows towards censored sites must be preceded by a DNS lookup, therefore in addition $W_9 = \emptyset$ and $W_{11} = \emptyset$. The number of flows towards censored sites created after a DNS lookup to a 3rd party resolver would be $w_5 + w_7$, and this is the interesting number to estimate.

By definition all web servers are located on the outside of the NetFlow boundary, and all clients on the inside of the NetFlow boundary. The set of relevant flows, W, is found using two criteria: First, only records relating to server TCP/UDP port 80 or 443 are considered. Second, only servers for which traffic both from and to the server is observed are considered, although the to/from traffic can relate to different flows to mitigate the effects of NetFlow sampling, following the same arguments as for DNS flows in Section 3.1. Flows are thereafter defined by aggregating NetFlow records by 5-tuple on a daily basis, and timestamped with the earliest timestamp on that day.

To estimate $w_5 + w_7$, the following steps are needed. Please refer to Table 4 for an overview of the different flow categories. An initial step is to identify the set of tainted and the set of shared IP addresses:

- T_{ip}: Let T_{ip}, the set of tainted IPs, be the set of DNS A record IPs returned by doing a DNS lookup towards a non-censoring DNS resolver of all the censored domains.
- S_{ip}: Let R_{ip} denote the set of IP addresses found in the Rdata field of A records of all responses from the default resolvers. As this because of the censoring will not include any non-shared IPs, R_{ip} thus contains all the non-tainted and all the shared IP addresses. The set of shared IP addresses, S_{ip}, can then be found as the subset of the tainted addresses, T_{ip}, that are also found in R_{ip}, $S_{ip} = T_{ip} \ltimes R_{ip}$.

These two IP address sets are then used split the full set of web flows W into sets of tainted, non-tainted, shared and non-shared flows corresponding to the four main categories (T, NT, S, NS) in Table 4:

- T and NT: Split the full set of flows, W, into the set of tainted flows $T = W_1 \cup W_2 \cup W_3 \cup W_5 \cup W_6 \cup W_7 \cup W_9 \cup W_{10} \cup W_{11}$ and the set of non-tainted flows $NT = W_4 \cup W_8 \cup W_{12}$. These can be determined based on whether or not one of the flow IP addresses can be found in T_{ip} such that $T = W \ltimes T_{ip}, NT = W \triangleright T_{ip}$.
- S: Find the set of shared flows, $S = w_1 \cup w_2 \cup w_5 \cup w_6 \cup w_9 \cup w_{10}$. This can be found using T as a tainted flow address is shared, if the server IP can be found in the default DNS responses, $S = T \ltimes S_{ip}$.
- $NS = W_7$: Find the number of non-shared (and by definition, censored) flows preceded by a 3rd party DNS lookup, W_7, by finding the total number of non-shared flows, NS, and exploiting that that $W_3 = \emptyset$ and $W_{11} = \emptyset$. $NS = W_7$ can be found using T as a tainted flow address is non-shared, if the server IP can not be found in the default DNS responses, $W_7 = NS = T \triangleright S_{ip} = T - S$.

The set of shared flows, S, consists of two subsets of flows, related to censored domains, $W5$, and non-censored domains, $W_2 \cup W_6 \cup W_{10}$. The next steps of the method focus on identifying which flows belong to which of these two subsets by various means. For this purpose, the concept of flow renaming will be used several times to determine which web flows are associated with which DNS responses. In our paper, flows and DNS responses are considered associated, if a flow is created no longer than θ minutes after the DNS lookup, if the client IP addresses match, and if the server IP of the flow is the IP found in the Rdata record of the DNS response. The effect of DNS caching at the user is assumed to be mitigated by the aggregation of flow records to the earliest timestamp during a specific day as mentioned above.

- W_2: Find the set of tainted, shared, non-censored flows preceded by a DNS lookup at the default servers, W_2. As $W_1 = \emptyset$ and $W_3 = \emptyset$ this can be found by renaming the flows of S by using all entries in the DNS response log, D, such that $W_2 = S \ltimes_\theta D$. The same method can in theory be applied to the set of non-tainted flows, NT, to find the untainted set W_4. However, the amount of data can be large, and the following steps therefore do not depend on the feasibility in practice of using renaming to distinguish between W_4 and $W_8 \cup W_{12}$.
- $w_6 + w_{10}$: The fraction of re-nameable flows within the non-tainted flow set and within the non-censored flow set is assumed to be the same, as none of these flows are censored. Therefore, $\frac{w_6+w_{10}}{w_2} = \frac{w_8+w_{12}}{w_4}$, where $w_6 + w_{10}$ is then easily found as w_2 is already known. Although W_4, W_8 and W_{12} cannot be identified (as elaborated above), the ratio $\frac{w_8+w_{12}}{w_4}$ can be found by renaming a sampled set of non-tainted flows, $\frac{w_8+w_{12}}{w_4} = \frac{w_{8_s}+w_{12_s}}{w_{4_s}} = \frac{nt_s - w_{4_s}}{w_{4_s}}$ where $sample(NT) = NT_s = W_{4_s} \cup W_{8_s} \cup W_{12_s}$, $W_{4_S} = NT_s \ltimes_\theta D$.
- w_5: Find the number of shared, censored flows preceded by a 3rd party DNS lookup, w_5, by subtraction: $w_5 = s - (w_2 + w_6 + w_{10})$

These steps provide the necessary values to calculate $w_5 + w_7$ which is the estimated number of flows towards censored sites that are associated with a DNS lookup to a 3rd party resolver.

Flow renaming is performed in the steps for finding W_2 and $w_6 + w_{10}$, and we consider this mechanism to be the largest cause of uncertainty to the result. The method as used in this paper is greedy in the sense that too many flows will be considered re-nameable and therefore as non-censored, both because flows and DNS responses are considered related based on a time interval (larger time interval is more greedy), but also because user IP addresses are anonymized by truncation. Therefore, the estimated value of $w_5 + w_7$ should be considered as the lower boundary of the real value. As shown in a later subsection, the estimation of the lower boundary instead of the actual value turns out to be a sufficient metric to support our conclusions.

The next step is to calculate the number of estimated, actual DNS responses, $\hat{p}$, that relate to the estimated, observed, flows $w_5 + w_7$. The techniques described in Section 2 for estimating the actual number of flows based on the observed

number of flows are not applicable in this case, as they depend on the availability of NetFlow records and not just the availability of an estimated flow count. Instead, we propose to identify all servers for which only port 80/443 flows are observed, let w_{web} denote the number of flows towards these servers and let p_{web} denote the count of DNS responses with an A record containing the IP addresses of these servers. Then we will estimate the number of DNS responses related to the censored flows as $\hat{p} = \frac{p_{web}}{w_{web}}(w_5 + w_7)$. As the value of $w_5 + w_7$ is considered a lower boundary, the value of $\hat{p}$ should also be considered a lower boundary.

5.2 Measurements and discussion

The estimation method detailed above is applied to DNS and NetFlow data from Telenor Denmark's network collected over a period of 4 days from 2021-09-23 to 2021-09-26. The most interesting metrics are summarized in Table 5. 1:1000 of the non-tainted flows are used to estimate $w_6 + w_{10}$. Results for two different values, $\theta = 1min$ and $\theta = 60min$, of the time interval allowed in the renaming process are presented in order to illustrate the importance of this parameter as discussed above. A $\theta > 60min$ does not give significantly different results.

In summary, we estimate that at least $\hat{p} = 477 \cdot 10^3$ DNS responses for censored domains have been answered by 3rd party DNS resolvers. This number can be compared to the number of censored DNS responses served by the default resolvers, $44,6 \cdot 10^3$, and the ratio between these numbers is $r = 10,7$.

Table 5. *Metrics for censorship evasion estimation.*

Metric	**Symbol**	**Count**	
Relevant flows	w	$1,03 \cdot 10^9$	
Shared flows	s	$196 \cdot 10^3$	
Non-shared flows	$ns = w_7$	$7,40 \cdot 10^3$	
Ratio of responses and flows	$\frac{p_{web}}{w_{web}}$	$18,1$	
Censored responses at default DNS resolvers	$d_{censored}$	$44,6 \cdot 10^3$	
Renaming interval	θ	1 min.	60 min.
Shared, non-censored flows preceded by default lookup	w_2	$103 \cdot 10^3$	$166 \cdot 10^3$
Shared, non-cens. flows not preceded by def. lookup	$w_6 + w_{10}$	$28,0 \cdot 10^3$	$11,1 \cdot 10^3$
Shared, censored flows preceded by 3rd party lookup	w_5	$65,5 \cdot 10^3$	$19,0 \cdot 10^3$
Estimated DNS responses related to censored flows	$\hat{p}$	$1,32 \cdot 10^6$	$477 \cdot 10^3$
Ratio of censored responses at default and 3rd party	r	$29,6$	$10,7$

Section 3.5 concluded that approximately 9% of the total DNS traffic was from 3rd party resolvers. If 3rd party resolvers were not used to circumvent censorship, it would be expected that $r \approx 0,09$. Censored 3rd party resolver responses are therefore at least two orders of magnitude more prevalent than expected, which suggests that 3rd party DNS resolvers are chosen to circumvent censorship. It is more challenging to consider if censorship circumvention is the *primary* reason for a user to choose a 3rd party resolver. Hypothetically, even if

this was the only reason for choosing 3rd party resolvers, the number of censored domains would still only be a small fraction of the total responses, as individual users will then also use the 3rd party resolver for non-censored domains.

As the number of censored responses from 3rd party servers is only an estimated number, it is not possible to assess how many users resolve censored domains using this method either. Even if this was possible, it would not be meaningful to compare this number of users to the number of users receiving censored responses from the default resolvers, without knowing more about the intentions of these users. One may argue that all of the responses from the default servers are caused by unintentional web page visits that will not be repeated by a user, whereas all the responses from the 3rd party servers could be caused by deliberate web page visits that will most likely be repeated by the user.

Although the results in this paper are based on only a single dataset, we find that the methods are independent of the dataset, and that the temporal length of the dataset is sufficient to present valid results for Telenor Denmark. We fully recognize that using the dataset of another ISP in another country could yield different results, both for technical reasons (such as differences in default DNS resolver setup) and cultural reasons (desire to circumvent censorship etc.).

6 Conclusion

In this paper we propose a method for estimating the amount of TCP/UDP/DoT/DoH DNS responses by using information from NetFlow records. This method is applied to estimate how much of the DNS traffic in an ISP is from 3rd party resolvers instead of the ISP's default resolvers. Using data from Telenor Denmark it is concluded that 8,9-9,7% of the total DNS traffic is from 3rd party resolvers (RQ1). This result supports and is supported by the most recent related work that uses a completely different method for obtaining the results [3]. Also, it is concluded that 1,1-1,5% of the total DNS traffic is from filtering resolvers (RQ2). Although it is expected that some traffic is from filtering resolvers, the specific number is not quantified by any existing research that we are aware of. The low number suggests that filtering resolvers are not commonly used by Telenor's customers, and this could represent an unexploited business opportunity to promote the use of such services.

Furthermore, we propose a NetFlow based method for estimating the amount of DNS responses from 3rd party resolvers that would have been censored by the ISP's default DNS resolvers. Using data from Telenor Denmark, it is concluded that DNS responses for censored domains are at least two orders of magnitude more prevalent at 3rd party resolvers than at the ISP's default resolvers (RQ3). We are not aware of any related work quantifying this number on an ISP scale. The high number suggests that 3rd party resolvers are actively chosen in order to circumvent censorship, which should be considered when the censorship legislation is up for evaluation.

It is correct that we only rely on a single dataset, however, we believe that the methods are independent of the dataset, and that the single dataset used is

sufficiently large to present valid results for the specific ISP. We fully recognize that using the dataset of another ISP in another country could yield different results. This is, however, more likely attributed to cultural differences (knowledge about cyber security in the population, the desire/need to circumvent censorship in a particular country, etc.) rather than the merits of the presented method.

The focus of this paper is purely technical, however for future work it could be interesting to compare the obtained results with a user questionnaire asking for the user's primary motivation for actively choosing 3rd party servers.

Although the specific results presented in this paper applies only to Telenor Denmark's customers, the methods are general, and it is our hope that they will be used by other ISPs and organisations to identify both business opportunities and regulatory challenges.

References

1. Ager, B., Mühlbauer, W., Smaragdakis, G., Uhlig, S.: Comparing DNS Resolvers in the Wild. IMC: ACM SIGCOMM conference on Internet measurement (2010), http://dx.doi.org/10.1145/1879141.1879144
2. Antunes, N., Pipiras, V., Jacinto, G.: Regularized inversion of flow size distribution. INFOCOM: IEEE Conference on Computer Communications (2019), https://doi.org/10.1109/INFOCOM.2019.8737406
3. Callejo, P., Cuevas, R., Vallina-Rodriguez, N., Ángel Cuevas Rumin: Measuring the Global Recursive DNS Infrastructure: A View From the Edge. IEEE Access (2019), https://doi.org/10.1109/ACCESS.2019.2950325
4. Cisco: Cisco Umbrella Privacy data sheet (2021), https://trustportal.cisco.com/c/dam/r/ctp/docs/privacydatasheet/security/umbrella-privacy-data-sheet.pdf
5. Cloudflare: 1.1.1.1 Public DNS Resolver (2020), https://developers.cloudflare.com/1.1.1.1/privacy/public-dns-resolver
6. Danish Ministry of Justice: Lov om ændring af retsplejeloven og forskellige andre love (2017), https://www.retsinformation.dk/eli/ft/201612L00192
7. Duffield, N., Lund, C., Thorup, M.: Properties and prediction of flow statistics from sampled packet streams. IMW: ACM SIGCOMM Internet Measurement Workshop (2002), https://doi.org/10.1145/637201.637225
8. Farnan, O., Darer, A., Wright, J.: Analysing Censorship Circumvention with VPNs via DNS Cache Snooping. IEEE Security and Privacy Workshops (SPW) (2019), http://dx.doi.org/10.1109/SPW.2019.00046
9. Fejrskov, M., Pedersen, J.M., Vasilomanolakis, E.: Cyber-security research by ISPs: A NetFlow and DNS Anonymization Policy. International Conference on Cyber Security And Protection Of Digital Services (2020), https://doi.org/10.1109/CyberSecurity49315.2020.9138869
10. Florio, A.D., Verde, N.V., Villani, A., Vitali, D., Mancini, L.V.: Bypassing Censorship: a proven tool against the recent Internet censorship in Turkey. IEEE International Symposium on Software Reliability Engineering Workshops (2014), https://doi.org/10.1109/ISSREW.2014.93
11. Google: Your privacy (2021), https://developers.google.com/speed/public-dns/privacy

12. Hubert, A., van Mook, R.: RFC 5452: Measures for Making DNS More Resilient against Forged Answers (2009), https://datatracker.ietf.org/doc/html/rfc5452
13. Khormali, A., Park, J., Alasmary, H., Anwar, A., Mohaisen, D.: Domain Name System Security and Privacy: A Contemporary Survey. Computer Networks (2021), https://doi.org/10.1016/j.comnet.2020.107699
14. Konopa, M., Fesl, J., Jelínek, J., Feslová, M., Cehák, J., Janeček, J., Drdák, F.: Using Machine Learning for DNS over HTTPS Detection. European Conference on Cyber Warfare and Security (2020), http://dx.doi.org/10.34190/EWS.20.001
15. Pearce, P., Jones, B., Li, F., Ensafi, R., Feamster, N., Weaver, N., Paxson, V.: Global Measurement of DNS Manipulation. USENIX Security Symposium (2017), https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-pearce.pdf
16. Quad9: Data Privacy Policy (2021), https://www.quad9.net/privacy/policy/
17. Radu, R., Hausding, M.: Consolidation in the DNS resolver market – how much, how fast, how dangerous? Journal of Cyber Policy (2019), https://doi.org/10.1080/23738871.2020.1722191
18. Reddy.K, T., Wing, D., Patil, P.: RFC 8094: DNS over Datagram Transport Layer Security (DTLS) (2017), https://www.rfc-editor.org/rfc/rfc8094.html
19. Roberts, H., Zuckerman, E., York, J., Faris, R., Palfrey, J.: 2010 Circumvention Tool Usage Report. The Berkman Center for Internet & Society (2010), https://cyber.harvard.edu/sites/cyber.harvard.edu/files/2010_Circumvention_Tool_Usage_Report.pdf
20. Sivaraman, M., Kerr, S., Song, L.: DNS message fragments (2016), https://www.ietf.org/staging/draft-muks-dnsop-dns-message-fragments-00.txt
21. Telecom Industry Association Denmark: Blokeringer (2021), https://www.teleindu.dk/brancheholdninger/blokeringer-pa-nettet/
22. The Danish Rights Alliance: Report On Share With Care 2 (2020), https://rettighedsalliancen.dk/wp-content/uploads/2020/06/Report-On-Share-With-Care-2_Final.pdf
23. The ICANN Security and Stability Advisory Committee (SSAC): SAC 032 - Preliminary Report on DNS Response Modification (2008), https://www.icann.org/en/system/files/files/sac-032-en.pdf
24. Trevisan, M., Drago, I., Mellia, M., Munafò, M.M.: Automatic Detection of DNS Manipulations. IEEE International Conference on Big Data (2017), https://doi.org/10.1109/BigData.2017.8258415
25. Yandex: Terms of use of the Yandex.DNS service (2021), https://yandex.com/legal/dns_termsofuse/