Malicious DNS Traffic in Tor: Analysis and Countermeasures

Michael Sonntag
Institute of Networks and Security, Johannes Kepler University, Altenbergerstr 69, A-4040 Linz, Austria

Keywords: Anonymization, Tor, DNS, Malicious Behaviour.

Abstract: Anonymization is commonly seen as useful only for people that have something to hide. Tor exit nodes are therefore associated with malicious behaviour and especially the so-called "darknet". While the Tor network supports hidden services, and a large share of these serve illegal purposes, most of the traffic in the Tor network exits to the normal Internet and could be, and probably is, legal. We investigate this by taking a look at the DNS requests of a high-bandwidth exit node. We observe some malicious behaviour (especially DNS scans), questionable targets (both widely seen as immoral as well as very likely illegal in most countries), and careless usage. However, all these, while undoubtable undesirable, make up only a small share of the exit traffic. We then propose some additions to reduce the detected malicious use.

1 INTRODUCTION

It is commonly claimed that the Tor anonymisation network (Dingledine/Mathewson/Syverson, 2004) is used for undesirable/illegal activities - but so is the "normal" Internet. The Tor network routes traffic over three nodes with multiple layers of encryption to anonymize the IP address of the source. While it can be used for any kind of TCP connection, it is overwhelmingly used for web surfing. In this way, visitors of websites may remain anonymous to the sites (unless they log in) and avoid blocks to them by their ISP. This definitely has appeal for illegal activities - but so it has for content which is officially labelled as "undesirable", e.g. in countries with strong censorship.

Inspecting the Tor traffic was done e.g. by Ling/Luo/Wu/Yu/Fu (2015), which discovered a large amount of malicious traffic. However, only 9 % of their alerts were related to actual malware. As we operate a high-bandwidth exit node, we investigated its exit traffic for signs of such undesirable (according to several ways of classifying it as such) traffic. In this paper we report on the results from observing the DNS traffic of the exit node regarding malicious behaviour, as opposed to Sonntag (2018), where we investigated the use by country of destination and categorization of second-level domains. Investigating DNS traffic is especially useful, as it would allow blocking undesirable behaviour before expending bandwidth, which is usually in low supply for exit nodes of the Tor network. Additionally, the DNS traffic is public anyway to a large degree: what cannot be answered immediately from the cache is sent to some external DNS resolver and is observable from the outside, e.g. the ISP of the user and the operator of the DNS server. This could lead to additional problems for exit nodes, e.g. complaints or blocks based on scans exiting from it.

2 DATA COLLECTION METHOD

Data was collected during five month, from 1.2.2018 until 30.6.2018 in one-hour periods. The method for collection was to install our own DNS caching server and use this as the DNS server of the exit node. As that computer is not used for anything else, all DNS queries can be attributed to the exit node. The cache logs all queries to disk. The logs are rotated hourly and investigation takes place per hour to better preserve privacy. To ensure as detailed data on exit traffic as possible, the timeout this caching server returns to the exit node is set to a very low value of 1 minute. Note that this is not directly effective, as Tor itself sets the timeout to 5 minutes for very small timeouts it receives (and 60 minutes for longer ones) to protect against attacks (DefecTor: Greschbach/Pulls/Roberts/Winter/Feamster 2017). Because of this, we did not modify these settings.On the Internet side of the cache, no changes are made- whatever the upstream servers send is used.

Sonntag, M.
Malicious DNS Traffic in Tor: Analysis and Countermeasures.
DOI: 10.5220/0007471205360543
In *Proceedings of the 5th International Conference on Information Systems Security and Privacy (ICISSP 2019)*, pages 536-543
ISBN: 978-989-758-359-9

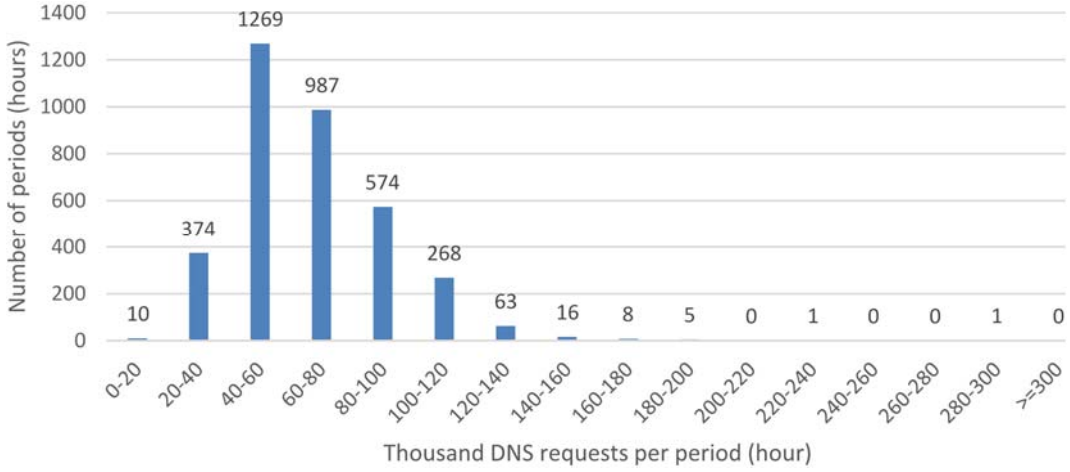

Figure 1: Histogram of DNS requests per hour.

3 SUSPICIOUS BEHAVIOUR

The simplest statistics is the number of DNS requests occurring per hour. For normal traffic this should correlate to the traffic, i.e. for each outgoing TCP connection one DNS lookup can be expected. Because of internal caching this number must be reduced significantly, as e.g. a web page does not consist of a single file only, but e.g. of several HTML pages, stylesheets, script files, multiple pictures etc. On average there were 66,542 requests, which translates to approx. 18.5 requests per second (see also at the end under ethical considerations). In sum there were 237,953,608 DNS requests during the whole observation period. However, this number varies significantly over time. The minimum number per hour encountered was 3,698, while the maximum was 291,472. To better understand these variations, a histogram of classes of counts was created (see Figure 1). From this it is apparent that the variations are much lower than it appears at first, as extreme outliers influence especially the maximum value. Regarding the number of connections, in total 2,429,411,680 connections were observed during the same period, which translates to 680,126 flows per hour. This produces 10.2 flows per logged DNS request (note the DNS caching; but some connections are established directly to IP addresses too). As most of the traffic is web surfing, this looks correct.

3.1 Reverse DNS Scanning

The single extreme outlier in Figure 1 was investigated individually: this hour had 221,621 .arpa PTR requests (i.e. about 70,000 other requests, which is perfectly average for a single period). Generally, very few reverse lookups are to be expected, as e.g. web traffic (taking up almost all of the traffic), does not need this at all. This was a reverse scan of several large networks (the names in the parentheses stem from the WhoIS database): 158.172.0.0/16 (ORGANISMO AUTONOMO DE CORREOS Y; this seems to be the Spanish postal service), 158.227.0.0/16 (Universidad del Pais Vasco), 158.42.0.0/16 (Universitat Politecnica de Valencia), 158.49.0.0/16 (Universidad de Extremadura). It was further investigated whether there exists an associated spike in traffic: we do not have any information on individual targets, but the whole traffic during this hour was not different from other hours at all, neither in number of connections nor the amount of data transferred. Therefore, this scan was not accompanied by actual connections to these IP addresses, it was “merely” a reverse DNS scan.

The reason for performing such a scan via Tor is not obvious: the targeted institutions would not note such a scan, unless they operate their nameservers themselves (or were specifically informed of it). As these are class B subnets, that is however likely the case - and was in this instance. The nameservers for the Spanish post (193.148.159.170, .171) are within a different network, but these addresses also belong to the post. For the universities, at least some of the nameservers lie within the address area scanned (158.227.82.16; 158.42.1.5; 158.49.8.2). We can therefore conclude that such large scans would likely have been noticed by the targets and potentially traced back. Performing them via Tor avoids that possibility as any trace back to the origin would stop at our exit node. As we did not discover any legitimate (or business) cause (e.g. checking for rogue

computers can be done from any IP address not affiliated to the institution), less than honourable intentions can be surmised, e.g. discovering which systems exist and gathering information on them.

3.2 DNS Scanning

Domains asked for but not existing are a significant portion of the queries: on average 6,577 domain names are asked for each hour, which do not exist. This translates to 10% of all requests. As it is unlikely that humans enter that many incorrect domain names and even notoriously non-specification-conforming HTML usually gets the host part of links right, a different explanation is required. After manual investigation of these errors, we could identify the following subgroups:

- These seems to be a lot of checking for existing (or not) domains going on with very good lists or sensible automation. Few nonsensical names are tried, as almost all do make (some) sense. For example, these contained (beside numerous similar others) the following series of queries: worldwiderevеal.com, worldwiderevenue.net, worldwidescort.com, worldwidescubatravel.com, worldwideshoponline.com, worldwideshopspot.com, worldwidetomatosociety.com, worldwidetowers.com, worldwidetowinginc.com, worldwidetravelmembership.com, worldwideunderstanding.com, world-wide-web-host.com, worldwidewebstersonline.com, worldwidewebtec.com, worldwideweed.xyz. However there seems to be no obvious generator being used, as definitely many more "worldwide*" names exist, multiple TLDs are used (but with different second-level names), and e.g. typos are perhaps also part of it (worldwidescort should probably be worldwideescort). Also, if being merely dictionary-based, many more combinations than the ones above would be tried. A possible explanation for this is that multiple exit nodes might be used, so we only saw a portion of all queries. Note that unlike the examples below, all these domain names were only queried for a single time over the whole observation period. This therefore seems unlikely to be a prelude to attacks, but more searching for opportunities to buy domain names, or creating respectively maintaining a list of existing top-/second-level domains.
- Numerous non-existing domain names are queried for multiple times: for example, the top one is "geo.mozilla.org" with 37,395 queries in total over all five months, a domain name that however did exist in the past. The next most common one (15,929 times) is cdn.api.twitter.com, which seems to have been a working (but non-official) server which has since been shut down. A small amount of queries are mistakes of websites, at least partly because of changing/removed server names not followed by changes in the websites.
- Some domain names are obviously simply erroneous, like "index.php" (3,778 queries) or "wp-login.php" (1,320 times), which are probably meant as a path and not as a host. Or "web.archive.orghttp" (2,433 queries), "web.archive.org.https" (occurred 19 times) or "web.archive.org.localhost" (4 queries), which are typos or signs of misconfigurations or mistakes. Even aggregated these do not constitute a significant number of queries in total.
- Not directly explainable are the huge amount of queries for domains of the form "forum.*". 714,174 such non-existing domain names were queried for. And as each of the top names ("forum.eurostimul.com", "forum.zawya.com", "forum.roots-archives.com" etc) occur more than 2,900 times this cannot be merely a scan. According to Google searches, these domain names do not exist or existed, although there might have been forums on these sites (e.g. "eurostimul.com/forum/memberlist.php" *is* in the result list). As it is unlikely that several thousand scans with the same lists occur, this is looking more like an error while performing scans.
- Apart from non-existing domain names, also many queries receive a "no-data" reply. The technical reason is when a specific type of DNS record is queried for and the domain name does exist, but not this kind of record. Because of the limitation of Tor in DNS queries (only A=IPv4, AAAA=IPv6, and PTR=reverse lookup; are possible), the explanation is simple: these are queries for IPv6 addresses, where only IPv4 data exists (or potentially the reverse). This can be exemplified by the most common name in this category: e13829.x.akamaiedge.net was queried for 1,111,778 times! This domain name does exist, but only serves IPv4, but was often queried for its non-existing IPv6 record. The same applies to the second largest count in this category: shops.myshopify.com (363,117 queries; IPv4 data only). These requests are therefore legitimate and not signs of a scan, but of the increasing share of IPv6 being used.

DNS scans can also be used as attacks: little outgoing traffic causes large return traffic. Together with falsifying the source address a DoS attack becomes possible. As the exit node determines the source address of query packets, this is not relevant here. But the fact remains, that a DNS server must produce a large answer (and expend computing time for producing it), thereby, although not allowing

Figure 2: Measurements against the hour of the day.

reflection attacks, potentially supporting DoS attacks against name servers. This would be especially prominent in reverse scans, as these are all going to the same nameserver(s) if a single/few TLDs are chosen.

3.3 Ad-/Malware Domains

What is surprising too is the number of domains queried, which are on a malware/adware blacklist (Black). This list is a compilation of several other lists with duplicates removed and contains slightly below 60,000 domain names. Variants with additional categories like fake news, gambling, porn etc exist, but these were not used as many of these extensions are legal in most countries.

Comparing all queries to this list results in 3,403 matches per hour, so about 5.1% of all requests are on this list (again: merely containing ca. 60,000 domain names!). However, there is a possible explanation for this: beside just obtaining anonymity people use Tor also for getting around restrictions, e.g. state or company censorship. Such measures are typically implemented on firewalls and use similar lists (for security purposes or to restrict non-business-related Internet use). So while “normal” websites can be visited directly, “forbidden” ones are more likely to be visited through Tor - and more likely to end up on such lists. Therefore the share of such websites would be larger.

Another element is, that despite its name, the list not only contains “bad” sites (malware/adware), but also many sites which are merely advertisements or user tracking (for example, 125 domains of the form *.oewabox.at are on the list; this is the “Austria Web analysis” used by most Austrian newspapers, online shops etc).

Therefore, this rather large share of domains found on the list is not solely a measure of illegal/dangerous activity, but still noteworthy. Additionally, these sites are “problematic” in the sense of posing dangers to visitors, so the “criminal behaviour” is at least often not on the party using Tor for visiting them, but on the website operators.

3.4 Result Validation via Time-of-Day

The total DNS traffic depends on the hour of the day, which is unsurprising as so does the total traffic. The maximum is at 23 o’clock local time (Austria), i.e. 22 UTC, while the minimum is between 5 and 7 UTC (see Figure 2 ; “Total”). We can compare this with the normal “European” traffic as evidenced by the throughput at DE-CIX (https://www.de-cix.net/de/locations/germany/frankfurt/statistics).
There the minimum is at 4 o’clock and the maximum at approximately 21 hours UTC. From this comparison it is evident, that our traffic is (on average) shifted 2 hours later. This would imply that the “average” user is slightly east of our location. If exit nodes are selected randomly and not deliberately, then a completely “flat” curve would have been expected, as humans are distributed across the whole world (except the oceans). Another factor to take

into account is, that people might use Tor differently than other web traffic, e.g. predominantly in the evening or preferably during work. No definite conclusion is possible, but either there are proportionately more users located in the western part of Russia and the middle east (or generally in Asia than America, which seems more likely), or users prefer Tor in the evening and shun in during the day.

As can be seen from Figure 2, the columns "NX" are practically independent of the time. This not only looks like this, the correlation between the total traffic and not-existing domains is merely 0.145. From this we can reinforce the discovery of scans going on - these are independent of actual end-user traffic and therefore do not rise/fall with it. Normal users are unlikely as being their source, as these would type wrong names in the same ratio all day. Independence is even further reinforced when comparing it to "ND", the replies that the domain exists, but no data is present. This *does* vary with the hour of the day and the correlation factor to traffic is 0.995, so IPv4/IPv6 issues are directly related to the traffic of users.

Regarding the lines in Figure 2 it is important to note, that these are individually scaled to be better visible and comparable, so their values are not according to the left axis. But what can be seen from them is, that the number of domains found in the Malware/Advertisement list is similar to the total traffic. This can be explained by the fact, that many sites use advertisements for commercialization. The correlation between those two is however not that strong (Malware/Ad vs Total traffic is 0,797). This leads to the conclusion, that "problematic" sites are visited in a larger share during the evening than during the day (see gaps/touching lines in Figure 2).

3.5 WhoIs Scans

Domain name queries were classified according to their third-level domain. Domain names may consist of up to 63 labels, and often the third from the right tells what service is being accessed (e.g. www.company.com → "www" → website). Today however many queries do not contains such a third level element any more at all (like in "google.com"; 114,475,510 such queries occurred in total).

What becomes apparent from these results is, that WWW traffic is by far the most prominent one, especially as the classes "Server" and "CDN" will in many cases be web elements too. But what is surprising is the large number (656,752) of "WhoIs" queries. This ties in with a previous finding showing significant such traffic based on ports accessed (Sonntag/Mayrhofer, 2017). One possible explanation is, that this is related to the reverse domain name requests: checking whether an IP address is associated to a domain name and then asking for its owner. However, verification would require detailed investigation of individual traffic content (which website was queried for in the WhoIs connection) and correlation with domain queries and was therefore not performed. Whether this kind of traffic will continue in the future is unclear, as e.g. according to the EU GDPR much less data will be contained in the WhoIs databases, and even less immediately publicly accessible, so queries might be of less use.

3.6 Dangerous Usage

Also noteworthy are the smaller but still significant counts of queries regarding mail servers (mail./smtp./imap./…): 262,220 queries. Although traffic with many of them will be encrypted, this is not guaranteed. Also note that we do not allow port 25 (=SMTP) on our exit node, so this must be mail retrieval, not sending. Even more surprising and potentially dangerous are queries regarding FTP servers (13,004). While small on comparison, this is still a very large absolute number, where the transmission of credentials would take place unencrypted. These could be "secure" in the sense that only anonymous logins to public servers are used, but whether this is the case cannot be determined without inspecting the actual traffic.

3.7 Illegal Content

What people are looking for via Tor has been investigated via categorization of the domain names requested. Categorization was performed through "Shalla's Blacklists" (Shalla's Blacklists). These lists provide categorizations or URLs and is with a count of 1.7 million entries quite comprehensive. This list contains both domains and URLs. While we could easily extract the domains from the URLs, this would be problematic, as e.g. the download link for the microsoft.com website (classified as "Downloads") does not mean that the whole of microsoft.com is purely a download site. Unfortunately we were able to categorize only 10 % of all traffic (89,99% is not in the classification list). But for the 10% found the results are as follows (only categories with at least 1% are listed individually):

Table 1: Successful categorization of DN queries.

Category	DN requests	Share
Porn	3,400,700	14.3%
Socialnet	3,001,751	12.6%
Shopping	2,765,896	11.6%
Adv	2,074,556	8.71%
News	2,063,338	8.55%
Forums	1,504,763	6.32%
Movies	1,385,006	5.82%
Tracker	1,339,224	5.62%
Searchengine	1,264,996	5.31%
Imagehosting	797,450	3.35%
Downloads	637,854	2.68%
ISP	520,920	2.19%
Chat	355,395	1.49%
Government	352,259	1.48%
Webmail	239,451	1.01%
Other		8.87%

While this list does not directly show "problematic" or "illegal" traffic, it clearly shows that many visits are likely legal: shopping, social networks, news, forums etc are predominantly legal, as are advertisements. Potentially problematic content is porn (depending on kind and country this can be illegal), forums/chat (depends on topic) and movies/imagehosting/downloads (a significant share of information about files violating copyright is to be expected - less so the files themselves because of the limited bandwidth).

Also interesting is the large share of webmail: using Tor to access a mail account does not guarantee anonymity for the E-Mail address at all, this requires different anonymization methods. Tor brings here only one advantage: the association between the user of a (free - anonymous paying is complicated) account and an E-Mail address remains hidden. So it seems there is a significant desire for not only using an "anonymous" E-Mail address, but also ensuring that this E-Mail address cannot be traced back to the computer accessing it. But see also above for directly accessing E-Mail servers in section 3.6.

Potentially "problematic" categories are comparatively rare: downloads (2.68% of queries that could be categorized), spyware (0.82%), warez (0.81%), gamble (0.49%), anonvpn (0.09%; i.e. another anonymisation layer on top of Tor!), hacking (0,08%), drugs (0.07%). While not common, these are still a relevant amount, e.g. "drugs" refers to 15,924 of 238 million queries (=0,0067% of all queries, so one in 14,946). No numbers for the "normal" internet could be found, but this tiny part looks not very extraordinary and is definitely not a major share of the total Tor usage.

4 POSSIBLE COUNTERMEASURES AGAINST MALICIOUS USE

What can be done against such attacks? We are discussing here only measures to be implemented on exit nodes. Educating users, securing their browsers etc are out of our scope. Similarly, existing countermeasures, like removing the WhoIs Port from the exit policy to prevent such connections completely (countering section 3.5), are not discussed.

4.1 DNS Queries without Traffic

DNS scans (sections 3.1 and 3.2) are either trivial to detect or very hard. If a single Tor circuit issues numerous DNS queries but does not open any connection to them, then this is technically easy to detect. This would merely require defining a "minimum" of actual content traffic per DNS request, as it should be very uncommon to ask for a specific domain name and then not even try to send any data to it. So a limit of 2-5 requests without data traffic (=RELAY_RESOLVE as opposed to RELAY_BEGIN; see src/feature/relay/dns.c of the Tor source code) could be easily enforced. This comes with a potential problem however: state storage. The exit node would have to store this additional information for each Tor circuit until a data connection is at least tried, potentially allowing DoS attacks against the exit node.

This approach would not completely prevent DNS scans, but at least render them much more difficult to perform as a new Tor circuit would have to be established every few requests, creating a significant slow-down. This would work even better for reverse scans (PTR queries), as these are so uncommon in normal traffic that any even slightly increased use is very likely a misuse.

A potential problem, however, could be web browser prefetching: requesting a DNS lookup for domains of links on the current page, which the user might click on later to reduce latency and browsing speed (see Nidd/Kunz/Arik, 2000). But see above: an average of 10 flows per DNS request point rather in the opposite direction.

Still, a permanent prevention of scans is impossible. This would require either to correlate multiple Tor circuits (all going to the same subnet or "similar" domain names - technically difficult and requiring lots of resources) or identifying that they are originating from the same system - something the Tor system is specifically designed to prevent.

4.2 Delaying Responses

A softer approach would be to artificially introduce delays. The first query of a Tor circuit is answered immediately, but each further query without data traffic is delayed by an additional e.g. one second (3rd query: 2 seconds and so on) before the response is sent back. In this way scans would be similarly discouraged, but the countermeasure would be harder to detect (which is less useful than it sounds, as this fact would very soon become public knowledge, both generally and specific to exit nodes).

4.3 Blacklist Filtering

Filtering with blacklists is another countermeasure that would be possible to reduce illegal usage, especially as discussed in sections 3.6 and 3.7. However, the problem is to define what is illegal. The exit node can of course ban what is not allowed at its location, but this need not be identical to illegality where the end-user is. Additionally, blacklists are notoriously problematic regarding their maintenance: adding new sites to block and removing old ones with changed content. There exists another issue here: blocking can only in some cases be performed based on DNS, as e.g. a site might contain legal as well as illegal content under different URLs. Differentiating them would only be possible by investigating the content of the exit traffic and no longer by DNS queries alone. As now most exit traffic is encrypted, this is impossible anyway. Blocking based on lists should therefore (and as well based on general considerations about censoring, too) be avoided.

5 ETHICAL CONSIDERATIONS

What we are investigating here is Tor exit node traffic, i.e. intended to be anonymous. The most important priority of research is therefore to keep it like this. A DNS name, other than the full URL, usually does not tell anything about the user visiting this site by itself. However, that is not guaranteed, like websites about specific medical problems. Together with the exact time of the DN query it could potentially be useful to deanonymize specific users through correlation attacks. To avoid any reduction in anonymity, even though the exit node alone will not help without the other two nodes, the recorded data is stored and evaluated in one-hour chunks. The exact time of the requests, resp. replies, is removed immediately after evaluation (and not used anyway, but cannot be avoided in the DNS cache's log).

We observed a minimum of 3,698 DNS requests per hour, resulting in approx. one DNS query per second. The average over all one-hour periods are 18 requests/second, with a maximum of 81 queries. The timestamp precision is typically one second, therefore the lower boundary is close to supporting individual identification.

Note that DNS information is not confidential: iterative DNS requests are typically sent to the next server in full, not merely the necessary subpart (see QNAME minimisation for privacy improvements: RFC 7816; Bortzmeyer 2016). Therefore, third parties may observe parts or all of the information anyway, as it is not encrypted at all (DNSSec is not widely used and would have to be added to Tor exit nodes via a proxy anyway). As the IP addresses of exit nodes are publicly known, if they perform name resolution themselves, this is obvious. In our case the same, solely dedicated to Tor services, network is employed, so queries can still be identified as related to the exit node.

6 CONCLUSIONS

While we detected malicious behaviour in the DNS traffic, it is on a very low level. Specifically, DNS-only behaviour is scanning, both forward (asking for IP addresses of multiple domain names) and reverse (asking for the domain name of many IP addresses). For both we have identified potential countermeasures, where the most promising seems to be limiting such queries per Tor circuit and/or delaying them. While this would not prevent such scans, it would make them more costly (continuously creating new Tor circuits) or more suspicious (actually initiating a connection to these hosts). Drawbacks from such measures are not apparent but should be tested. In this way malicious behaviour through Tor could be reduced to some degree.

ACKNOWLEDGEMENTS

We would like to thank both the Johannes Kepler University Linz as well as the AcoNet for supporting this project by granting permission respectively providing the necessary bandwidth.

REFERENCES

Black, S.: Unified hosts file with base extensions. https://github.com/StevenBlack/hosts.

Bortzmeyer, S.: DNS Query Name Minimisation to Improve Privacy. RFC 7816, 2016. https://tools.ietf.org/html/rfc7816.

Dingledine, R., Mathewson, N, Syverson, P.: Tor: The Second-Generation Onion Router, In: *Proceedings of the 13th conference on USENIX Security Symposium - Volume 13 (SSYM'04)*, Vol. 13. USENIX Association, Berkeley (2004).

Greschbach, B., Pulls, T., Roberts, L. M., Winter, P., Feamster, N.: The Effect of DNS on Tor's Anonymity. *NDSS '17*, Internet Society, San Diego (2017).

Ling, Z., Luo, J., Wu, K., Yu, W., Fu, X.: TorWard: Discovery, Blocking, and Traceback of Malicious Traffic Over Tor, *IEEE Transactions on Information Forensics and Security*, Vol 10/12, 2515 - 2530 (2015).

Nidd, M., Kunz, T., Arik, E.: Prefetching DNS Lookup for Efficient Wireless WWW Browsing. *Proceedings of Wireless 97*, 409-414.

Shalla's Blacklists. http://www.shallalist.de/

Sonntag, M., Mayrhofer, R.: Traffic Statistics of a High-Bandwidth Tor Exit Node. In: Mori, P., Furnell, S., Camp, O. (eds.) *Proceedings of 3rd International Conference on Information Systems Security and Privacy*, 270-277. SCITEPRESS (2017).

Sonntag, M.: DNS Traffic of a Tor Exit Node - An Analysis. In: Wang G., Chen J., Yang L. (Eds): Security, Privacy, and Anonymity in Computation, Communication, and Storage. SpaCCS 2018. Springer, Lecture Notes in Computer Science, vol 11342, 33-45 (2018).