Measuring Accessibility of Popular Websites when using ProtonVPN

Willemijn Tutuarima[1] , **Stefanie Roos**[1]
[1]TU Delft

Abstract

Censorship and privacy issues have led people to use VPNs when accessing the internet. These VPNs not only try to protect their user but they are also associated with criminality and cyber attacks. Because of this, websites have started to resort to blacklisting the IP addresses that are used by the VPNs, thus blocking both genuine and malicious users. This forces users to sacrifice privacy for accessibility. This paper provides a method on how to measure the amount of blocking that VPN users experience and to be able to determine what type of blocking is occuring. This method is then used in an experiment using a web crawler where nodes from ProtonVPN are used to measure the amount of blocking that occurs while browsing the internet's most popular websites. This experiment shows that on average 1.12% of the domains perform some type of blocking directed towards the VPN user and that the majority of this blocking consists of a total block, which means that the user is entirely excluded from any use of the website. Next to this it is shown that not all VPN nodes show the same amount of blocking and that there was no large difference in blocking found between days while using the same VPN node. It also shows that the categories which perform the most blocking are Business, Online Shopping and News.

1 Introduction

Many countries restrict the amount of information citizens can access through the internet. Censorship of certain websites is becoming common all over the world and causes many people to be restricted from media and information sources. This censorship can be caused by different parties in the internet network and can have different motivations behind the censorship. As Aase et al. [1] describes, the motivations behind the censorship can differ because of laws, government policy and the effort of different parties to minimize abuse of networks. Next to understanding the motivations behind limiting accessibility of internet users, it is also important to understand the amount of censorship that occurs and how this censorship occurs. For this purpose, many applications have been developed to be able to create an overview of where and how this censorship is distributed over the world.

Tools have been created to target specific countries such as China [2], Pakistan [3], and India [4] that show that accessibility of websites is limited in these countries. A survey done by Aceto et al. [5] shows that many different tools have been made focusing on different types of censoring techniques and detection techniques. These tools also show that censorship is a problem for many people and that it drives the disadvantaged users to steer to censorship circumvention tools to still be able to have access to the information they require. There are different techniques that can help to circumvent censorship such as using a DNS resolver in a different country and ignoring spoofed TCP RSP packets as described by Verkamp et al. [6] or other packet-based evasion strategies such as used in the Geneva tool [7]. Still when using this sort of techniques, a lot of personal data can be gathered from the user such as usage patterns that are used for targeted advertisements [8].

To prevent this break of privacy and anonymity, people can use an anonymity network to reduce the exposure of their data such as Tor [9], I2P or VPNs [10] [11]. The problem with these networks is that they, as pointed out by Polyakov [12], are starting to be associated with activities involving criminality such as cybercrime attacks [12] (e.g. DDoS attacks) but also real-world crime such as distribution of drugs as described by Nihal [13]. Because of this, websites have started to resort to using blacklists of IP addresses associated with these types of activities. Since the IP addresses of VPN services are shared and thus being used by both malicious and genuine user, genuine users wrongfully lose access to these websites [14]. Since this can be viewed as enforcing censorship on VPN users, it is important to know the amount of blocking this causes for users of these networks. Since VPNs are increasingly being used throughout the world with millions of users, a lot of people can be affected by blocking of the IP addresses used by these services.

In this paper, the amount of blocking of websites is measured while using the commercial VPN ProtonVPN basic. ProtonVPN basic is used because it is a large commercial VPN that has multiple IP servers in the Netherlands, the country where the control IP address is located, which is the non-VPN IP address used to compare with the VPN IP address.

There are several questions that should be answered to determine to what extend blocking occurs. Firstly, we ask how blocking would be defined and what types of blocking can occur. Next to this, a method is developed to be able to detect these types of blocking and with this method an experiment is performed to determine what types of blocking occur and in what frequency.

In the paper, we show that the main methods of blocking are block pages or captchas which contribute to an average of 1.12% of blocked domains. Within the experiment, it is shown that different VPN nodes show different percentages of blocking and that for different days using the same VPN node there was no considerable difference found. Next to this, the most common categories for blocked websites are business domains, online shopping and news pages. For which most of the blocking is done through a block page or an empty page.

1.1 Related work

As mentioned before, a lot of different tools have been developed to measure internet censorship across the world. These have mainly focused on measuring general censorship imposed by governments and internet service providers (ISP) contrary to blocking imposed by websites themselves, on which this paper is focused. They all focus on different types of attack methods against internet users such as TCP/IP blocking, DNS manipulation, HTTP(S) blocking and content manipulation. Tools such as Censored Planet [15], UBICA [16] and Encore[17] are used for large scale measurements that focus on blocking caused by ISPs and government entities. These types of blocking do not involve the websites but are enforced by external entities that do not have a direct role in the data exchange. This paper focuses on detecting website-based blocking where the ISP and country of origin remain constant and the behaviour of these parties is equal for a VPN request and a non-VPN request. This way the website's behaviour can be monitored and the blocking it performs against VPN users. Others such as Filtered Web [18], Concept Doppler [19] and Quack [20] also focus on keyword filtering, which means that websites are filtered by certain keywords found in HTML files. This is also mainly done by ISP and thus does not give information on website-based blocking.

Next to this, there has also been some research done that discusses IP blocking from blacklists such as Censmon [21] and C-saw [22]. These provide good methods for measuring different types of blocking including block pages but they both give VPNs as a solution to this kind of blocking instead of being the cause. Some tools such as ICLab [23] do actually use privacy networks in their measurement. These networks are used to make sure that the data gathering is done in a secure and anonymous way and no actual users of the internet are affected by participating in the measurements[3]. But although these do use VPNs, they do not measure the effect of the actual VPN but only of the location of the vantage points. Thus, all of these tools do not measure the blocking of anonymity networks by websites. Research done by Singh et al. [24] does show some measurements of blocking for the privacy network Tor where it measures the amount of IP blocking done by websites for Tor exit nodes. Next to this also Khattak et al. [25] shows that users of Tor experience blocking and content manipulation through CAPTCHA and block pages. Both of these give a good view of the blocking occurring when using Tor but they do not provide any measurements for using a VPN. As can be seen from this previous work, there has been a lot of research done on measuring blocking on the internet from different perspectives and using different techniques but there is still a lack of research on the direct effect of using a VPN on the accessibility of websites.

2 Methodology

To measure the extent and frequency of blocking that occurs when using ProtonVPN basic, data is gathered about internet

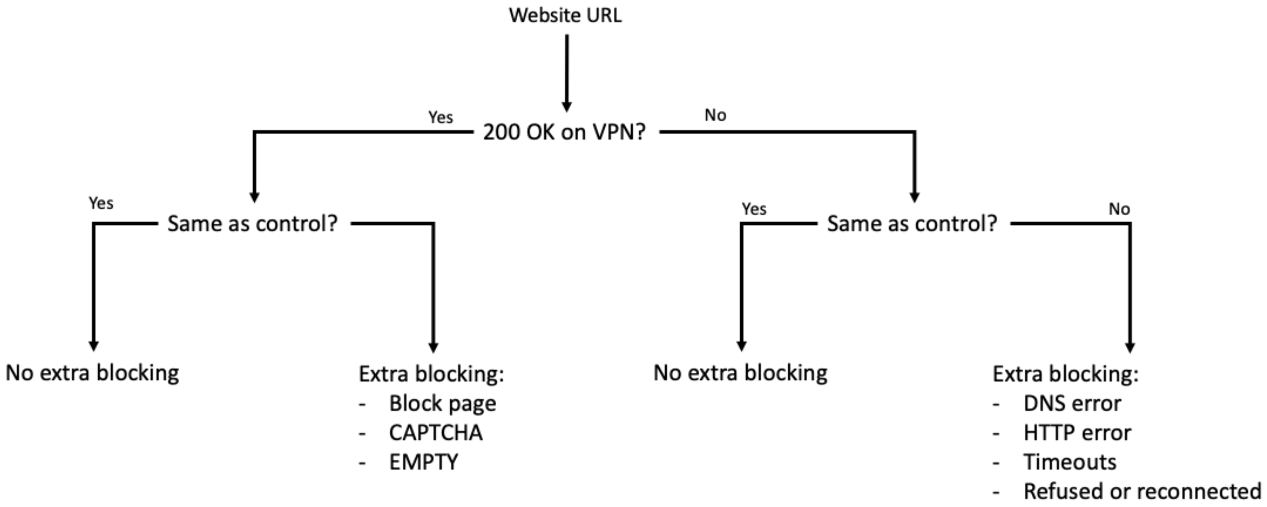


Figure 1: Process of categorizing HTTP responses

traffic when using the VPN and is compared to the internet traffic when using a control connection that is not using a VPN. To do this, traffic information for popular websites is categorized based on comparison between responses of the VPN connection and a non-VPN connection. This method was chosen since it gives the opportunity to isolate the effect of only the VPN when all other variables (e.g. local network, browser or operating system) remain constant. Next to this it provides a control webpage that can be used to indicate block pages.

To gather data about the internet traffic a web crawler is used that takes as input a list of popular websites. For every website on this list, it performs a HTTP request. For some pages also a maximum of 3 sub pages is requested and the response for these are recorded. Both the connection status is recorded and if a good connection is made, a screenshot of the web page is taken. Another method that could be used is comparing DOM files. The problem with this is that a block can cause only a small change to the DOM file and thus will not be detected. Next to this the screenshot gives the best representation of the web page since this will also be what the user of a VPN will be seeing.

Every response of the VPN connection will first be categorized as an error or as a successful connection. If an error occurred and the HTTP status code is not 200OK, the error is compared to the response of the control connection. If the responses differ, then the error is categorized in one of the following categories:

- DNS lookup errors: caused by the used DNS server not being able to locate the requested URL or the website server either not responding or responding too slow. When using the VPN, the DNS request is handled by the VPN.
- 3xx or 4xx HTTP status code: this indicates that the connection failed and that either a problem occurred on client side or on server side.
- Timeouts: either TCP timeouts or TCP connection loss , the limit set is 30 seconds
- Refused or reconnected: the server has either refused or tried to reconnect and caused an error.

If the response is HTTP 200 OK, a screenshot of the response of the VPN will be compared to a screenshot of the control response. It will either be different or equal. If the response is different, it is categorized into the following 3 categories:

- Total block: the website indicates that the connection is refused for reasons of avoiding DDoS attacks or a VPN connection being detected and purposely blocking it.
- CAPTCHA: the website provides a page that asks the user to fill in a CAPTCHA or redirects the user through a browser check to verify that the user is human.
- Empty page: the VPN connection shows a page without any content while the control connection shows a properly loaded website.

There are also other types of blocking that could occur. Examples are functionality blocking where a certain functionality is not usable for the website visitor or feature blocking, where certain parts of a websites are completely left out of the response. These types of blocking are unique for a certain web page and do not give complete inaccessibility to the user. Thus in this paper there is no focus on these specific kinds of blocking. Next to this, some domains might be wrongfully flagged for being blocked because their website does not load fast enough and thus the page is assumed to be empty. This is partly avoided by adding a 30 second delay during which the website can load. After this, if the page is not loaded, it is assumed to be blocked. Because of this, some of the pages might still just be too slow and thus will create a small bias towards the percentage of blocking measured.

An overview of the method of determining the category can be seen in figure 1.

After categorizing the websites, the percentages of the differentiated responses are compared using statistical analysis to decide what effect a VPN connection has on the amount of blocking that occurs when visiting popular websites.

3 Experimental Setup

To retrieve websites and their subpages, a web crawler is used. Many different crawling and scraping tools are suitable for this purpose but not all have the correct qualifications required for this experiment. Requirements for the tool is that it supports JavaScript handling and can simulate browser behaviour such as pop-ups, automatic download and other browser settings. Next to this bot detection by websites should be minimized since this will lead to false positives when performing a request. Scrapy [26] is very fast, which could be good for gathering large quantity of data, but it does not handle JavaScript and is known for triggering bot related blocking. Urllib [27] and Mechanical Soup [28] also do not handle JavaScript and are thus not suitable. Both Selenium [29] and Puppeteer [30] are very suitable and meet the requirements. The only disadvantage is that they can be very slow and thus decrease the size of the data set. Puppeteer is only compatible with Chrome while Selenium is also compatible with different browsers such as Safari and Firefox. Both will yield proper results but since Selenium is compatible with multiple browsers and also multiple programming languages (Python, Java, JavaScript, etc.) and their libraries, Selenium will be used for the experiment. This way the experiment will be reproducible using different browsers and compatible with more libraries. The browser used is the Chrome webdriver using Selenium implemented in Python. The Chrome browser was chosen because it is currently the most used browser with a market share of 64.73% in May 2021 [31].

For completing the setup, two browser extensions are used. Firstly, an extension to handle cookie acceptance. Many pages ask the user to accept cookies through pop-up windows in the page. These can cause the webpage to have a different appearance as it would usually have, thus is it preferable to block these pop-ups. To do this, the *I don't care about cookies* [32] extension is used that immediately upon page loading tries to accepts all cookies thus reducing the number of popups shown. Next to this, ads can be a large part of a web page and since they continually change, they can give false positives when looking for block pages during screenshot comparison. Thus, an ad blocker is used to reduce the number of ads loaded. The ad blocker *AdBlock Plus* [33] is used for this purpose.

3.1 Vantage Points

The control connection for the experiment is a regular commercial connection and is setup in the Netherlands. To reduce geo location-based blocking, the location of the VPN IP addresses is also chosen to be located in the Netherlands. ProtonVPN basic offers 9 VPN IP addresses in the Netherlands. These are iterated over the course of the experiment. One bias that could occur is that it is assumed that when ProtonVPN states that an IP address is located in the Netherlands, that this is also actually the case. However since “advertised server locations cannot be relied upon” as stated by Weinberg et al. [34] who showed that at most 70% of the locations of the IP addresses shown by VPNs are actually in these locations, it is important to make sure that the IP addresses used in the experiment are actually located in the Netherlands. Through a IP tracking service online it can be stated that all of the IP addresses provided by ProtonVPN basic that are located in the Netherlands are located in either Amsterdam, Naaldwijk or Roosendaal and thus this gives a positive expectation that the connection through ProtonVPN basic does actually route through the Netherlands and thus reducing geolocation blocking in the experiment.

3.2 Website list

For the popular website list, the Alexa top domain [35] list is used. This is a website list created by Amazon by listing the websites with the most daily unique visitors and average page views per visitor over the last three months. Since the list is updated constantly, it gives a good overview of the most world-wide popular websites at time of testing. The list is retrieved once at the beginning of the experiment (on May 25 2021) and used for all different iterations. The Alexa is also used during multiple censorship research such as the work of Raman et al. [15], Darer et al. [18] and Niaki et al. [23].

3.3 DNS server

ProtonVPN’s default settings use it’s own internal DNS servers. Since we want to mainly flag website based blocking, the amount of DNS blocking should be minimized and be the same for both connections. For this purpose an external DNS server is used for both connections. The DNS server used is the Google public DNS main and additional server. This server is used since it is the number 2 topped ranked DNS server on the list of top DNS servers in 2021 [36]. The number one is Cloudflare. This is not used as Cloudflare is often the cause of VPN blocking and thus might give a bias if the DNS server is also blocking VPN connections. There is no way of telling that Google DNS does not perform any VPN blocking but this does not seem to be the case based on the results.

3.4 Perceptual Hashing

To be able to determine whether a webpage consist of blocking, a screenshot is made for both the control connection and the VPN connection. These are compared to determine whether the VPN response is equal or different to the webpage. To reduce the number of required manual checks, the pool of screenshots is reduced by using a threshold for the difference in perceptual hash. Perceptual hashing is a hashing method that gives similar hash values to images that are similar, in contrary to regular hashing methods. This way the difference between two screenshots of webpages can be expressed numerically through the difference in hash value. A training suite of 500 website screenshots pairs was manually checked to identify the minimum value where blocking occurs. From this training set the threshold value of 26 was determined that was afterwards checked with a test suite of a different 500 website screenshots pairs for which no block pages were found below this value. For the selection of the domains for both suites, a 1000 domain list was extracted out of the Alexa top website list and each domain was randomly assigned to either of the suites. This way the popularity level of both suites are similar and no bias is introduced. This threshold is the minimum difference value for which manual checks are done. For websites with a difference value below this threshold, it will be assumed that the response of the VPN connection is equal to the response of the control connection. For pages with a difference value above the threshold, a manual visual comparison is done and the page is categorized accordingly. This process can cause some block pages which are more subtle, thus having a low difference value, to be categorized as not blocked. Since the control suite gave the same results as the training suite, it is assumed that this amount of wrongfully categorization is small.

3.5 Sub pages

For a set of websites, three sub pages will be requested. This means that links that are embedded in the HTML file of a web page are extracted. These links will then be processed in a new request and the response will be processed in the same way as regular responses. The links to be requested are selected in sequential order where links that are equal to the original domain are disregarded.

4 Results

In total two different experiments were executed. Firstly, there was an iteration for each VPN node (total of 9). Secondly, there was an iteration for 5 different days when using the same VPN node (NL-FREE#1). For every iteration, a total of 1500 domains was requested from the Alexa top website list. For the first 500 domains, both the main page and a maximum of 3 sub pages was requested and for the remaining 1000 only the main page was requested. All the categorized responses for all iterations can be found in Appendix A and all categorized blocking can be found in Appendix B.

For the first experiment, an overview of the percentage of different responses and thus assumed blocking can be seen in figure 2. The average percentage of blocking is 1.12% with a standard deviation of 0.43% and a median of 0.90%.

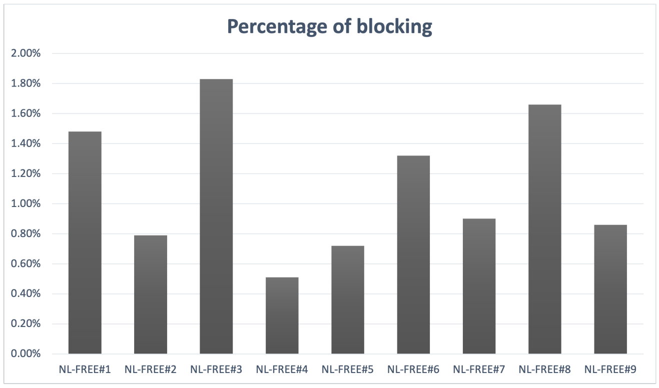


Figure 2: Percentages of blocking for all different VPN nodes

To determine whether there is a difference in blocking between the different VPN nodes, a Chi Square test was performed on each pair of distributions from 2 different VPN IP addresses used. This type of test is used since it fits the categorical data and the question that is asked, namely whether the distribution of the two nodes is equal or not. The test is performed with a significance level of 99% and 7 degrees of freedom. The hypothesis used for the test are the following:

$$H_0 = \text{ The amount of blocking for VPN Node A is equal to VPN Node B}$$

$$H_1 = \text{The amount of blocking for VPN Node A is different from VPN Node B}$$

The resulting p-values for all pair can be found in Appendix C. From the total of 36 pairs, 30 pairs have a p-value between 0.1 and 0.99 that indicates that the outcome is insignificant and thus no conclusion can be made from these pairs. For 5 pairs, the p-value is below 0.01 and thus the null hypothesis is rejected and there is an indication that there is a difference in blocking for these pairs. For one pair, the p-value is above 0.99 and thus the null hypothesis is not rejected, and it is assumed that the amount of blocking for both nodes is equal. Using this many pairs can cause getting a significant result when there is none (Type 1 error) thus the Holm-Bonferroni method can be used to determine for which values, the test can be seen as significant and for which it cannot. The target alpha level is 0.01 and the number of tests is 36. For the 3 pairs with the smallest p-value, the test is significant and the null hypothesis can be rejected with confidence but for the next 2 pairs it shows that they are not significant. For these pairs the null hypothesis is not rejected. In total this results in only 3 out of 36 pairs being significant to reject the null hypothesis and thus showing that the amount of blocking is equal for both nodes. Even though for most of the pairs the statistical test gives no significant results, it can still be seen from the chart that the percentage of blocking is not constant over the VPN nodes. The average of the types of blocking can be seen in figure 3. The numbers for these distributions are too low to do a proper statistical comparison test.

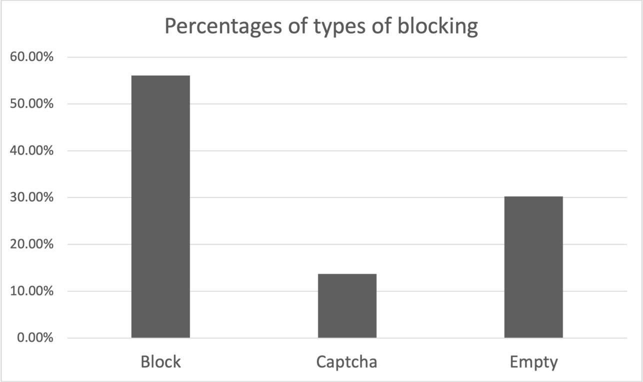


Figure 3: Percentages of blocking types for all different VPN nodes

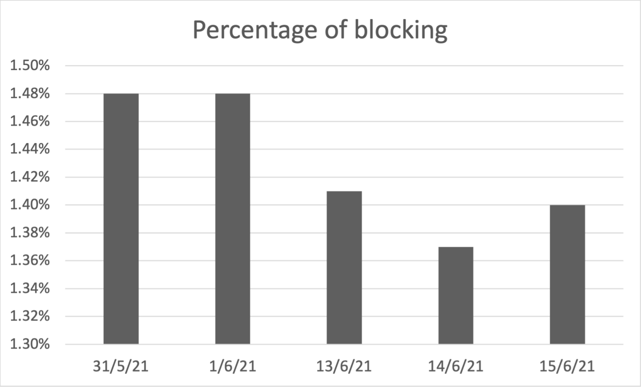


Figure 4: Percentages of blocking for all different days

For the second experiment, the overview can be seen in figure 4. The average percentage of blocking is 1.43% with a standard deviation of 0.04% and a median of 1.41%. For these outcomes, the same Chi Square test was performed from which the p-values can be seen in Appendix D. The hypothesis used for these tests are the following:

$$H_0 = \text{ The amount of blocking for day A is equal to day B}$$

$$H_1 = \text{The amount of blocking for day A is different from day B}$$

Out of 10 pairs, 2 had a p-value below 0.01 and thus the amount of blocking between days is assumed to be different and 1 had a p-value above 0.99 that indicates that the amount of blocking is equal. For the other pairs, the value lays between 0.01 and 0.99 and are thus not significant enough to make a conclusion. When using the Holm-Bonferroni Method on these p-values, it can be seen that all values under 0.01 are significant and thus there is no indication for any of these pairs to reject the null hypothesis. For this experiment it can also be seen that there is no significant conclusion to be taken from the statistical tests. The standard deviation for

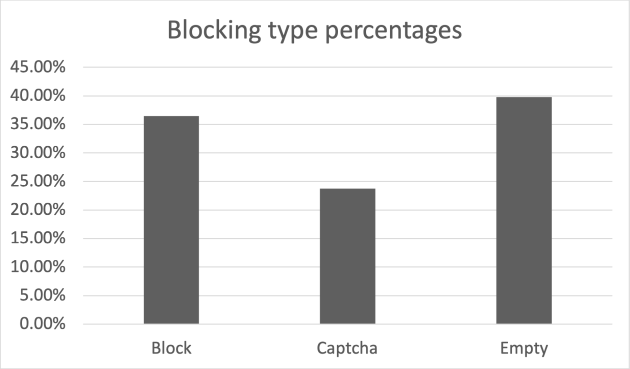


Figure 5: Percentages of blocking for all different days

this distribution is considerably smaller than the standard deviation for all the different VPN nodes (0.04% compared to 0.43%). It can thus be seen that there is not much variability between days. There is although not enough evidence to claim that there is no difference in days for the VPN nodes in general. To be able to do this, more experiments must be done using different VPN nodes. Because of the limited time span of the project this was not done. The average of the types of blocking can be seen in figure 5.

For all of the iterations, a list of websites is created that show blocking behaviour. These lists are combined to have a list of all domains which show blocking behaviour. For every domain in the list, the category of website content is decided through the Website Categorization API [37]. The categories with a frequency higher than 2%, which in practice also means more than 2 domains, can be seen in figure 6. It can be seen that there are three main categories that contain most of the blocking: Business, Online Shopping and News. It is also seen that for the News category, the majority of websites block using a total block page or an empty page thus completely disallowing the user access to the website. The full overview of distribution of categories for the different types of blocking can be found in appendix E.

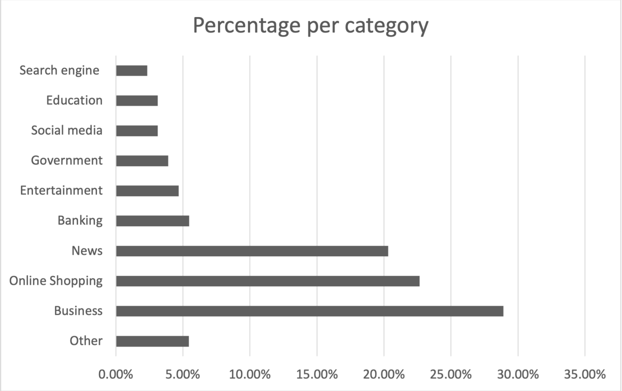


Figure 6: Categories of common blocked websites

5 Responsible Research

Research related to computer networks can infer with a lot of different parties. Namely all the parties that are part of the used network such as DNS servers, website servers, ProtonVPN and the control connection of the network. Firstly, for the DNS servers the free Google DNs server and additional server is used. Google itself claims that these servers are used by 10% of the internet users [38] thus indicating that the server usage is very high and the experiment will not impact the Google DNS system. Next to this, Google DNS is available all over the world and thus can be used to reproduce the experiment in any environment. Secondly, the website servers could be impacted by the amount of requests that are being sent during the experiment. Since the Alexa ranking is partially based on the amount of visitors a website received per day, the websites are used to a large amount of visitors and thus a large amount of traffic. Every page will be loaded a maximum of 8 times (1 main pages and 3 sub pages for both VPN and control connection) per day, this is a very small number compared to the total amount of traffic. Thus the chance of affecting the website servers is very small.

Next to this, some IP addresses from ProtonVPN will be used, and through performing extensive web crawling may become blocked by some websites. Since the amount of requests is small, this is unlikely. The amount of traffic routed through the ProtonVPN network could also impact the performance of the ProtonVPN servers. Again, since the amount of crawling is small and the ProtonVPN servers have a minimum of 1 Gbps bandwidth [39], the probability of doing any damage is very small. Finally, the control connection could be affected by the IP address being blocked for bot behaviour. Again since the amount of crawling is low per web page this will be unlikely.

Next to research being ethical to all parties, it is also important to be reproducible, For this, the tools and materials, such as the website list, used should be available to anyone. This is the case for ProtonVPN since the basic version is used that is free. It is usable with all different operating systems and all countries. Google DNS is also free to use and accessible from all over the world. Lastly, the Alexa 1 million list of domain is also freely available for download. The domain list and the used code for the experiment can be found in the projects github [40]. Lastly, both extensions *I don't care about cookies* and *Adblock Plus* are free to use tools that are available for different browsers such as Microsoft Edge, Firefox and Chrome. Thus all the tools are accessible and make the study reproducible.

6 Conclusion

The results from the experiment show that the average frequency of blocking is 1.12% when looking at all different VPN nodes from ProtonVPN. This indicates that around 1 out of 100 websites are blocked for users of the VPN. It also shows that the frequency is not always the same and can range from 0.51% to 1.83%. This range shows that some VPN nodes are 3 time more likely than others and that the fact that a websites is blocked for one node should not give the assumption that it is also blocked for others.

The results also show that the amount of blocking varies in a smaller amount when using the same node over different days. There is still a small variation which could mean that websites do not have a constant blocking on IP addresses but may some times block while other times it will not. Still, more data should be gathered to be make a conclusion about effect of difference in days on the amount of blocking on different VPN nodes.

It is also shown that a major part of the blocking is of the total block type or empty page that means that the user is totally disallowed to even access the website in any way. This does not give the user a possibility of proving themselves against the website that it is not malicious, which would be the case for CAPTCHA blocking. The problem with this is that CAPTCHA does not protect the website against any cyber attacks, which blocked pages do.

7 Future Work

This paper gives a method to detect different types of VPN blocking by websites. The main block types that it focuses on are total block, CAPTCHA and empty pages. Other types of blocking are not detected but could still occur such as functionality blocking or feature blocking where the website is actually visible but a section or part of the website is not functional for the user. Future work could focus more on this, to get a more accurate view of the amount of blocking. This does mean that the amount of blocking found in this paper can be seen as a minimum since other types of blocking will add to this existing amount thus increasing the percentage of blocked websites.

Next to this, the method created is only tested in a limited experiment both in time, and data set. Firstly when dealing with time, only a short amount of iterations throughout different days was conducted and there was no possibility to compare the amount of blocking over a span of weeks or even months. This could be one of the goals of future research to get a clear overview of how blocking progresses throughout time.

The data set for the experiments done is limited to only Dutch VPN nodes since the control node necessary was located in the Netherlands. ProtonVPN supports VPN nodes in 55 different countries with a total of 1246 nodes at this time thus the 9 nodes tested in this paper is a very limited part of the total ProtonVPN network. ProtonVPN is also one of the many current commercial VPN providers and thus there should be more experiments done to get a understanding of whether the frequency of blocking is similar for different services.

Next to this, only one data set of domains is used with a size of 1500 which is a very limited list of domains when considering that the Alexa top domains keeps a 1 million domain list. Other website lists could also be explored whether there is a difference between popular and often used websites and smaller less popular websites. Future work could expand on both the VPN node issue and the website list issue.

Currently, there are several tools such as Nymble [41] that offers websites the possibility of blocking only malicious VPN users. This is done by using anonymous authentication tokens to register users through external servers. This method keeps the users privacy intact while giving the website the opportunity to blacklist malicious users without blacklisting an entire IP address. The problem with this approach is that external servers are needed which would have to be large scale when used for popular websites. To avoid these limitations, more research should be done to find a method that would fit into networks which have a high amount of traffic and how this would be integrated into the internet. This would be necessary to allow VPN users to be able to have access to all of the internet without compromising their privacy.

8 Acknowledgements

During the process of writing this paper several topics were researched through shared knowledge within the Research Project group. The types of blocking were researched by A. Pingle[42], a literary review was done by both F. Biazin do Nascimento [43] and P. Iacoban [44] and research regarding different possible website lists was conducted by J. Mulder [45]. This research was shared and used for several topics in this paper.

References

[1] N. Aase, J. Crandall, A. Diaz, J. Knockel, J. Molinero, J. Saia, D. Wallach, and T. Zhu, "Whiskey, weed, and wukan on the world wide web: On measuring censors' resources and motivations," in *2nd USENIX Workshop on Free and Open Communications on the Internet (FOCI 12)*, (Bellevue, WA), USENIX Association, Aug. 2012.

[2] J. Liu and J. Zhao, "More than plain text: Censorship deletion in the chinese social media," *Journal of the Association for Information Science and Technology*, vol. 72, no. 1, p. 18–31, 2021.

[3] N. Zubair, "The anatomy of web censorship in pakistan," in *Proceedings of the 3rd USENIX Workshop on-Free and Open Communications on the Internet*, vol. 13, 2013.

[4] T. K. Yadav, A. Sinha, D. Gosain, P. K. Sharma, and S. Chakravarty, "Where the light gets in: Analyzing web censorship mechanisms in india," in *Proceedings of the Internet Measurement Conference 2018*, IMC '18, (New York, NY, USA), p. 252–264, Association for Computing Machinery, 2018.

[5] G. Aceto and A. Pescapé, "Internet censorship detection: a survey," *Computer Networks*, vol. 83, p. 381–421, 2015.

[6] J. Verkamp and M. Gupta, "Inferring mechanics of web censorship around the world," in *2nd USENIX Workshop on Free and Open Communications on the Internet (FOCI 12)*, (Bellevue, WA), USENIX Association, Aug. 2012.

[7] K. Bock, G. Hughey, X. Qiang, and D. Levin, "Geneva: Evolving censorship evasion strategies," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, CCS '19, (New York, NY,

USA), p. 2199–2214, Association for Computing Machinery, 2019.

[8] T.-F. Yen, Y. Xie, F. Yu, R. P. Yu, and M. Abadi, "Host fingerprinting and tracking on the web: Privacy and security implications.," in *NDSS*, The Internet Society, 2012.

[9] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *Proceedings of the 13th Conference on USENIX Security Symposium - Volume 13*, SSYM'04, (USA), p. 21, USENIX Association, 2004.

[10] C. Burkert, J. McDougall, H. Federrath, and M. Fisher, "Analysing leakage during vpn establishment in public wi-fi networks," in *IEEE International Conference on Communications 2021*, 2021.

[11] A. Chaabane, T. Chen, M. Cunche, E. D. Cristofaro, A. Friedman, and M. A. Kaafar, "Censorship in the wild: Analyzing internet filtering in syria," 2014.

[12] V. Polyakov, "Criminalistics specifics of methods of committing computer crimes and peculiarities of their prevention," *Religación. Revista de Ciencias Sociales y Humanidades*, vol. 4, pp. 90–97, Nov. 2019.

[13] T. Nihal, "A criminological examination of the use of cyberspace to traffic drugs in durban south africa," *The International Journal of Social Sciences and Humanities Invention*, p. 5864–5871, 2020.

[14] U. Taheral, M. Asma, and Q. M, "Blacklisting misbehaving users for enhancing security in anonymizing networks," *IOSR Journal of Computer Engineering*, vol. 9, no. 5, pp. 15–20, 2013.

[15] R. Sundara Raman, P. Shenoy, K. Kohls, and R. Ensafi, "Censored planet: An internet-wide, longitudinal censorship observatory," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, CCS '20, (New York, NY, USA), p. 49–66, Association for Computing Machinery, 2020.

[16] G. Aceto, A. Botta, A. Pescapè, N. Feamster, F. Awan, T. Ahmad, and S. Qaisar, "Monitoring internet censorship with ubica," in *Conference: International Workshop on Traffic Monitoring and Analysis*, pp. 143–157, 04 2015.

[17] S. Burnett and N. Feamster, "Encore: Lightweight measurement of web censorship with cross-origin requests," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, SIGCOMM '15, (New York, NY, USA), p. 653–667, Association for Computing Machinery, 2015.

[18] A. Darer, O. Farnan, and J. Wright, "Filteredweb: A framework for the automated search-based discovery of blocked urls," in *2017 Network Traffic Measurement and Analysis Conference (TMA)*, pp. 1–9, 2017.

[19] J. Crandall, D. Zinn, M. Byrd, E. Barr, and R. East, "Conceptdoppler: A weather tracker for internet censorship," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, CCS '07, (New York, NY, USA), p. 352–365, Association for Computing Machinery, 2007.

[20] B. VanderSloot, A. McDonald, W. Scott, J. A. Halderman, and R. Ensafi, "Quack: Scalable remote measurement of application-layer censorship," in *Proceedings of the 27th USENIX Conference on Security Symposium*, SEC'18, (USA), p. 187–202, USENIX Association, 2018.

[21] A. Sfakianakis, E. Athanasopoulos, and S. Ioannidis, "S.: Censmon: A web censorship monitor," in *In: USENIX FOCI 2011*, 2011.

[22] A. Nisar, A. Kashaf, I. A. Qazi, and Z. A. Uzmi, "Incentivizing censorship measurements via circumvention," in *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, SIGCOMM '18, (New York, NY, USA), p. 533–546, Association for Computing Machinery, 2018.

[23] A. Niaki, S. Cho, Z. Weinberg, N. Hoang, A. Razaghpanah, N. Chrisin, and P. Gill, "Iclab: A global, longitudinalinternet censorship measurement platform," in *Proceedings of the 41st IEEE Symposium on Security and Privacy*, 2020.

[24] R. Singh, R. Nithyanand, S. Afroz, P. Pearce, M. C. Tschantz, P. Gill, and V. Paxson, "Characterizing the nature and dynamics of tor exit blocking," in *Proceedings of the 26th USENIX Conference on Security Symposium*, SEC'17, (USA), p. 325–341, USENIX Association, 2017.

[25] S. Khattak, D. Fifield, S. Afroz, M. Javed, S. Sundaresan, V. Paxson, S. Murdoch, and D. Mccoy, "Do you see what i see? differential treatment of anonymous users," in *The Network and Distributed System Security Symposium 2016*, feb 2016.

[26] "Scrapy — a fast and powerful scraping and web crawling framework." https://scrapy.org.

[27] "urllib3 · pypi." https://pypi.org/project/urllib3/.

[28] "Mechanicalsoup · pypi." https://pypi.org/project/MechanicalSoup/.

[29] "Seleniumhq browser automation." https://www.selenium.dev/.

[30] "Puppeteer." https://pptr.dev/.

[31] "Browser market share worldwide." https://gs.statcounter.com/browser-market-share.

[32] "I don't care about cookies 3.3.0." https://www.i-dont-care-about-cookies.eu/.

[33] "Adblock plus — the world's #1 free ad blocker." https://adblockplus.org/.

[34] Z. Weinberg, S. Cho, N. Christin, V. Sekar, and P. Gill, "How to catch when proxies lie: Verifying the physical locations of network proxies with active geolocation," in *Proceedings of the Internet Measurement Conference 2018*, IMC '18, (New York, NY, USA), p. 203–217, Association for Computing Machinery, 2018.

[35] "Alexa - top sites." https://www.alexa.com/topsites.

[36] "Best free and public dns servers in 2021." https://www.techradar.com/news/best-dns-server.

[37] "Website categorization api." https://main.whoisxmlapi.com/.

[38] "Google dns turns 8888 years old." https://security.googleblog.com/2018/08/google-public-dns-turns-8888-years-old.html.

[39] "Protonvpn - security features." https://protonvpn.com/secure-vpn.

[40] "Github for project including code and website list." https://github.com/WillemijnTutu/RPWillemijn_Tutuarima.

[41] P. C. Johnson, A. Kapadia, P. P. Tsang, and S. W. Smith, "Nymble: Anonymous ip-address blocking," in *Privacy Enhancing Technologies* (N. Borisov and P. Golle, eds.), (Berlin, Heidelberg), pp. 113–133, Springer Berlin Heidelberg, 2007.

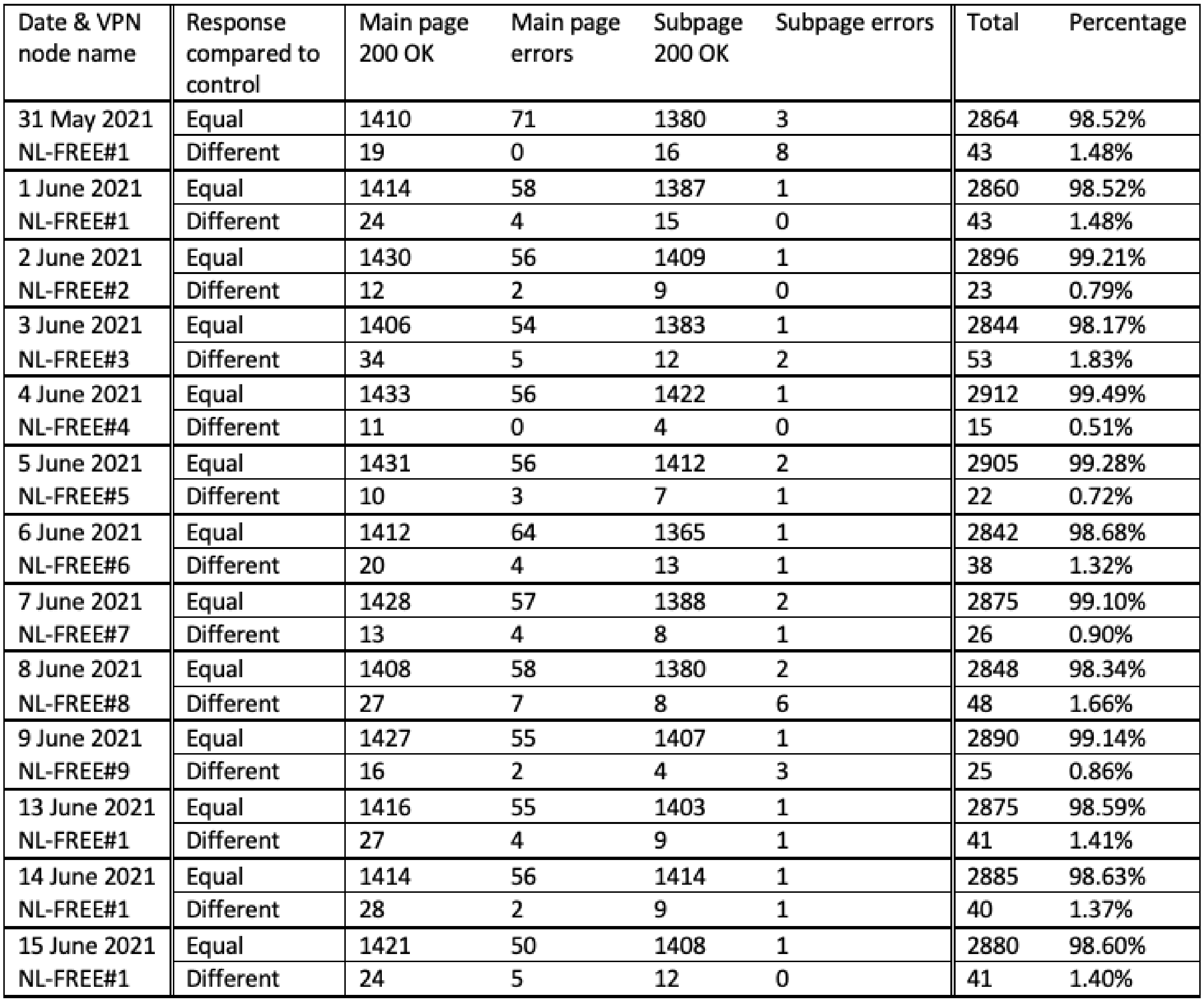
[42] A. Pingle, "Measuring accessibility of popular websites while using tor."

[43] F. B. do Nascimento, "Measuring accessibility of popular websites while using mullvad."

[44] P. Iacoban, "Measuring accessibility of popular websites while using the i2p anonymity network."

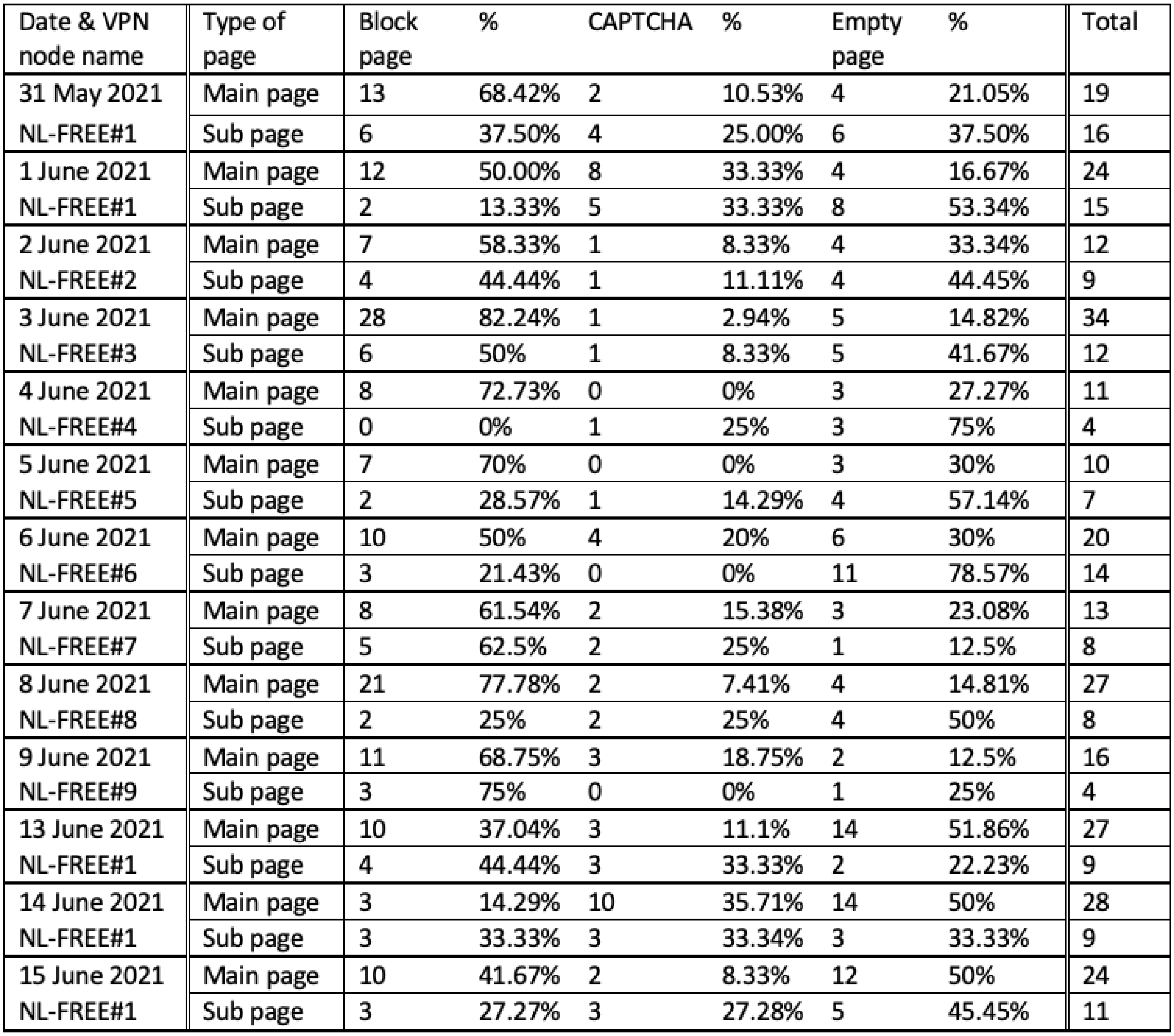
[45] J. Mulder, "Measuring the blocking of an.on users by popular websites through web scraping."

A Results for reponses



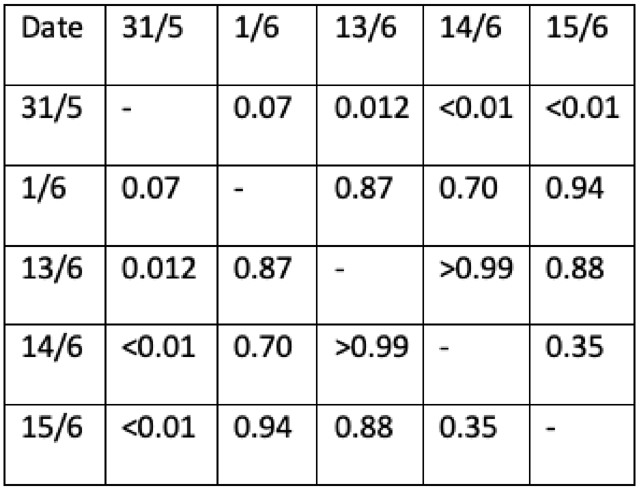
Date & VPN node name	Response compared to control	Main page 200 OK	Main page errors	Subpage 200 OK	Subpage errors	Total	Percentage
31 May 2021 NL-FREE#1	Equal	1410	71	1380	3	2864	98.52%
	Different	19	0	16	8	43	1.48%
1 June 2021 NL-FREE#1	Equal	1414	58	1387	1	2860	98.52%
	Different	24	4	15	0	43	1.48%
2 June 2021 NL-FREE#2	Equal	1430	56	1409	1	2896	99.21%
	Different	12	2	9	0	23	0.79%
3 June 2021 NL-FREE#3	Equal	1406	54	1383	1	2844	98.17%
	Different	34	5	12	2	53	1.83%
4 June 2021 NL-FREE#4	Equal	1433	56	1422	1	2912	99.49%
	Different	11	0	4	0	15	0.51%
5 June 2021 NL-FREE#5	Equal	1431	56	1412	2	2905	99.28%
	Different	10	3	7	1	22	0.72%
6 June 2021 NL-FREE#6	Equal	1412	64	1365	1	2842	98.68%
	Different	20	4	13	1	38	1.32%
7 June 2021 NL-FREE#7	Equal	1428	57	1388	2	2875	99.10%
	Different	13	4	8	1	26	0.90%
8 June 2021 NL-FREE#8	Equal	1408	58	1380	2	2848	98.34%
	Different	27	7	8	6	48	1.66%
9 June 2021 NL-FREE#9	Equal	1427	55	1407	1	2890	99.14%
	Different	16	2	4	3	25	0.86%
13 June 2021 NL-FREE#1	Equal	1416	55	1403	1	2875	98.59%
	Different	27	4	9	1	41	1.41%
14 June 2021 NL-FREE#1	Equal	1414	56	1414	1	2885	98.63%
	Different	28	2	9	1	40	1.37%
15 June 2021 NL-FREE#1	Equal	1421	50	1408	1	2880	98.60%
	Different	24	5	12	0	41	1.40%

B Response for blocking



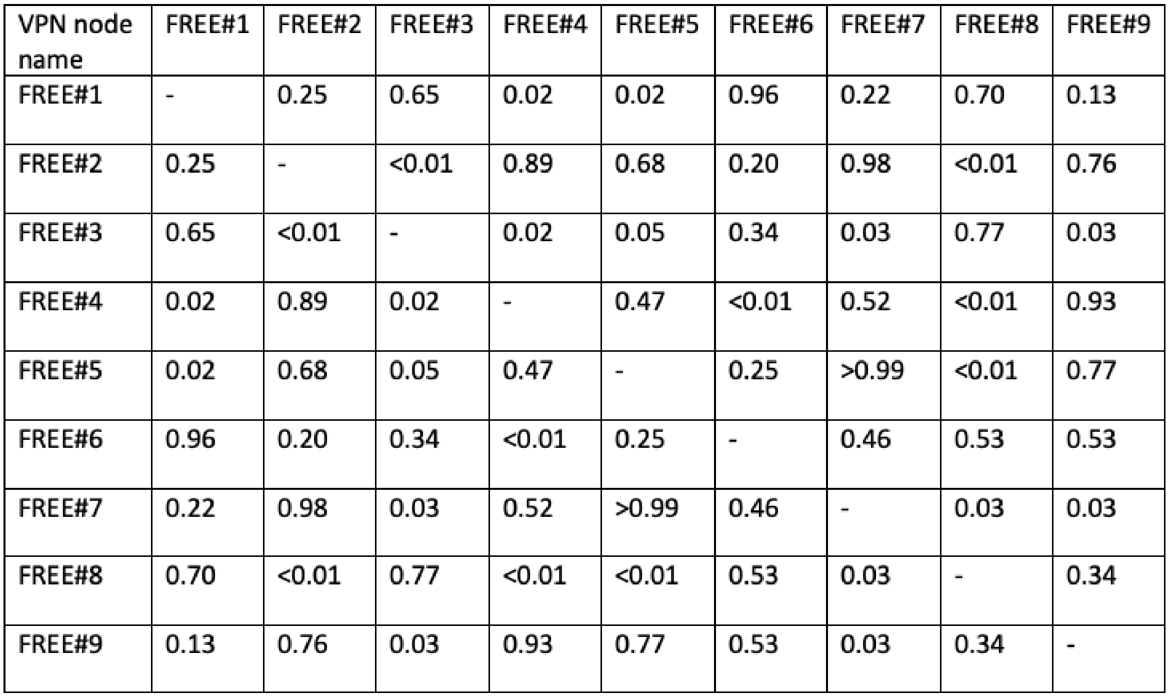
Date & VPN node name	Type of page	Block page	%	CAPTCHA	%	Empty page	%	Total
31 May 2021	Main page	13	68.42%	2	10.53%	4	21.05%	19
NL-FREE#1	Sub page	6	37.50%	4	25.00%	6	37.50%	16
1 June 2021	Main page	12	50.00%	8	33.33%	4	16.67%	24
NL-FREE#1	Sub page	2	13.33%	5	33.33%	8	53.34%	15
2 June 2021	Main page	7	58.33%	1	8.33%	4	33.34%	12
NL-FREE#2	Sub page	4	44.44%	1	11.11%	4	44.45%	9
3 June 2021	Main page	28	82.24%	1	2.94%	5	14.82%	34
NL-FREE#3	Sub page	6	50%	1	8.33%	5	41.67%	12
4 June 2021	Main page	8	72.73%	0	0%	3	27.27%	11
NL-FREE#4	Sub page	0	0%	1	25%	3	75%	4
5 June 2021	Main page	7	70%	0	0%	3	30%	10
NL-FREE#5	Sub page	2	28.57%	1	14.29%	4	57.14%	7
6 June 2021	Main page	10	50%	4	20%	6	30%	20
NL-FREE#6	Sub page	3	21.43%	0	0%	11	78.57%	14
7 June 2021	Main page	8	61.54%	2	15.38%	3	23.08%	13
NL-FREE#7	Sub page	5	62.5%	2	25%	1	12.5%	8
8 June 2021	Main page	21	77.78%	2	7.41%	4	14.81%	27
NL-FREE#8	Sub page	2	25%	2	25%	4	50%	8
9 June 2021	Main page	11	68.75%	3	18.75%	2	12.5%	16
NL-FREE#9	Sub page	3	75%	0	0%	1	25%	4
13 June 2021	Main page	10	37.04%	3	11.1%	14	51.86%	27
NL-FREE#1	Sub page	4	44.44%	3	33.33%	2	22.23%	9
14 June 2021	Main page	3	14.29%	10	35.71%	14	50%	28
NL-FREE#1	Sub page	3	33.33%	3	33.34%	3	33.33%	9
15 June 2021	Main page	10	41.67%	2	8.33%	12	50%	24
NL-FREE#1	Sub page	3	27.27%	3	27.28%	5	45.45%	11

C P values for Chi Square test for different days



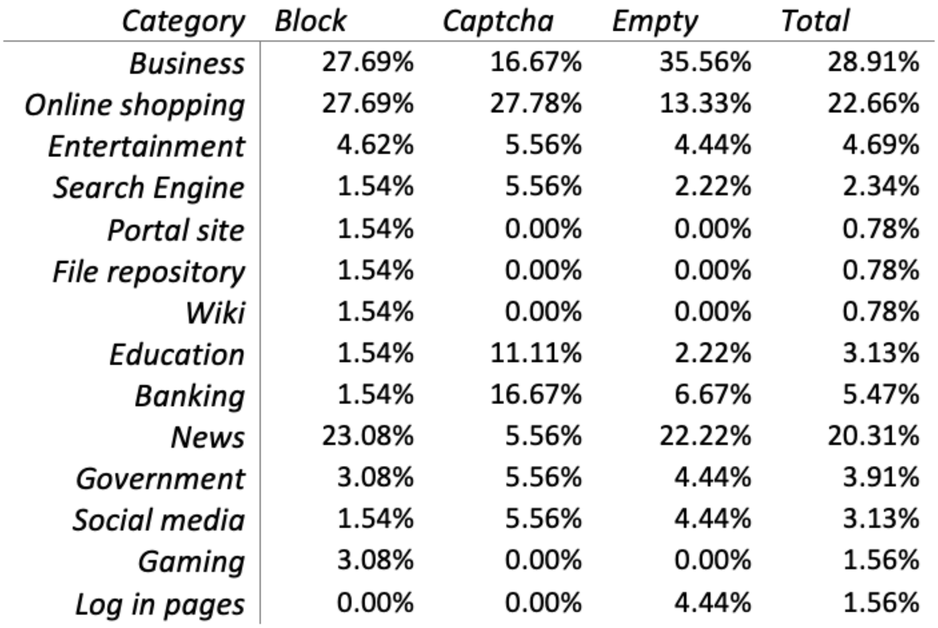
Date	31/5	1/6	13/6	14/6	15/6
31/5	-	0.07	0.012	<0.01	<0.01
1/6	0.07	-	0.87	0.70	0.94
13/6	0.012	0.87	-	>0.99	0.88
14/6	<0.01	0.70	>0.99	-	0.35
15/6	<0.01	0.94	0.88	0.35	-

D P values for Chi Square test for different VPN nodes



VPN node name	FREE#1	FREE#2	FREE#3	FREE#4	FREE#5	FREE#6	FREE#7	FREE#8	FREE#9
FREE#1	-	0.25	0.65	0.02	0.02	0.96	0.22	0.70	0.13
FREE#2	0.25	-	<0.01	0.89	0.68	0.20	0.98	<0.01	0.76
FREE#3	0.65	<0.01	-	0.02	0.05	0.34	0.03	0.77	0.03
FREE#4	0.02	0.89	0.02	-	0.47	<0.01	0.52	<0.01	0.93
FREE#5	0.02	0.68	0.05	0.47	-	0.25	>0.99	<0.01	0.77
FREE#6	0.96	0.20	0.34	<0.01	0.25	-	0.46	0.53	0.53
FREE#7	0.22	0.98	0.03	0.52	>0.99	0.46	-	0.03	0.03
FREE#8	0.70	<0.01	0.77	<0.01	<0.01	0.53	0.03	-	0.34
FREE#9	0.13	0.76	0.03	0.93	0.77	0.53	0.03	0.34	-

E Percentages of blocking per category



Category	Block	Captcha	Empty	Total
Business	27.69%	16.67%	35.56%	28.91%
Online shopping	27.69%	27.78%	13.33%	22.66%
Entertainment	4.62%	5.56%	4.44%	4.69%
Search Engine	1.54%	5.56%	2.22%	2.34%
Portal site	1.54%	0.00%	0.00%	0.78%
File repository	1.54%	0.00%	0.00%	0.78%
Wiki	1.54%	0.00%	0.00%	0.78%
Education	1.54%	11.11%	2.22%	3.13%
Banking	1.54%	16.67%	6.67%	5.47%
News	23.08%	5.56%	22.22%	20.31%
Government	3.08%	5.56%	4.44%	3.91%
Social media	1.54%	5.56%	4.44%	3.13%
Gaming	3.08%	0.00%	0.00%	1.56%
Log in pages	0.00%	0.00%	4.44%	1.56%