An analysis of TOR

Team 2B

Sara Sverdrup-Thygeson - sarasver
Julia Vister - juliavi
Jannina Veluppillai - janninav

Department of informatics
The faculty of mathematics and neutral science

UNIVERSITY OF OSLO
2024

Introduction

In this paper we have done an analysis of the Tor blocking issues that occur with the use of the network. We will first explain in Part1 what Tor is and how it works, and then look at the challenges that Tor faces. In Part 2, we will present our findings from the analysis and compare them with the conclusions from three articles; *A bestiary of blocking: The motivations and modes behind website unavailability, How the Great Firewall of China is Blocking Tor, characterizing the nature and dynamics of Tor exit blocking*. These articles provide an overview of general TOR blocking issues, and we will show how our findings align with or differ from their insights.

Part I

What is TOR and how does it work?

In the 1990s, the lack of security in the internet and its ability to be used for tracking became more and more significant. David Goldschlag, Mike Reed, and Paul Syversen at the U.S. Naval Research Lab's (NRL) answer to this was to create and deploy the first research design and prototypes of onion routing(Torproject, 2024). The goal of onion routing is to use the internet with as much privacy as possible, the idea was to route traffic through multiple servers and encrypt for each step(Torproject, 2024). TOR is an onion routing network with over 2M daily users and over 7000 supporting servers (Singh et al., 2017, p. 326).

The network is designed to provide its users with anonymity by routing internet traffic through a series of volunteer-operated servers. Each relay in the Tor network peels away one layer of encryption, making it nearly impossible to trace the original source of the traffic. Anonymity systems like Tor provide a useful service to users who wish to access the Internet without revealing their intended destination to any local monitoring, or their network-layer identity to the destination.

Challenges

The article *A bestiary of blocking: The motivations and modes behind website unavailability* address different issues regarding Tor and blocking mechanisms. These issues involve

country-based restrictions, security-motivated blocks, and more recently, GDPR-related blocks. Each form of blocking has unique motivations, and understanding them is critical to analyzing the extent of online censorship (Tschantz, M. C 2018, s.2).

Country Based Blocking

Country-based blocking is one of the most recurring issues. It restricts access to content conditional to the users geographical location. Even though the article emphasizes this method is a technical approach rather than a specific motivation. It's clear that different countries have varying points of internet censorship(Tschantz, M. C 2018, s.4). An example of this is the difficulty to access uncencored content in China through TOR, due to the strict firewall policies in the country.

China's Great Firewall (CGF) blocks access to the Tor anonymity network. However, the Chinese government has developed methods to block it. CGF targets the entry points to Tor, known as entry guards and bridge relays. The firewall dynamically detects Tor traffic by searching for specific patterns of data that identify a connection as Tor-related. Once identified, the GFC initiates a scan using arbitrary Chinese computers that try to "speak Tor" to the suspected bridge. If this interaction succeeds, the firewall blocks the bridge, preventing users in China from accessing the Tor network and bypassing censorship. (Winter & Lindskog, 2012).

GDPR based blocking

GDPR-motivated blocking is the second type of blocking. This arises when a website blocks incoming network traffic from the EU, in order to avoid compliance with the European Union's General Data Protection Regulation (GDPR). This occurs more often at websites based in the U.S, as they have lower requirements for data protection than the GDPR and find it easier to block all EU-based traffic rather than reconfigure their operations to obey(Tschantz, M. C 2018, s.3). Thus, browsing through Tor with an EU exit node, many users might encounter sites that are inaccessible due to GDPR compliance concerns(Tschantz, M. C 2018, s.6).

Security motivated blocks

One of the most noticeable and significant challenges that Tor faces is security-motivated due to its strong association with anonymity and perceived security risks. While Tor aims to provide users with privacy, many websites implement protective measures to guard against malicious traffic often linked to the network.

A major issue stems from the abusive behavior of some users, which leads to discrimination from online service providers. These providers argue that Tor's anonymity is frequently exploited for activities like spamming, vulnerability scanning, and data scraping(Singh et al., 2017, p. 326). As a result, Tor users are often subjected to restrictive practices, such as solving CAPTCHAs or being entirely blocked from accessing certain websites, which often leads to users encountering "403 Forbidden" or "451 Unavailable for Legal Reasons" (Tschantz, M. C 2018, s.5) .
The discrimination arises because all Tor users share the same set of exit relay IP addresses, making it difficult for service providers to distinguish between malicious actors and legitimate users. When one user engages in harmful behavior, the shared IP address can be blacklisted, penalizing benign users as well. This collective treatment adds to the ongoing friction between Tor's mission of privacy and the efforts of websites to safeguard their platforms.

For example, if you attempt to access certain U.S. news websites from an EU exit node, it's blocked due to GDPR restrictions. Security-related blocking is also recurrent. While visiting e-commerce sites and even forums, the sites were regularly met with CAPTCHA challenges, which thought the traffic was wary due to the use of TOR. Even though these challenges are manageable, it does add friction to the browsing experience. Country-based blocking can be less of an issue for certain individuals, but it's noticeable when accessing different streaming services due to its regional licensing agreements.

Facts

Many websites and services proactively block Tor traffic, leading to difficulties accessing websites and services for legitimate users. Around 88% of Tor exit relays are blacklisted, compared to the endpoints of VPNGate and HMA VPN services that is at 9 and 69 %, and

discrimination is common on websites, with about 20% of Alexa's Top 500 websites blocking or restricting Tor traffic. This number is further increased by search and login functionalities with 3.9% and 7.5 % percent(Singh et al., 2017, p. 326).

Some of the websites practiced proactive blacklisting and others reactive, where proactive was blacklisting based upon the networks pre existing reputation or online service policy and reactive is the blacklisting in response to the abuse. On the research provided out of the 84 blacklists, 6 were proactive and 78 were reactive(Singh et al., 2017, p. 326).

Part II

Methodology

This report analyzes website accessibility using the Tor browser. Over a week, we used Tor as the main browser, logging the date, website name, link, and whether access was granted. If access was denied, we noted the specific error.

We then examined the data for patterns in access restrictions across websites and industries, identifying trends. The report also highlights issues like slow load times, access denial, and other disruptions to browsing.

Our findings and experiences

To identify potential patterns, we segmented the data by types and sectors to determine if any trends emerged within these categories. The industries were e-commerce, banking, digital media, public sector, aviation, hotels, and streaming services. We found that e-commerce, aviation websites and streaming services were most prone to blocking, where 11 out of 20 e-commerce websites were blocked and half of the aviation as well. In many instances we experienced "access denied" errors, indicating that many online stores either block anonymous traffic or are more sensitive to Tor's slow speed. *LOT.com*, *Converse.com*, *HM.com*, *Lyko* and are some examples that were blocked, with access denied errors.

While browsing we also experienced irregular behavior and issues with poor performance due to latency issues. For example, the streaming site *hdtoday.tv* exhibited functionality

problems, where the site barely worked and the files did not load. The fact that this is a pirated streaming site could be the reason for this. Other websites, like *Bigordi.com*, suffered from extremely slow performance. *Chat GPT* was just a blank, static website with an input field that did not work.

Impact of testing location

The data collection was done at 4 different physical locations and the data collected was more or less split equally between the faculty and each of the homes of the testers in order to see if that would have an impact on the data collected. However, the success rate of the websites did not show a clear pattern based on the testing location. Both environments experienced issues, suggesting that the problems encounterer, whether related to site blocking or performance, are largely independent of the location.

Conclusion

Tor's work to offer an anonymous browsing experience faces significant challenges due to the occurrence of several blocking methods. Tor users experience restricted access in several ways, from country-based restrictions to GDPR compliance and security-motivated blocks. These blocks vary depending on the exit node and the region users are browsing from, making a fragmented and sometimes frustrating internet experience. While Tor aims to provide users with privacy, many websites aims to protect themself against malicious traffic, creating a challenge to balance both needs.

Bibliography

Singh, R., Nithyanand, R., Stony Brook, Sadia Afroz, Paul Pearce, Michael Carl Tschantz, Phillipa Gill, & Vern Paxson. (2017, August 16). *Characterizing the Nature and Dynamics of Tor Exit Blocking*. Usenix.

Winter , P., & Lindskog, S. (2012). *Karlstad University.* Retracted from: How the Great Firewall of China is Blocking Tor: https://www.cs.kau.se/philwint/static/gfc/

Tschantz, M. C., Afroz, S., Sajid, S., Qazi, S. A., Javed, M., & Paxson, V. (2018). *A bestiary of blocking: The motivations and modes behind website unavailability*. In 8th {USENIX} Workshop on Free and Open Communications on the Internet ({FOCI} 18). https://www.usenix.org/system/files/conference/foci18/foci18-paper-tschantz.pdf

Tor project
https://www.torproject.org/

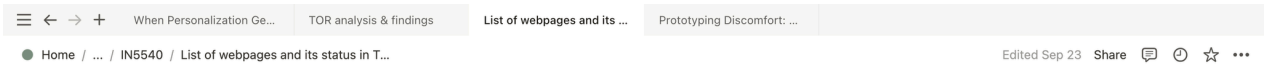
Appendix

Datacollection in excel

Date	Name of website	Link	Managed to acess	ErrorCode in case of no
19-Sep	VG	https://www.vg.no/	Yes	
19-Sep	How to geek	https://www.howtogeek.com/272	Yes	
19-Sep	.onion Wikipedia	https://en.wikipedia.org/wiki/.oni	Yes	
19-Sep	Tor blog	https://blog.torproject.org/breaki	Yes	
19-Sep	IFI	https://www.mn.uio.no/ifi/	Yes	
19-Sep	IFI	https://www.mn.uio.no/ifi/om/hm	Yes	
19-Sep	IFI	https://www.mn.uio.no/ifi/om/hm	Yes	
19-Sep	IFI	https://www-int.mn.uio.no/ifi/live	Yes	
19-Sep	IFI	https://www.titan.uio.no/andre-te	Yes	
20-Sep	Lufthansa	https://www.lufthansa.com/us/er	No	Not permission based on ser
20-Sep	Ryan Air	https://www.ryanair.com/nl/nl	Yes	
20-Sep	SAS	https://www.flysas.com/en/book/	No	Thinks I am a bot
20-Sep	Norwegian	https://www.norwegian.no/	Yes	
20-Sep	Booking.com	https://www.booking.com/search	Yes	
20-Sep	AirBnb	https://www.airbnb.com/	No	Not permission based on ser
20-Sep	Hotels.com	https://hotels.com/	No	Not permission based on ser
20-Sep	DNB	https://www.dnb.no	No	Not permission based on ser
20-Sep	Sparebanken	https://www.sparebank1.no/nb/s	Yes	
20-Sep	Nordea	https://www.nordea.com/en	Yes	
20-Sep	Skatteteaten	https://www.skatteetaten.no/pers	Yes	
20-Sep	Altinn	https://info.altinn.no/	Yes	
22-Sep	Harvard business review	https://hbr.org/1990/03/the-man	Yes	
22-Sep	Chat GPT	https://chatgpt.com/c/66f02355-	No	Fikk ikke lastet inn
22-Sep	Chat GPT	https://chatgpt.com/	No	Hvit
22-Sep	Harvard buinsess publishing educati	https://hbsp.harvard.edu/student	Yes	
22-Sep	Accenture	https://www.accenture.com/us-e	Yes	
22-Sep	Politiet	https://www.politiet.no/	Yes	
23-Sep	Skatteteaten	https://www.skatteetaten.no/pers	Yes	

20-Sep	Altinn	https://info.altinn.no/	Yes	
22-Sep	Harvard business review	https://hbr.org/1990/03/the-man	Yes	
22-Sep	Chat GPT	https://chatgpt.com/c/66f02355-	No	Fikk ikke lastet inn
22-Sep	Chat GPT	https://chatgpt.com/	No	Hvit
22-Sep	Harvard buinsess publishing educati	https://hbsp.harvard.edu/student	Yes	
22-Sep	Accenture	https://www.accenture.com/us-e	Yes	
22-Sep	Politiet	https://www.politiet.no/	Yes	
23-Sep	Skatteteaten	https://www.skatteetaten.no/pers	Yes	
23-Sep	The Guardian	https://www.theguardian.com/eu	Yes	
23-Sep	The Guardian	https://www.guardian2zotagl6tmj	Yes . Onion	
23-Sep	Finacial Times	https://www.ft.com/	Yes	
23-Sep	Forsvaret	https://www.forsvaret.no/	Yes	
23-Sep	Fredrik og Louisa	https://fredrikoglouisa.no/	Yes	
23-Sep	Kicks	https://www.kicks.no/	Yes	
23-Sep	Lyko	https://lyko.com/	No	403 Forbidden
23-Sep	Tannlege	https://colosseumtannlege.no/	Yes	
23-Sep	Hm	https://www.hm.com/	No	don't know

Navn	Lenke	Type nettsted	Annet	Funket nettsiden?	Errors/Anomalities
hdtoday	https://hdtoday.tv/	Streaming	på skolen		Site barely works, the images do not load. The same with the media. Lots of pop up windows.
UiO Canvas	https://uio.instructure.com/	Teaching platform	skolen		
Goodreads	https://www.goodreads.com/genres/fantasy	Book catalog site	hjemme		
Ark	https://www.ark.no/search?spr%C3%A5k=Engelsk	Nettbutikk	hjemme		
Adlibris	https://www.adlibris.com/	Nettbutikk	hjemme		Got "Connection timed out" the first time i went to the site.
Bigordi	https://bigordi.pl/	Nettbutikk	hjemme		Extremely slow
LOT	https://www.lot.com/us/en/	fly	på skolen		Access denied
Converse	https://www.converse.com/	Nettbutikk	på skolen		Access denied



List of webpages and its status in TOR

Table

Untitled

Aa WebPages	Status	Error Code
VG	Full Access	
E24	Full Access	
NRK	Full Access	
Netflix	Full Access	
Nordea	Full Access	
Sbanken	Full Access	
UiO	Full Access	
Skatteetaten	Full Access	
Zalando	Full Access	
Pent	Full Access	
Microsoft	Full Access	
Finn	Full Access	
NRK	Blocked	
DNB	Blocked	Access Denied You don't have permission to access

When Personalization Ge... | TOR analysis & findings | List of webpages and its ... | Prototyping Discomfort: ...

Home / ... / IN5540 / List of webpages and its status in T...

Edited Sep 23 Share

Aa WebPages	Status	Error Code
NRK	Blocked	
DNB	Blocked	**Access Denied** You don't have permission to access "http://www.dnb.no/en" on this server. Reference #18.3f677b5c.1727072375.9abdf143 https://errors.edgesuite.net/18.3f677b5c.1727072375.9abdf143
Canvas	Blocked	**403 Forbidden**
Disney +	Blocked	Sorry, an unexpected error has occurred.Please try again later.
HM	Blocked	**IT'S NOT YOU, IT'S US!** *For some technical reason the page cannot be viewed. Please contact customer service.*
Weekday	Blocked	**Access Denied** You don't have permission to access "http://www.weekday.com/en-ww/" on this server. Reference #18.86792617.1727073966.26d1f9cc

When Personalization Ge... | TOR analysis & findings | List of webpages and its ... | Prototyping Discomfort: ...

Home / ... / IN5540 / List of webpages and its status in T...

Edited Sep 23 Share

Aa WebPages	Status	Error Code
		https://errors.edgesuite.net/18.86792617.1727073966.26d1f9cc
ChatGPT	Partly blocked	You get access to the website, but unable to use its functions
Figma	Partly blocked	
Pintrest	Partly blocked	

+ New

Calculate