Internet censorship in the European Union

Doctoral Thesis

to acquire the academic degree of
Doctor rerum politicarum
(Doctor of Economics and Management Science)

submitted to the

School of Business and Economics of
Humboldt-Universität zu Berlin

by

(MSc. Vasilis Ververis)

President of Humboldt-Universität zu Berlin:
Prof. Dr. Peter Frensch (kommissarisch)

Dean of the School of Business and Economics:
Prof. Dr. Daniel Klapper

Reviewers:
1. Prof. Dr. Benjamin Fabian
2. Prof. Dr. Stefan Lessmann
3. Prof. Dr. Stefania Milan

Date of Colloquium: 06.09.2022

– 0 –

Contents

Abstract

This is a thesis on Internet censorship in the European Union (EU), specifically regarding the technical implementation of blocking methodologies and filtering infrastructure in various EU countries. The analysis examines the use of this infrastructure for information controls and the blocking of access to websites and other network services available on the Internet. The thesis follows a three-part structure. Firstly, it examines the cases of Internet censorship in various EU countries, specifically Greece, Cyprus, and Spain. Subsequently, this paper presents a new testing methodology for determining censorship of mobile store applications. Additionally, it analyzes all 27 EU countries using historical network measurements collected by Open Observatory of Network Interference (OONI) volunteers from around the world, publicly available blocklists used by EU member states, and reports issued by network regulators in each country.

Abstract

Diese Arbeit befasst sich mit Internetzensur innnerhalb der EU, und hier insbesondere mit der technischen Umsetzung, das heißt mit den angewandten Sperrmethoden und Filterinfrastrukturen, in verschiedenen EU-Ländern. Neben einer Darstellung einiger Methoden und Infrastrukturen wird deren Nutzung zur Informationskontrolle und die Sperrung des Zugangs zu Websites und anderen im Internet verfügbaren Netzdiensten untersucht. Die Arbeit ist in drei Teile gegliedert. Zunächst werden Fälle von Internetzensur in verschiedenen EU-Ländern untersucht, insbesondere in Griechenland, Zypern und Spanien. Anschließend wird eine neue Testmethodik zur Ermittlung der Zensur mittels einiger Anwendungen, welche in mobilen Stores erhältlich sind, vorgestellt. Darüber hinaus werden alle 27 EU-Länder anhand historischer Netzwerkmessungen, die von freiwilligen Nutzern von OONI aus der ganzen Welt gesammelt wurden, öffentlich zugänglichen Blocklisten der EU-Mitgliedstaaten und Berichten von Netzwerkregulierungsbehörden im jeweiligen Land analysiert.

– 1 –
Introduction

1.1. Internet censorship in EU

1.1.1 Topic significance

The issue of internet censorship in the EUand western countries receives less research than other regions and countries, such as China and India. When I began this research years ago, I discovered that the country where I was living had authorized a non-government commission to block certain websites. They accomplish this by utilizing blocklists, which consist of (Uniform Resource Locator (URL)) listings of websites. Internet Service Providers (ISPs) are then required to censor and deny access to the URLs that are included on the blocklist. The details of the research, as well as its methodology, are presented in Chapter 2. In general, Western countries and the European Commission (EC) take the position that Internet censorship either doesn't exist or serves to protect individuals from accessing certain websites or network resources. Section 1.1.2 will provide a definition of Internet censorship. Unfortunately this stance that western countries are not censoring the Internet is apparent in many parts of our societies as well as academic conferences that review, discuss and publish research about Internet censorship. In order to provide further evidence of this issue, it is important to dedicate space to this thesis and include a quote from an anonymous reviewer of a well-established academic conference on network measurements and interference. Internet censorship exists in the EU, and this thesis outlines how ISPs in the EU restrict access to online resources.

> *First, the way such a sensitive topic is treated is not appropriate. The article does not make any effort to contextualize the observations and results to the Spanish and European legal framework and the political context. In many cases, the authors seem to consider all the reported instances of blocking as state-wide censorship, when in most cases –excluding perhaps the Catalan referendum and tsunami democratic which has a political component– it is about blocking unlawful content. [...]*
>
> Anonymous reviewer

1.1.2 Definition of Internet censorship

Internet censorship, or the unintentional blocking of Internet services and network interference, is a relatively well-researched topic. However, most studies focus their research on non-Western European countries and regions. The research focuses on EU countries in order to fill the gap of almost non-existent research on Internet censorship in the EU. The research is also justified because its revelations and discoveries add to and enrich the knowledge of the technical implementations deployed and used to control information and block content or services on the Internet in EU countries. In addition, the research

presents several examples of how Internet blocking infrastructure can be misused to restrict freedom of expression, block access to information, and suppress democratic rights.

This thesis presents research on the development of Internet censorship in the European Union. We present case studies from different countries where we perform and analyze network measurements from different angles to detect and methodologically record the techniques, software and other infrastructure used to control information or block access to specific network resources (such as specific websites). During the research, we have identified several missing network measurement methodologies that are needed to better detect network disruptions caused by filtering or blocking in networks.

1.1.3 Background

How I become interested in this area of study

I started reading about Internet censorship around the world, but mostly from countries outside of Western Europe. I started looking for more technical details on how Internet censorship is implemented in countries in Western Europe and found very little information, let alone any published work on the technical implementation. I became interested in learning more after realizing that countries in Western Europe have been blocking Internet services for some time. I hope that my work will help others to better understand the technical implementation of Internet censorship in Western European democratic countries and raise awareness about the growing phenomenon of Internet censorship and information controls.

1.1.4 Methodology

The foundation of this thesis relies heavily on network measurements and data analysis. A background in network protocols (such as the Internet protocol suite Transmission Control Protocol/Internet Protocol (TCP/IP)) specifications as well as application services (such as the World Wide Web (WWW), email, file sharing) specifications has been extremely useful in helping me to design and develop the methodologies used to conduct my research on Internet censorship detection and analysis. Through this research, I have been able to improve Internet censorship detection methodologies, develop techniques to improve and expand the longitudinal collection of network measurements from various geographically relevant vantage points. Participated in the creation of the largest freedom data repository of network measurements used for Internet censorship research, licensed under a free and permissive license available to everyone.

1.2. Contributions

Internet censorship in the EU is an under-researched field. This thesis addresses the challenges of Internet censorship in EU democracies through technical analysis. The research includes developing novel methodologies to detect network traffic interference, historical data analysis of network measurements, development of software to perform unattended longitudinal network measurements as well as Internet censorship circumvention.

- *An understanding of Internet blocking in Greece*
  Chapter 2 analyses the techniques and policies used to block content of gambling websites in Greece and the implications of the detected underblocking and overblocking. Our results highlight issues

related to how transparently Internet filtering is implemented in democratic countries and could indicate the presence of unfair competition between ISPs.

- *Internet censorship capabilities in Cyprus*
  Chapter 3 discusses a number of unreported Internet censorship cases, non-transparently implemented blocking regulations, and collateral damage due to blocking of email delivery to the regulated domains by the National Betting Authority of the Republic of Cyprus. Furthermore our results indicate the presence of at least two distinct regimes on the island based on network measurement data collected in Cyprus from five major residential ISPs.

- *Longitudinal Internet censorship analysis in Spain*
  Chapter 4 analyzes data of networks measurement research that spans over 2016 to 2020. Our analysis indicate the existence of advanced network interference techniques that grow in sophistication over time. We identified evidence of network interference from all the major ISPs in Spain, (91% of mobile and 98% of broadband users) and several governmental and law enforcement authorities. Furthermore we contribute an enhanced domain testing methodology to detect certain kinds of Transport Layer Security (TLS) blocking now included in the free and open source licensed software OONI Probe. Finally we made our research reproducible by providing the software and data analysis source code and methodology publicly available online.

- *App store censorship detection methodologies*
  Chapter 5 presents a novel methodology for querying the public search engines and APIs of major app stores (Google Play Store, Apple App Store and Tencent MyApp Store) to detect application censorship per country. Our research finds that users in specific countries do not have access to popular app stores due to local laws, financial reasons, or because countries are on a sanctions list that prohibit foreign businesses to operate within its jurisdiction. The source code is publicly available online allowing for reproducing of our methodology.

- *Website Blocking in the European Union: Network Interference from the Perspective of Open Internet*
  Chapter 6 conducts a comprehensive analysis encompassing the 27 EU member countries. This analysis is founded upon data derived from three primary sources. Firstly, it incorporates a substantial volume of historical network measurements compiled in 2020 by OONI data. Secondly, the investigation integrates publicly accessible compendiums of blocked entities employed by EU member states. Lastly, the study integrates reports disseminated by National Regulation Authorities (NRAs) spanning the time frame of May 2020 to April 2021.

In addition, the thesis contributions have seen real-world practice and adoption to censorship detection and network measurement software. detection of possible events of Internet censorship. See Section 1.2.1 for more details about the Raspberry Pi distribution Lepidopter developed to perform unattended longitudinal OONI network measurements and a hands-on collaborative documentation Magma guide (see Section 1.2.2) to conduct research, perform and analyze network measurements suitable for research fellows, human rights activists, lawyers, network engineers and technologists.

1.2.1 Lepidopter distribution

Instructing a Raspberry Pi device to run tests and report censorship is the first step to getting more and more people to provide results and help by contributing more data that could be used to conduct research on the motives, criteria, techniques used, and possibly expose the censor itself. This could be achieved if the collected data are open and available to everyone and the tests are easy to run.

My proof of concept implementation, Lepidopter [266] Raspberry Pi [258] boot image. The operating system, utilities and diagnostics on an Operating System (OS) distribution image, ready to boot and start performing tests.

Lepidopter is the codename used for the Raspberry Pi image tha comes with the necessary software and si configured to perform longitudinal network measurements from Raspberry Pi devices without requiring any physical presence or technical expertise.

The Raspberry Pi device was chosen because of the small footprint (minimal power requirements) of this embedded device (the size of a credit card), making it ideal for distribution to organisations (and possibly motivated individual volunteers) interested in becoming official OONI operators and providing measurement reports for cases of filtering or censorship. Deploying a Raspberry Pi to run tests and report censorship events is the first step towards contributing more measurement data and increasing the number vantage points. This deployment will be used to conduct research on the motives, criteria, and techniques being used, and perhaps even to directly expose censorship.

The source of Lepidopter has been released under a free/libre code license, and its detailed developer instructions [266], user documentation, including a screenshot- based guide on how to copy the image to an SD card and access the device are also publicly available for various platforms and operating systems [268].

1.2.2 Magma guide

In recent years, a number of research fellows, journalists, human rights activists, lawyers as well as a larger activist community, have been working in high-risk contexts, which create the need to consider their qualitative and quantitative research data as highly sensitive. Albeit their competitiveness and high qualification in their respective areas (social and political science, usability, law, political economy analysis), they can rarely claim to have a specific expertise or extensive experience regarding networks services and systems, telecommunication infrastructure, applied data analysis of network measurements, internet censorship, surveillance and information controls.

Magma aims to build a scalable, reproducible, standard methodology on measuring, documenting and circumventing internet censorship, information controls, internet blackouts and surveillance in a way that will be streamlined and used in practice by researchers, front-line activists, field-workers, human rights defenders, organizations and journalists.

Ideally, researchers working with various network measurement tools and frameworks such as the OONI, should have qualified technical help and assistance, thus enabling them to develop appropriate testing methodologies, suiting exactly their research environment and needs.

Magma aims to build a research framework for people working on information controls and network measurements, facilitating their working process in numerous ways. As such, this framework will enable them to properly structure an activity plan and make informed choices regarding the required tools.

Specifically the Magma guide, provides documentation in network measurements, internet censorship research, assessment of ISP network, surveillance probing and data analysis in order to:

- Asses the risks by providing, implementing and maintaining technologies demanded by researchers on front-lines and areas where the need of operational security, anti-surveillance and censorship circumvention is of paramount importance.

- Provide tailored technical assistance, developing at the same time appropriate testing methodology for network measurements, evaluation and analysis of data and reports that correspond to the respective research questions.
- On a long-term basis, build a scalable and reproducible methodology for collecting, evaluating and analyzing data and reports' self-defense for front-line researchers, front-line activists, field-workers, human rights defenders, organizations and journalists, by keeping exact documentation.

1.2.3 Thesis structure

The remainder of this thesis is structured as follows. Chapter 2 discusses the large scale website blocking deployed by the Greek government to censor access to gambling websites. The data analysis was conducted from collected network measurements by major broadband and cellular ISPs that are a representative sample of Internet usage in Greece. The chapter is based on Appendix A.1.

Chapter 3 discusses the findings of an open and collective effort towards a cross comparison study of web content blocking regulations and practices, in different parts of Cyprus. The chapter is based on Appendix A.1.

Chapter 4 provides the data analysis and Internet censorship detection methodologies in Spain over 2016 to 2020. The chapter is based on Appendix A.1.

Chapter 5 presents a novel methodology for detecting application store censorship and the availability of mobile applications across countries. The chapter is based on Appendix A.1.

Chapter 6 investigates EU member states' implementation of network monitoring and website blocking infrastructures in compliance with EU law, highlighting the lack of prior research in this area. Using data from multiple sources, it reveals diverse blocklist practices by NRAs and ISPs, along with transparency issues in addressing blocked websites. The chapter is based on Appendix A.1.

Chapter 7 concludes the thesis by summarizing and discussing future work.

All thesis chapters are based on previously peer reviewed published papers and are referenced in Appendix A.1.

– 2 –
Internet Censorship analysis in Greece

2.1. Introduction

There are many incidents of Internet censorship, which have usually been reported from countries without democratic political systems. In the following article we present methods and techniques to investigate Internet censorship based on empirical measurements. Furthermore, we present a case study from a democratic country where we analyze the state of censorship in depth, demonstrating the viability and usefulness of our approach.

In Greece, there have been several incidents reported that indicate ongoing issues of Internet censorship. Starting as early as October 2006, the administrator of a Greek blog Rich Site Summary (RSS) aggregator service blogme.gr has been sued, arrested and jailed for hosting a link via an RSS feed from a blog post containing allegedly offending content [146]. As a consequence, the server of blogme.gr was shut down, the hard drives and the computer systems used by the server administrator were confiscated even though the service provided by blogme.gr was unrelated in any way other than linking the offending blog post via automatic RSS syndication.

In February 2010 ISP Tellas/Wind Hellas blocked the Piratebay site [249]. In May 2012 the Greek Organization for Intellectual Property Collective Administration went to court against every Greek ISP demanding to censor Ellinadiko.com, a music sharing forum and Music-Bazaar.com, an MP3 webstore, both under the accusation of infringing copyright laws [215]. The court ordered the ISPs to block the IP addresses of the referred websites [65].

Later in September 2012 a citizen was arrested on charges of malicious blasphemy and religious insult after posting a Facebook page that ridiculed a well-known Greek Orthodox monk [243]. Following in January 2013, a politician filed a defamation lawsuit against a Greek Wikipedia user and administrator, insisting to remove content hosted on a Greek Wikipedia page related to his name. Nonetheless he sued the Greek Free / Open Source Software society that he mistakenly believed to be the organization running the Wikipedia project [210].

In this article we will focus on censorship implied by content regulation policies, and particularly an anti-gambling policy in article 52, law No. 4002/2011 [123]. So far, there have been only very few and selective data reports available to conduct research in this field in terms of studies regarding the type of censorship taking place and the techniques used in order to observe the criteria set by censors. Our article will try to close this research gap by systematic empirical measurements across multiple ISPs. The selected set of ISPs account for the majority of the fixed and wireless broadband customers [150], [261].

The rest of the article is structured as follows. First we present related work in Section 2.2. Then, in Section 2.3, we present our methodology, the infrastructure and tools used to conduct our censorship research. Following we provide an analysis of the collected set of data per blocking method and ISP in Section 2.4. Continuing in Section 2.5 we analyze the blocklist used to conduct the blocking of the

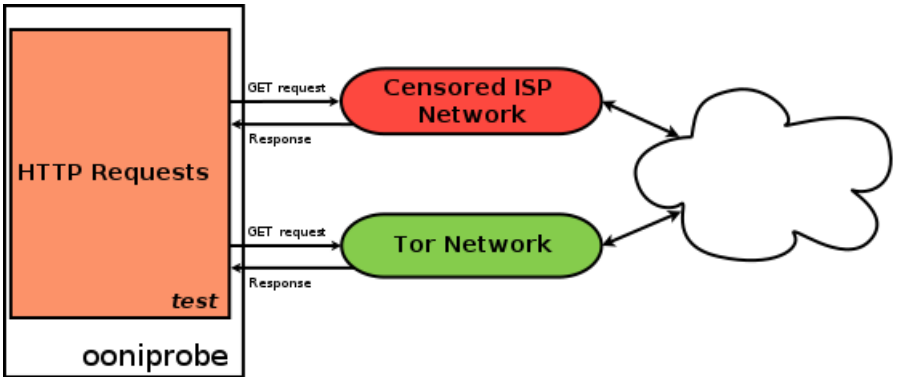

Figure 2.1: Ooniprobe HTTP request test diagram

resources, we reveal how the ISPs "broke" the email communication with these websites. Conclusions are given in Section 2.7.

2.2. Related work

So far to the authors' best knowledge there is minimal technical literature based on Internet censorship in the western world. Breindl et al [39] examine the debates surrounding network filtering in France and Germany, focusing on the arguments used by opponents and proponents of Internet blocking. The authors analyse the outcomes and, the various challenges posed by Internet blocking to democracy. Aase et al [2] collected measurements and social sciences aspects from three different contexts; public wireless networks in USA and microblogging and chat programs in China. By this comparison they attempt to illustrate the importance of the elements of motivation, resources and time in Internet censorship.

Furthermore we are providing a brief description of some censorship related research located outside the western world which we consider relevant as it provides a technical perspective related to our research focus. Wright et al [287] examine the problem of Internet censorship from a user perspective rather than on a national level. In their paper they discuss the possibility to detect the effects of Internet filtering thought different providers and services. In a recent article [117] Geddes et al. presented the arising issues when using covert communication channels to circumvent Internet censorship. In his revision of an anti-censorship technologies taxonomy, Leberknight et al. discusses the challenges and opportunities of censorship resistant systems [167]. An in-depth analysis by Nabi [13] provides a time-line (starting from 2006) on the implemented censorship activities in Pakistan. He analyses a variety of technical methods being used and provides some trivial censorship circumvention options. In the same context Arya et al [151] examined Iranian censoring techniques and tools applied. Their work consists of a topographical map that enumerates the categories of the censored websites. Further contributions regarding Internet censorship include reports that have been gathered by the OONI project on several countries, including Zambia [290], Palestina [126], USA [244], Uzbekistan and Turkmenistan [245].

2.3. Methodology

In our censorship research we have used a variety of common Free and Open Source Software networking tools for gathering, categorizing, distributing, analyzing data and comparing the results. Since acquiring results from many ISPs is crucial to form a representative sample, we have probed five ISPs offering landline broadband connections and three cellular mobile operators during our research. An analysis can

```
Certain censorship: 300
Certain censorship (single requests): 57
Possible censor mistakes (404): 72
Total Censored (Certain + Single + Mistakes): 429
----
Total Single responses: 65
Single responses over Tor (exclude from stats): 0
Control failure: 18
```

Listing 2.1: Reports Parser Output

be found in Section 2.4

2.3.1 Set of data used for the tests

We have used the blocklists issued by the Hellenic Gaming Commission (EEEP) (see Section 2.5.3), titled:*"List of sites providing gaming and betting services without authorization in Greece"*. The newly formed Independent Administrative Authority's Hellenic Gaming Commission acts as the public body responsible for the control and supervision of gambling services in Greece and publishes a blocklist of websites that are offering unlicensed gambling and betting services to Greek Internet users. According to article 52, law No. 4002/2011, A180 [123] gambling and betting websites without a license specifically for the Greek market is a serious criminal offense for users (players) that interact with or via these websites, but also for companies that run gambling businesses and ISPs that allow users to access the unlicensed websites.

The blocklist was transmitted to 211 ISPs, credit institutions and competent prosecuting authorities (misdemeanors Athens prosecutor, financial and e-crime prosecution department of financial police). According to [87] each entry in the blocklist has been revised by EEEP and was checked by two different echelons, through checks carried out on two different days and at different times-of-day [87].

2.3.2 Collection of censorship analysis reports

The collection of the network measurements took place during the months of June and August 2014. For our censorship research we used ooniprobe, an application developed by the OONI project [105] used by users and organizations to probe their network for signs of network tampering, surveillance or censorship. Developed with the idea of ensuring the detection of any interference with network communications, it aims to collect and provide high quality reports by using open and transparent data methodologies freely available to anyone that would like to process, read and research using a standard common file format YAML.

Ooniprobe is the application being used to conduct the measurements on the ISP networks (both landline and cellular networks) where we detected network tampering and content blocking (see Section 2.4). Ooniprobe provides a variety of test cases and classes that could be used to probe the networks. In our research we have deployed the Hypertext Transfer Protocol (HTTP) Requests test [202]. This performs an HTTP GET request from a specified list of URLs containing potentially censored websites over the test network (censored ISP network) and over the Tor network; the process is illustrated in Figure 2.1. It then compares the response headers and checks if the two responses (the one over Tor and the one over the censored network) match and if the proportion of differences between the expected body lengths is under

a specific threshold. In our test cases we have used a tolerance factor of 80 percent between the two body lengths. The page body length difference was proven to be the most effective similarity indicator [158]. Moreover, in order to avoid false positives, results were further analysed by hand. See Section 2.3.3 for a summary of an ooniprobe test.

2.3.3 Reproducing the results with ooniprobe

We first install ooniprobe following the install instructions for our system [148]. We have used the command in Listing 2.2 to generate our reports the *file* parameter contains a formatted list of the EEEP blocklist [206]. The parameter "file" denotes the list of URLs to perform GET and POST requests. We have used a text file (EEEP_Blacklist.txt) with the blocklisted entries published from EEEP [206].

```
ooniprobe blocking/http_requests --file EEEP_Blacklist.txt
```

Listing 2.2: Running ooniprobe HTTP requests test

For the collection and storage of ooniprobe reports we have used remotely managed Raspberry Pi embedded devices. The small footprint, minimal cost and power requirements make Raspberry Pi an ideal candidate for distributing it across individuals and organizations, who would like to contribute results by probing their network for instances of censorship. We have implemented Lepidopter; a custom boot image [266] based on Debian GNU/Linux offering a ready to boot image that eases the installation, configuration, as well as management and assist in the execution of ooniprobe network measurements tests. Apart from our network measurements Lepidopter aims to increase the coverage of censored networks around the world. The distribution image with the source code is freely available under the GNU general public license version 3 [266]. Our network measurements have been collected from multiple probes and locations under the same ISPs. Specifically, we have used at least two different probes per ISP and conduct measurements on different days and time frames to provide consistent reports. Initially, we have encountered inconsistent results between same ISPs on different connections. The cause of these inconsistencies were due to the Domain Name System (DNS) resolvers being used, since many of the ISPs use DNS Hijacking to enforce the blocking implied by EEEP. Additionally, we have identified that some of the ISPs were blocking the TCP requests to port 443, which resulted to a network timeout when a probe (or a user) was trying to access any URL entries from the EEEP blocklist. Note that the users should be receiving a block page (see Figure 2.2) that provides information on why this URL (or domain name) is blocked rather than a network timeout that could confuse the users to think that there is some network or server failure.

```
dig +short A www.netbet.com @213.249.17.10
Response: <@\textcolor{red}{213.249.29.111}@>

dig +short A netbet.com @213.249.17.10
Response: <@\textcolor{red}{(no answer)}@>

dig +short A sport.netbet.com @213.249.17.10
Response: <@\textcolor{red}{(no answer)}@>
```

Listing 2.3: DNS responses of netbet.com entries

The size of each result set, over 10 MiB, makes it very impractical to easily extract results. To be able to extract meaningful information from ooniprobe's YAML output we have created a sample parser in python. The parser looks for a number of criteria in the headers or body of the test results from ooniprobe

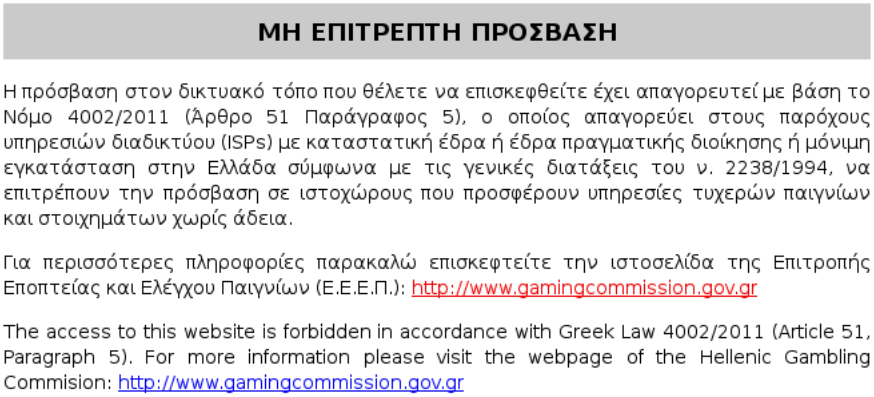

Figure 2.2: Blocked Webpage Screenshot

and if there's a match, it categorizes the test as being censored or not. Among these criteria there is a check for *headers_match* parameter of ooniprobe tests, a check for ISP redirect URLs in response headers and a search for *gamingcommission.gov.gr* in response body. The parser also catches some ooniprobe test failures, for example not being able to fetch results over Tor. Finally, it displays a summary of the parser output (see Listing 2.1).

In this section we are reviewing and processing the results taken from eight major Greek ISPs, five of them offering standard landline services (Cyta, Hol, Forthnet, Ote, Wind) while the other three ISPs are offering mobile services (Cosmote, Vodafone, Wind).

2.4. Analysis of blocking per ISP

Most Greek ISPs have not issued any public report that notifies their customers and users about the content blocking of the EEEP blocklist. On 2 August 2014, we contacted all ISPs via email communication, mentioning our research and the related network measurements. We have inquired the ISPs for clarification about their filtering policies, specifically how they renew the blocklist, and by which technical means and implementations the blocking of the content takes places. Additionally, we have asked them to inform us how they communicate the content blocking with their customers. Finally, we have sent a request for comments regarding the blocking regulation imposed by EEEP, how the process could be improved and if they are forced to block other content upon request or based on another blocklist. Out of the eight ISPs representatives and support teams that we contacted, only one (Cyta) replied and directed us to the EEEP website [88] which was irrelevant to our specific inquiries. In our reply we repeatedly

ISP	Blocked Entries	Blocking Method	Server Fingerprint	Overblocking
Cosmote	438 (100%)	DNS Hijacking	Apache/2.2.15 (CentOS)	✗
Cyta	357 (81.5%)	DNS Hijacking, HTTP 404	Apache	✓
Forthnet	96 (21.91%)	DNS Hijacking	BigIP	✓
Hol	438 (100%)	DNS Hijacking	lighttpd/1.4.31	✓
Ote	438 (100%)	DNS Hijacking	Apache/2.2.15	✓
Wind	325 (74.2%)	DNS Hijacking, HTTP 404	Tellas HTTP Server	✗
Wind Mobile	325 (74.2%)	DNS Hijacking, HTTP 404	Tellas HTTP Server	✗
Vodafone	425 (97.03%)	DNS Hijacking, DPI	WebProxy/6.0	✗

Table 2.3: Per ISP list of server fingerprints, overblocking indication, blocked entries and methods

pointed to our inquiries but as of the date of this article submission no further email communication was received.

2.4.1 Blocking methods

DNS hijacking

ISPs are in control of the DNS servers being used by their clients' xDSL routers. Since they can manipulate their DNS servers' responses, they can redirect the requesting clients to anywhere they want. Taking advantage of this privilege, ISPs modify their resolvers to override censored domains' legitimate DNS replies by creating local zone entries [29]. These entries usually point to a server that they control where they run a web server that displays a censorship warning message to the users.

Deep Packet Inspection

Deep Packet Inspection (DPI) is the basis of the most advanced form of censorship. Special appliances have the ability not only to look into Layer 3 and Layer 4 headers but to also look inside the payload of each and every packet. They can distinguish packets going to a server and either stop them from reaching their target, change the server's response, or even redirect the packets to another server. These devices perform a hostile, active, man-in-the-middle attack on every client connecting to the network, Internet or Intranet, through them.

Vodafone DPI In this section we demonstrate the case of DPI filtering found during our analysis of Vodafone ISP. We use the curl program to fetch the HTTP headers of the domain *rivernilecasino.net*. The output of curl suggests that the request is not blocked and transmitted to the legitimate server of the domain. The complete response is listed in Listing 2.4. Similarly, when changing the URL to *www.rivernilecasino.net* (adding the *www* subdomain), the request also passes through to the legitimate server. The response is listed in Listing 2.5. Finally, using the same URL as the one published in the blocklist (*www.rivernilecasino.net/index.asp*), gets the request proxied via Vodafone ISP. The HTTP headers of the response are listed in Listing 2.6.

```
HTTP/1.1 302 Moved Temporarily
Date: Sun, 31 Aug 2014 12:38:01 GMT
Server: Microsoft-IIS/6.0
X-Powered-By: ASP.NET
X-AspNet-Version: 4.0.30319
Location: http://www.vegaspartnerlounge.com/generic/informer.asp?Subgid=987228&
    ↪ Country=Greece&btag=rivernilecasino.net&btag2=16&btag3=&btag4=&btag5=
Set-Cookie: RiverNileCasino=btag=rivernilecasino.net&btag2=16&btag3=&btag4=&btag5=;
    ↪ domain=rivernilecasino.net; expires=Mon, 01-Sep-2014 12:38:00 GMT; path=/;
    ↪ HttpOnly
Cache-Control: private
Content-Type: text/html; charset=utf-8
Content-Length: 279
Connection: keep-alive
```

Listing 2.4: Vodafone ISP: HTTP headers request to 'rivernilecasino.net'

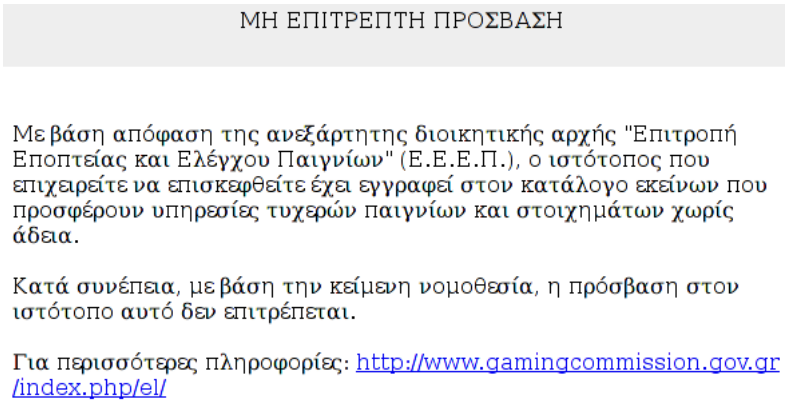

Figure 2.4: Cosmote and Ote EEEP blocked website screenshot

```
HTTP/1.1 302 Moved Temporarily
Date: Sun, 31 Aug 2014 12:38:28 GMT
Server: Microsoft-IIS/6.0
X-Powered-By: ASP.NET
X-AspNet-Version: 4.0.30319
Location: http://www.vegaspartnerlounge.com/generic/informer.asp?Subgid=785274&
    ↪ Country=Greece&btag=www.rivernilecasino.net&btag2=16&btag3=&btag4=&btag5=
Set-Cookie: RiverNileCasino=btag=www.rivernilecasino.net&btag2=16&btag3=&btag4=&
    ↪ btag5=; domain=rivernilecasino.net; expires=Mon, 01-Sep-2014 12:38:28 GMT;
    ↪ path=/; HttpOnly
Cache-Control: private
Content-Type: text/html; charset=utf-8
Content-Length: 283
Connection: keep-alive
```

Listing 2.5: Vodafone ISP: HTTP headers request to 'www.rivernilecasino.net'

```
HTTP/1.1 301 Moved Permanently
Server: WebProxy/6.0
Date: Sun, 31 Aug 2014 12:39:01 GMT
Content-Length: 0
Location: http://1.2.3.50/ups/no_access_gambling.htm
Connection: keep-alive
```

Listing 2.6: Vodafone ISP: HTTP headers request to 'www.rivernilecasino.net/index.asp'

2.4.2 ISP analysis

Cosmote

Cosmote ISP (AS29247) has blocked all the entries of the published blocklist. The HTTP headers included in the hijacked response indicate the presence of an Apache/2.2.15 web server running on a Linux based OS (CentOS). The implied method of blocking is DNS hijacking and the indication of censorship includes the URL string *http://www.gamingcommission.gov.gr/index.php/el/* in the body of the HTTP response.

Cyta

Cyta ISP (AS6866) has blocked 81.5% entries of the published blocklist. The collected HTTP headers indicate the use of the Apache web server. The method of blocking used is DNS hijacking and 404 HTTP errors. During our network measurements, we detected at least 80 blocklist entries incorrectly responded with a 404 HTTP error, *"Not Found"* instead of displaying the filtering warning page. Even though this is probably because of a misconfiguration of the ISPs' web server hosting the warning page, this technique (the fake 404 HTTP error) results in the user not knowing the reason why the resource is inaccessible and is therefore a transparency issue. Such censorship technique has reportedly been used in many countries already [191], [129]. The HTTP response body contains the URL string *http://www.gamingcommission.gov.gr/index.php/el/* but only for the entries that return a blocking page. We have been able to identify the fake 404 HTTP errors by comparing the results with other ISPs. Cyta introduces a Google Analytics script on their blocking webpage [208] which can be used to track users that have tried to access the blocked content. Although the blocking page is hosted on the same website (Cyta main website) the user tracking application differs.

Forthnet

Forthnet ISP (AS1241) has blocked 21.91% of the published blocklist. This is the lowest percentage of blocked URLs among all ISPs, and this fact is quite known to the Greek gambling community who advise users that experience blocking to switch to this ISP. The number of subscribers of this ISP has been steadily increasing since 2013 [152], bursting the total subscriptions to a historical record of 1,145,948 [214]. The server fingerprint collected by the HTTP headers indicate the use of Big IP as part of the network filtering infrastructure. Upon receiving the hijacked DNS response the user is being redirected (HTTP 302) to the blocked page URL *http://eeep.forthnetgroup.gr*.

Hol

HOL ISP (AS3329) has blocked all entries of the published blocklist. The HTTP headers collected state the use of a lighttpd web server software (with build version 1.4.31). This ISP returns a fake DNS response that redirects (HTTP 301 code) the user to the blocking page with URL string *http://eeepnotice.hol.gr/*.

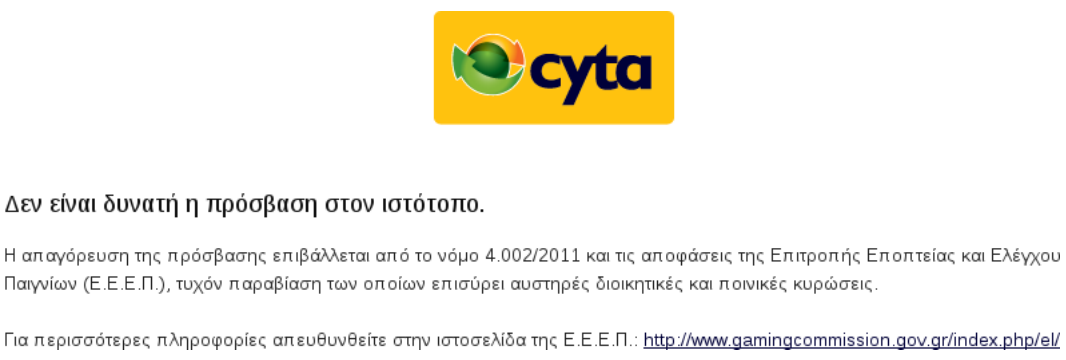

Figure 2.5: Cyta EEEP Blocked website screenshot

Figure 2.6: HOL EEEP Blocked website screenshot

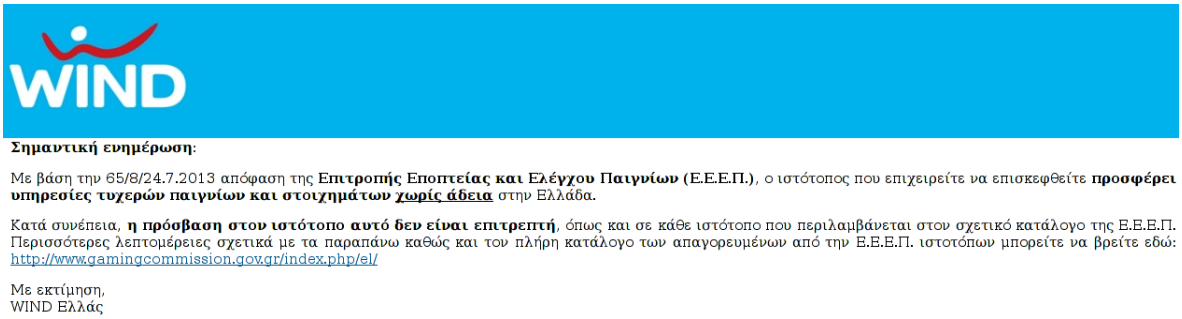

Figure 2.7: Wind and Wind Mobile EEEP blocked website screenshot

Ote

OTE ISP (AS6799) has blocked all the entries of the published blocklist. It is the only ISP that takes care to preserve the DNS Mail Exchanger (MX) records for some of the filtered domains. The HTTP headers (Apache/2.2.15 (CentOS)) and the URL string returned on the response HTTP body (http://www.gamingcommission.gov.gr/index.php/el/) extracted from our measurements are identical with the Cosmote ISP, which is part of the same company group.

Wind

Wind (AS25472) and **Wind mobile** (AS15617) ISPs (same company group) have blocked 74.2% of the published blocklist. The HTTP headers include the string *Tellas HTTP Server*. Apart from the DNS hijacking blocking method, we detected that at least 65 entries erroneously responded with a 404 HTTP error. Four other entries of the blocklist were redirecting to a page of the Wind ISP website with the HTTP body string *"landline services provided by Wind have been suspended"*. These ISPs mislead users by not providing them with the block page that informs about the gambling law.

Vodafone

Vodafone ISP (AS12361) uses DNS hijacking for 58 entries that EEEP has published either using Hypertext Transfer Protocol Secure (HTTPS) URLs on the blocklist or entries that are not prefixed with *http://* (the blocklist entries can be found in [206]. For the rest of them it uses some kind of DPI/proxy, using Bluecoat's WebProxy/6.0. If an HTTP URL does not exactly match the one published at the blocklist it is passed on to the original server, if it matches then the request gets redirected to *http://1.2.3.50/up-*

s/no_access_gambling.htm. The process of determining the DPI blocking by Vodafone is presented in Section 2.4.1.

Another interesting case with this ISP is that for domains that it has filtered using DNS hijacking, subdomains of those do not even have an A record (dns_lookup_error). That means that some URLs on the blocklist that contain subdomains are not getting redirected to *http://1.2.3.50/ups/no_access_gambling.htm*, they cannot be resolved and are not accessible at all from clients. We use the DNS lookup utility DiG and queried the A DNS records for the resources *www.netbet.com*, *netbet.com* and *sport.netbet.com* using Vodafone Name Server (NS). Out of the three queries only one (*www.netbet.com*) provided a response which pointed to the host of Vodafone used for blocking. The DNS lookup queries are listed in Listing 2.3. Vodafone ISP has blocked 425 entries of the published EEEP blocklist, 15 of them returned DNS lookup errors.

In summary Figures 2.2 and 2.4 to 2.8 illustrate the landing blockpages of each ISP.

2.4.3 Collateral damage

In an announcement [86] published by the commission addressing the public about the blocklist, they reply to the question: *"How will (gambling) players be able to contact the companies since they are now blocked?"* and their response is: *"By advising people to look at their previous bank transactions where contact details of these companies might exist"*. If the ISPs were actually blocking URLs, then emails towards the companies would still work, but because ISPs interfere and manipulate the DNS records of the blocked sites, most of them do not even pay attention to the DNS MX records of the gambling companies domains, and a user cannot email them any more since an MX record points to the appropriate mail server and specifies how email delivery should be routed for a given domain name.

In the process to deliver an email, an Simple Mail Transfer Protocol (SMTP) client will first query the destination domain for an MX record and if no record is found, it will fall back to look up an A record (or AAAA record if Internet Protocol version 6 (IPv6) is available) for the domain in question and attempt to deliver email based on these records. Apparently, without any MX records or legitimate A records for the blocklisted domains in question email delivery would be impossible. During our research, we found out that most ISPs have hijacked the MX records of the domains included in the EEEP blocklist.

In our case the ISPs have spoofed the A records of the gambling domains to point to a local server or

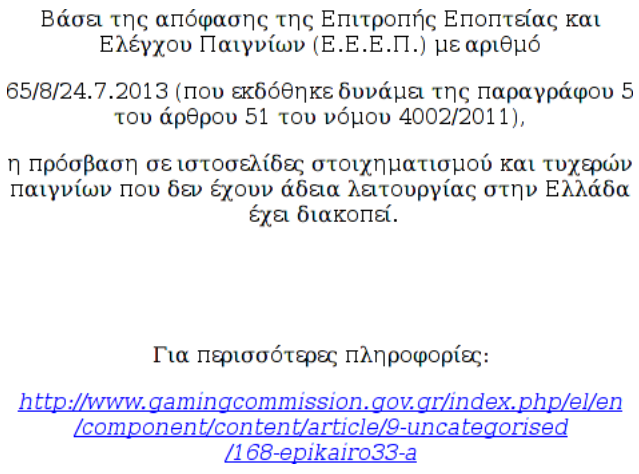

Figure 2.8: Vodafone Blocked Webpage Screenshot

ISP	DNS QUERY	ANSWER	RESPONSE
Cosmote	MX	NO	None
	A	YES	Hijacked
Cyta	MX	NO	None
	A	YES	Hijacked
Forthnet	MX	NO	None
	A	YES	Hijacked
Hol	A	Yes	Hijacked
Ote	MX	NO	None
	A	YES	Hijacked
Vodafone	MX	NO	None
	A	YES	Hijacked
Wind	MX	NO	None
	A	YES	Hijacked
Wind Mobile	MX	NO	None
	A	YES	Hijacked

Table 2.9: DNS MX and A records responses ISP summary table

a proxy server. As a result, any email delivery will fail and the user will only realize this after hours or even days (depending on their SMTP server configuration). This implies tremendous negative impacts and leads to a restriction of fair markets and business regulations, i.e., a user trying to communicate with any business (all of the censored websites are businesses) will find himself unable to do so. Out of eight ISPs only one (Ote ISP) found to sync the MX records from some (but not all) of the blocklisted domains. However, it remains unclear if and how often these records are updated. Table 2.9 summarizes the DNS MX and A record responses returned per probed ISP.

2.4.4 DNS MX record responses

In this section we manually inspect the DNS MX of two sample domain names, *770.com* and *880.com* that return incorrect or empty records. The non manipulated MX records of the domain names in question are illustrated in Listing 2.7.

```
; <<>> DiG 9.8.4-rpz2+rl005.12-P1 <<>> MX 770.com @8.8.8.8
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 764
;; flags: qr rd ra; QUERY: 1, ANSWER: 5, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;770.com.                IN   MX

;; ANSWER SECTION:
770.com.          2473    IN   MX   20 alt1.aspmx.l.google.com.
770.com.          2473    IN   MX   30 aspmx2.googlemail.com.
770.com.          2473    IN   MX   20 alt2.aspmx.l.google.com.
770.com.          2473    IN   MX   30 aspmx3.googlemail.com.
770.com.          2473    IN   MX   10 aspmx.l.google.com.
```

```
; <<>> DiG 9.8.4-rpz2+rl005.12-P1 <<>> MX 880.com @8.8.8.8
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 17952
;; flags: qr rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;880.com.                IN  MX

;; ANSWER SECTION:
880.com.         2753    IN  MX  0 smtp.secureserver.net.
880.com.         2753    IN  MX  10 mailstore1.secureserver.net.
```

Listing 2.7: Google public DNS MX record responses for the domains: 770.com and 880.com

Ote

OTE ISP provides a correct reply for the domain *888.com* but a bad one for the domain *770.com* as illustrated in Listing 2.8.

```
; <<>> DiG 9.8.4-rpz2+rl005.12-P1 <<>> 770.com
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 25114
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 0

;; QUESTION SECTION:
;770.com.                IN  MX

;; AUTHORITY SECTION:
770.com.         86400   IN  SOA localhost. hostmaster.localhost. 2014030701 10800
    ↪ 3600 1814400 86400

; <<>> DiG 9.8.4-rpz2+rl005.12-P1 <<>> 888.com
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 54428
;; flags: qr rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;888.com.                IN  MX

;; ANSWER SECTION:
888.com.         85914   IN  MX  5 smtplo.888holdings.com.
888.com.         85914   IN  MX  9 smtp.888holdings.com.
```

Listing 2.8: OTE DNS MX record responses for the domains: 770.com and 880.com

The remaining hijacked MX record responses for the remaining of the ISPs are illustrated in Listings 2.9 to 2.13.

```
; <<>> DiG 9.9.5-4-Debian <<>> 770.com
```

```
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 64761
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;770.com.                       IN      MX

;; AUTHORITY SECTION:
770.com.                900     IN      SOA     nyma_grid_dns1.dcn.cosmote.gr.
    ↪ please_set_email.absolutely.nowhere. 4 10800 3600 2419200 900
```

Listing 2.9: OTE DNS MX record responses for the domains: 770.com and 880.com

```
; <<>> DiG 9.9.5-4-Debian <<>> 770.com
;; Got answer:
;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;770.com.                       IN      MX

;; AUTHORITY SECTION:
770.com.                86400   IN      SOA     770.com. root.770.com. 42 10800 900
    ↪ 604800 86400
```

Listing 2.10: Cyta DNS MX record responses for the domains: 770.com and 880.com

```
; <<>> DiG 9.8.4-rpz2+rl005.12-P1 <<>> 770.com
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 19146
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;770.com.                       IN      MX
```

Listing 2.11: Cyta DNS MX record responses for the domains: 770.com and 880.com

```
; <<>> DiG 9.8.4-rpz2+rl005.12-P1 <<>> 770.com
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 8803
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 0

;; QUESTION SECTION:
;770.com.                       IN      MX

;; AUTHORITY SECTION:
770.com.                86400   IN      SOA     localhost. root.localhost.
    ↪ 2013091901 14400 7200 604800 86400
```

Listing 2.12: Wind Mobile DNS MX record responses for the domains: 770.com and 880.com

```
; <<>> DiG 9.9.5-4-Debian <<>> 888.com
```

```
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 50372
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;888.com.                       IN      MX

;; AUTHORITY SECTION:
888.com.                600     IN      SOA     ns0.panafonet.gr.  root.iapetos.
    ↪ panafonet.gr. 2014180301 28800 7200 1209600 600
```

Listing 2.13: Wind Mobile DNS MX record responses for the domains: 770.com and 880.com

2.5. Blocklist analysis

2.5.1 Blocklist distribution

The method that EEEP uses to distribute the blocklist to the ISPs remains unclear. Sources from ISPs claim that they have never received any updates apart from the very first time that the blocklist was communicated to them. At that time they were given instructions to visit the EEEP website [280] and manually download the blocklist. Since then, there have been three updates to the original blocklist released, all of them only published on EEEP's website in the form of Portable Document Format (PDF) files. One would have expected that the blocklist was at least in a machine parsable format to automate the procedure, but that is not currently the case. So far, there is no automated way to check for updates of the blocklist, the downloadable PDF changes filename each time there's an update, and there is no Application Programming Interface (API) to query for updates. The published PDF is not signed by any authority but at least EEEP's website is using HTTPS and is redirecting all HTTP requests to their HTTPS equivalent URLs.

2.5.2 Privacy and legal aspects

EEEP was originally created in 2004 by the Greek state as an administrative authority to control gambling and lottery games but was never activated until 2011. Law 4002/2011 transformed EEEP into an independent administrative authority [123] with capabilities to control and regulate the industry. This specific law gave EEEP the ability to publish blocklists and fine companies that do not comply with them. On January 2013 the Greek state published a public consultation regarding changes to the law that dictate the way EEEP operates. The only ISP that commented on that consultation was Ote.

Article 51, paragraph 5 of Greek law 4002/2011 states that ISPs who allow access to blocklisted websites can be fined by EEEP. Since most ISPs only filter DNS requests, and not the actual connections, it is an open question whether they will ever be fined by EEEP. Article 52, paragraph 3 of Greek law 4002/2011 [123] states that people participating in unregistered gambling activities can be jailed up to three months and fined from 5,000€ to 20,000€. This article poses an interesting question on how these people will be tracked and found. Paragraph 6 of the previous article [123] states though that 3rd parties who allow access (proxies) to unregistered gambling activities can be jailed up to 3 years and fined from 100,000€

31/07/2013	401 entries.
22/11/2013	423 entries.
22/02/2014	438 entries.
11/07/2014	437 entries.

Table 2.10: Timeline of EEEP blocklist publication

to 200,000€.

Even though ISPs are not legally bound to keep track of users trying to access unregistered gambling sites, it is unclear what really happens in practice. At least one ISP has added a Google analytics JavaScript to the website that the users are redirected instead of the original one. Other ISPs use HTTP Etags on their server responses that could be used instead of cookies to uniquely track users [115]. Since Etags and analytics applications are optional, it is unclear why ISPs have opted-in to use them.

2.5.3 EEEP blocklist analysis

Starting in July 2013 [87], EEEP published a blocklist [55] containing 401 entries of URLs that do not comply with the regulations of EEEP as analyzed in Section 2.3.1. Later in November 2013, there were 22 new entries added to the blocklist [56]. Following that, in February of 2014 there were 25 further entries added [57]. Finally, in July 2014 [265] there was one entry removed, summing up the blocklisted entries to 437. Our measurements tests took place between June and Augst 2014, using the third released version of the blocklist [57]. An overview of the EEEP blocklist timeline can be found in Section 2.5.3.

Upon the latest blocklist update [265] the entry: *http://www.pokerstarsblog.com/* was removed from the blocklist. We have probed again all ISPs to check if they have complied with the update of the EEEP blocklist, unfortunately many of them were still blocking the entry, the results (overblocking column) are listed in Table 2.3. The complete blocklisted entries can be found in [206].

Throughout the blocklist analysis we determined that the entries contain duplicate, malformed and unavailable entries. Malformed entries do not follow certain specifications such as URL canonicalization, contain spelling mistakes, and query parameters in URLs with reserved characters such as the strings `'#_'` and `'action='` that should be first encoded [36] prior to any distribution and publication of the blocklist.

- 28 entries (6.39%) are duplicate domains (with different URLs).
- 17 entries (3.88%) refer to pages or subdomains which are malformed.
- 3 entries (0.68%) are not hosting any gambling content (empty DNS A record, expired or parked domain names): *loosecannonholdem.com*, *unibet-1.com* and *venicegames.com*

While the Hellenic Gaming Commission has specifically asked for URL blocking, since it has posted a blocklist with URLs inside, ISPs would only be able to block them if they had previously installed some kind of DPI mechanism. That mechanism would give the ability to ISPs to look inside the payload of packets, and more specifically at the layer 7 contents where the actual URL of an HTTP request is referenced. A problem that arises though is what would have happened to the HTTPS URLs included in the blocklist. In

order to be able to filter HTTPS URLs one needs to actually perform an active man-in-the-middle attack on every HTTPS connection in order to decrypt the SSL/TLS, layer it and look at the unencrypted payload.

Instead of filtering using DPI, all but one (Vodafone) of the aforementioned ISPs make use of DNS hijacking to block access to the blocklisted domains. This censorship method provides a different IP address of the requested domain/site than the actual one, redirecting the user to some other destination. Each ISP has chosen to redirect users to a website of their own where they display a generic warning that the site is filtered.

We have discovered that the content blocking was sometimes implemented incorrectly and in many cases the users were not informed that a specific website adheres to the EEEP policy, leading them to assume that the website encountered a technical problem (HTTP 404 error or Connection timed out), which effectively obfuscates deliberate censorship. It is unclear how frequently the ISPs evaluate the effectiveness of their filtering rules, resulting in outdated and poorly implemented blocklists and in the surprising fact that none of them exactly complies with the EEEP policy. Unfortunately, without transparent methodologies and review on the blocking techniques it is quite difficult to be assured of the effectiveness of the filtering systems.

Furthermore, DNS hijacking can be quite easily circumvented and is highly ineffectual for blocking access to content, since the user could simply use a different DNS service. There are numerous issues introduced with DNS hijacking such as network security threats [29], phishing attacks [81] and privacy violation [282].

2.6. Circumventing censorship

Using a blocklist/allowlist model (sort of an ON/OFF model) is not granular enough and tends to fail; it will have negative impacts on users and customers of the ISPs that may or may not find routes around the blocking. Furthermore, creating such an ineffective blocking could give a false sense of security as the entities enforcing such censorship would assume that the content is being blocked although the blocking can be easily circumvented. One common practice to circumvent censorship implemented via DNS hijacking is to use an external DNS resolver or a different DNS server known to provide the legitimate DNS records of the blocklisted domain entries.

2.6.1 Using alternative DNS resolver

The DNS hijacking enforced by ISPs in Section 2.4 can be circumvented with a trivial network configuration and requires minimal technical expertise by using an alternative DNS resolver that does not manipulate DNS records. However this has a negative impact on users' privacy as the they are exposed to the global scale metadata collection that is currently happening on the Internet. Furthermore, using name servers that are geographically located at a greater distance than the user's network location degrades the Internet experience of the user as every DNS request will be significantly delayed. A more technical savvy user could set up a NS with the legitimate DNS records to circumvent censorship.

2.6.2 Further methods

Low-latency anonymity networks like the Tor Project [281], offer another way to access filtered sites and circumvent censorship. Tor is known to thrive in countries where governments deploy censorship to

block users from accessing parts of the Internet. The global scale of Tor is ideal for this kind of situation. Additionally pluggable transports can be used to circumvent sophisticated censorship that blocks the Tor traffic flow. Another options are Virtual Private Network (VPN) providers that offer their services in order to help users circumvent filters and censorship. Many of these services are offered free of charge or with a minimal fee.

2.7. Conclusion

Throughout our research we highlighted flaws in the implementation of betting website censorship. Nevertheless, we do not aim for a properly implemented censorship of any kind. Instead, our intention is to make all those problematic issues visible, which arise when censorship is invoked as a method to approach a social or public issue.

As of the publication of this research, no ISP has published any information page regarding censorship or filtering that is conducted by their side. There is no reference to any court orders or laws informing their users on which IPs or websites ISPs have to censor and how. Information is only presented upon visiting one of the blocked websites, but still that page does not provide either the full list of blocked websites nor the filtering methods in use. We found one ISP that has published press releases [153] declaring that they will be filtering websites from now on according to law 4002/2011.

ISPs in Greece have not provided any kind of notification to their customers informing them how the blocking took place, why this happened and if they can opt-out from the service. Lack of transparency on behalf of the providers permits them to block arbitrary Internet destinations according to their needs, thus following a blocking-at-will strategy. Internet destinations may be accessible or not, while users have no reasoning about it. That would lead to deliberate abuse of citizens' accessibility and view of the Internet.

Censorship of some betting sites in Greece was implemented as a way to forbid residents of Greece to bet on websites that do not intend to pay taxes, whereas the claimed goal of the censors is to prevent players from betting. As examined in Section 2.5, censorship implementations includes some major side-effects: users are not only forbidden to play on these websites, but they are also now unable to communicate (via email) with these companies. Censorship is thus not limited to a specific problem (tax evasion) but it is massively affecting user experience and communication.

2.8. Future work

Based in our methodology described in Section 2.3.2 performing network measurements daily, weekly and upon renewal of the EEEP blocklist per ISP significantly improves the contribution of our research study. Furthermore, we consider our research study of gambling censorship a low risk, but high impact network measurement that could be applied to different countries ISPs.

– 3 –
Internet censorship analysis in Cyprus

3.1. Introduction

This paper describes the initial findings of an open and collective effort to gather data, using OONI and open DNS resolvers in Cyprus, towards a cross comparison study of web content blocking regulations and practices between Cyprus and other countries in terms of implementation techniques. We suggest there is a need for a closer study of how censorship (the blocking of content, the top-down imposition of restrictions on information) is legislated and justified in political terms on the one hand, and on the other hand the actual extent and the procedural technicalities of its implementation as experienced by the citizen, in this case the Internet user or the ISP client. This investigation of *how* blocking is legislated and implemented on a local level contributes to discussions around transparency, accountability, and freedom of expression more broadly. The island of Cyprus presents an interesting geopolitical case study because it allows for the collection of data on what we have come to think of as more than two distinct regimes in terms of information policy: the one followed in the Republic of Cyprus (RoC) in the south of the island, which largely adopts EU policy, and the one followed in the area occupied by Turkey in the north of the island. The landscape regarding policy over Internet blocking may prove to be even more complex, considering the existence of two British sovereign military bases on the island, although our study does not yet include data from these areas. Our initial measurements are biased towards Internet blocking by ISPs following RoC protocols, with fewer observations revealing the policy of Internet blocking in the north (only one north Cyprus ISP, Multimax, is measured).

Our intention is to gather data on the capabilities of ISPs to perform censorship, or more specifically their capabilities to block access to specific information in Cyprus, and to provide comparable data about how the application of technologies for censorship, or control over information, is developing internationally.

The rest of the article is structured as follows. First we introduce the case of Cyprus and the specific legal circumstances around online gambling that allow us to investigate Internet blocking on the level of the ISP. We then briefly indicate similar research done in other countries, and present our methodology, the infrastructure and the tools we used. Following, we provide an analysis of the collected data set per blocking method and ISP and analyze the blocklist used to conduct blocking, its effects and collateral damage. We conclude with an outlook on how this kind of research might be used in the future.

3.2. The case of Cyprus

For the case of Cyprus we collected measurements from end-user connections located on various ISPs on both sides of the island. Cyprus has a population of 1,1 million. In comparison to other countries, access to Internet services is very good, as shown by the 2016 statistics of the ITU information society report: 71.2% penetration of Internet access in Cyprus. The share of fixed-broadband subscriptions of residents lies at 22.3%, with an additional 54.8% having active mobile broadband subscriptions. The average Internet

bandwidth per Internet user is measured at 89,791 Bit/s in 2016 [156]. This gives us a better understanding about user experience and allows for evaluating how each ISP has implemented the updated betting act directives. We investigate the extent to which ISPs may have over-blocked or under-blocked any entries included or deduced in the blocklist, and we analyze any collateral damages to unregulated websites. In recent years, ISPs in the RoC have implemented an Internet filtering infrastructure to comply with the laws and regulations imposed by the National Betting Authority (NBA). Our starting point was to find out how the technical infrastructure to block or filter unregulated web resources (the ones implied by the NBA) has taken place and discover cases of under or over blocking and to find collateral damage caused by blocking Internet resources that were not meant to be blocked (such as email).

3.3. Previous research

The RoC is considered a safe haven for freedom of speech. It is important to note that Freedom House reports that mention and catalog Internet censorship related events in the years 2006 [137], 2007 [138], 2008 [139], 2011 [140], 2012 [141], 2013 [142], 2014 [143] document that citizens are able to access the Internet on a regular basis and are not subject to any known government restrictions, although they do report a difference between the years 2012 and 2013. However, Freedom House numerical rating reports for Cyprus are based on conditions on the south of the island only. Worth mentioning is research on media pluralism that considers risks to freedom of expression and right to information in Cyprus as low risk [45]. We have not been able to find any previous work that discusses Internet censorship in Cyprus, and there has been no attempt to compare information across the island's divisions.

This case study on Cyprus is related to two previous OONI case studies. In the first instance we refer to previous research on large scale content blocking in Greece [275]. Similarly with the NBA in Cyprus, in Greece this kind of blocking is initiated by the EEEP, an independent administrative authority that acts as the public body, responsible for the control and supervision of gambling services. The Greek case-study analyzed the techniques and policies used to block content of gambling websites in Greece and presented the implications of under-blocking, over-blocking, and collateral damage by blocked email communication. It also highlighted issues of transparency in Internet filtering and unfair competition between ISPs. In the second instance we refer to a case study in Turkey that attempted to track changes to Internet traffic during the coup d'etat of July 2016. The study brings up the technical aspects of potential Internet blocking in Turkey and highlights the importance of a grassroots understanding of ISP blocking capabilities [84].

3.4. Detecting network interference and the Republic of Cyprus gambling law of 2012

Identifying signs or conclusive results of network interference that can be caused by Internet filtering or surveillance is a challenging process that requires adequate knowledge of the underlying network infrastructure on the side of the ISPs or their upstream providers. In this article, we focus on censorship by content regulation policies, and particularly the gambling law of 2012, L. 106(I)/2012 [66]. The law implies that the ISPs are obliged to apply a *blocking system* that will prevent users and ISP clients from accessing gambling services providers who are not licensed (do not hold a Class B license) or service providers who possess, operate infrastructure or provide online casino services in Cyprus. According to the NBA, a *blocking system* is defined as:

> A system installed by the Internet service provider which prevents the routing and the movement from the terminal equipment of the Internet user to particular Internet website addresses URL.

According to the RoC gambling law of 2012, non compliance is punishable with a term of imprisonment not exceeding five years or a fine not exceeding three hundred thousand Euro or to both such sentences. Upon notification from the NBA, ISPs are obliged to block URLs of gambling services that do not follow regulations within seventy two hours. Although the law does not specify the way in which URLs should be submitted to the ISPs for blocking, the current means seem to be a publicly available blocklist; a file with a list of URL entries named as *Blocking List* [77], located on the official website of the RoC NBA [78].

3.4.1 Analysis of the Republic of Cyprus NBA blocklist

NBA publishes a blocklist usually in a text file format that contains a number of URL entries of websites with complete file paths, not just domain names (such as `http://m.downloadatoz.com/apps/com.microgenius.casino777,482188.html`) that offer non-licensed gambling services in Cyprus. NBA was established in 2012 as an independent authority, consisting of a president and six members. One of the authority's duties is to notify ISPs in an electronic manner on every Internet URL through which gambling services are offered that are not covered by a class A or B licensed bookmaker, or anyone offering services prohibited in the present gambling law [66]. Although the law was issued in 2012, the first public release of the blocklist (that we were able to detect from the online archives) was in February, 2013 [67]. NBA does not provide a blocklist versioning system similar to other countries [275]. We assume 10 blocklist versions from February 2013 to May 2017 [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], though we cannot with certainty confirm the existence of additional blocklists in the past. Our findings are derived from Internet archives [27], [25] that provide historical snapshots of websites. Starting in February 2013, the NBA publishes a blocklist containing 95 entries of URLs [67] that increases to a total of 2563 (in April 2017) URL entries [76], approximately 27 times more than the initial size of the blocklist. Figure 3.1 illustrates a timeline with the date and URL entries of the blocklist published by the NBA.

During our analysis of the blocklist, we identified a number of malformed entries (mainly URLs and domain names) such as *1xbet.??* as well as duplicate entries and at least one entry that does not seem to host gambling related content; *https://www.commission.bz*, an advertisement affiliate program. The malformed URL entries of the blocklist may introduce technical issues to the filtering implementation of the blocklist as URLs that contain malformed characters (such as *??*) may not be parsed correctly. Additionally, a number of domain names in the blocklist were found to be expired or not registered, meaning that these domain names are not hosting any gambling related content (actually not hosting any content since they are not registered) but are still blocked by many ISPs in the Republic of Cyprus.

The NBA list implies that ISPs should do URL blocking as the entries of the blocklist contain URLs. ISPs would only be able to block them if they had previously deployed a blocking mechanism that would give the technical capability to ISPs to look inside the payload of the network packets, and more specifically at the layer 7 contents where the actual URL of an HTTP request is referenced, that technology is named as a DPI. In order to be able to filter HTTPS URLs the ISP needs to intercept the connection between the client (user of the ISP) to the server and perform an active man-in-the-middle attack on every HTTPS connection in order to decrypt the SSL/TLS, layer it and look at the unencrypted payload. Currently the SSL/TLS connections (HTTPS URLs) destined to the ISPs censorship infrastructure are not being handled (port 443 is unreachable). The connection times out and the user is not receiving any notification about the blocking in place apart from a connection error (error: couldn't connect to host).

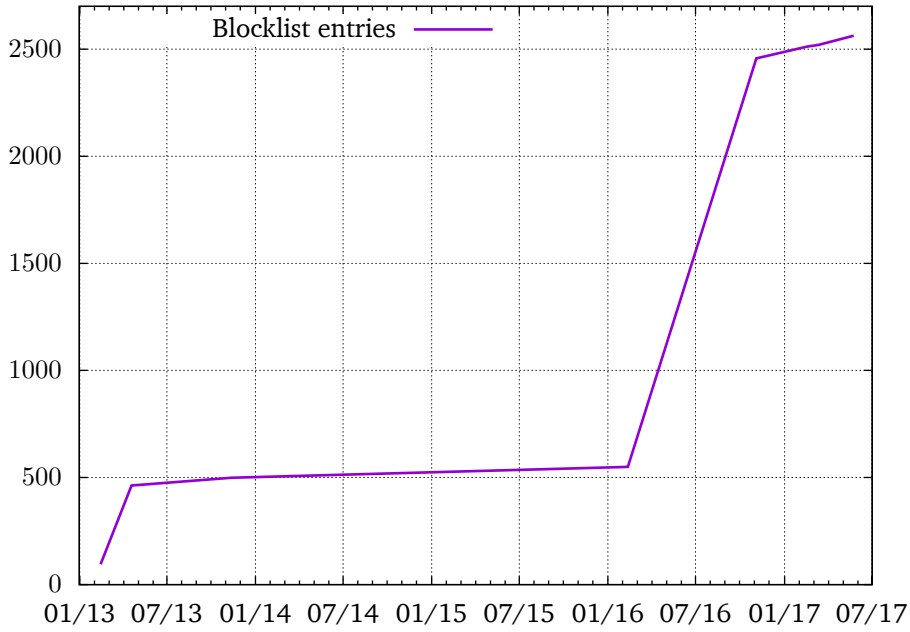

Figure 3.1: Timeline of the NBA blocklist publication

3.5. Methodology for data collection and analysis

We are using a variety of common free and open source software networking tools for gathering, categorizing, distributing, analyzing data and comparing the results. Acquiring results from a number of different ISPs is crucial to form a representative sample. We have conducted network measurements and used publicly available data based on OONI reports [201] submitted by volunteers. We were able to collect and process network measurement data from the following residential landlines and cellular Autonomous Systems (ASes): AS15805 (MTN Cyprus Limited), AS24672 (CallSat International Telecommunications Ltd.), AS35432 (Cablenet Communication Systems Ltd.), AS6866 (Cyprus Telecommunications Authority), AS8544 (Primetel) and AS197792 (Multimax Iletisim Limited). Even so, this remains a limited sample and the findings presented here are tentative and preliminary.

3.5.1 Data set used for the tests

First, we compiled a list of all URLs that are reported to be blocked in Cyprus as published and curated by the RoC NBA [77], the Greek gambling authority's blocklist [265] the Lumen database [149] for Turkey, and the community-collected global test list maintained by Citizenlab[165]. Additionally, we have used the public open DNS servers list provided by Digineo GmbH [119].

3.5.2 Collection of network measurements

The collection of the network measurements took place during the months of March to May 2017, though we were able to process relevant data submitted by volunteers from the months January and February earlier in 2017. Volunteers collected and submitted network measurements by using a custom set of tools and test lists [196] populated from the data sets enumerated in Section 3.5.1. For our censorship research we used ooniprobe, an application developed by the OONI project [105] and used by volunteers

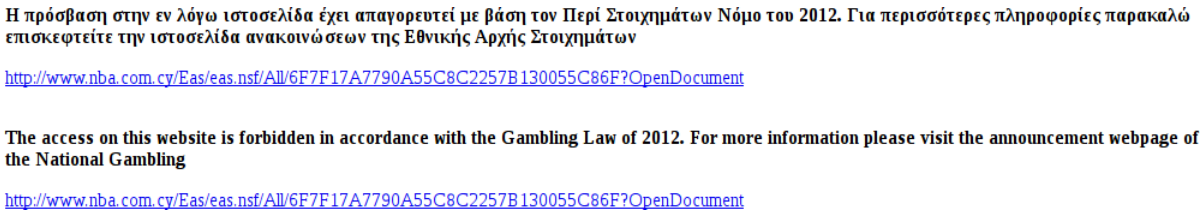

Figure 3.2: Callsat ISP NBA regulation landing page

and organizations to probe their network for signs of network tampering, surveillance or censorship. Developed with the idea of ensuring the detection of any interference to network communications, it aims to collect and provide high quality reports by using open and transparent data methodologies freely available to anyone that would like to process and analyze.

Ooniprobe is the application that was used to conduct the measurements on the ISP networks (both landline and cellular networks) where we detected network tampering and content blocking. Ooniprobe provides a variety of test cases and classes that could be used to probe the networks. More analytically, in our research we have deployed and analyzed a number of network measurements tests, precisely instant messaging, HTTP header fields manipulation and invalid request line tests, Tor and pluggable transports reachability tests as well as the web connectivity test. We were not able to identify any certain case of network interference in all of the tests apart from the web connectivity test. However this does not necessarily mean that there is no other sort of network interference happening on the network during different date periods or from different vantage points.

Web connectivity is an ooniprobe test methodology where we were able to identify and detect if a website is reachable and the reason or cause in case a website is not reachable. This test reaches a non censored control measurement endpoint (test helper) to assist with the comparison of the measurements for a given website. At first, the test performs an *A* DNS lookup to a special domain name service in our experiments; *whoami.akamai.com* that will respond to the *A* DNS lookup request with the resolver of the probe. Upon DNS resolver identification, the test will perform a DNS lookup querying the *A* record of the default resolver for the hostname of the URL tested. Following the test will try to establish a TCP session on port 80 or port 443 if the URL in question begins with the prefix *http* or *https* accordingly for the list of all IPs returned by the previous DNS query. Finally, the test performs a HTTP GET request for the path specified in the uniform resource identifier using the most widely used web browser user agent; *Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/47.0.2526.106 Safari/537.36* [240] as the HTTP header. Upon completion of the test, the gathered data are compared with the ones of the control measurement test helper.

3.6. Preliminary findings

We were able to perform measurements on the following ISPs: Cytanet (AS6866), Cablenet (AS35432) and Multimax (AS197792). Additionally, we were able to identify block pages based on reports contributed by volunteers to the OONI data repository [201] on ISPs Callsat (AS24672) and MTN (AS15805).

The most common identified method of content blocking on Cypriot ISPs is DNS hijacking. Since ISPs are in control of the DNS servers used by their users in residential broadband or cellular connections, they can manipulate the DNS servers' responses and can redirect the requesting users to anywhere they want. Taking advantage of this privilege, ISPs modify their resolvers to override censored domains' legitimate DNS replies by creating local zone entries [29]. These entries usually point to a server that they control

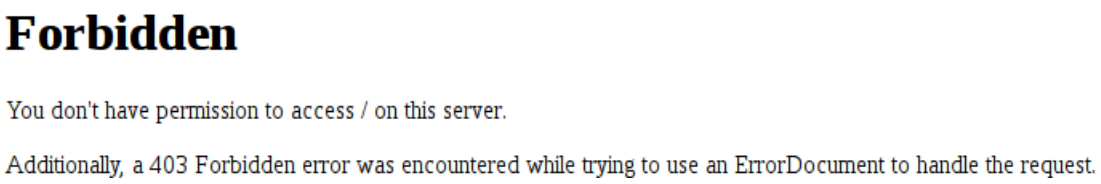

Figure 3.3: Cablenet ISP NBA regulation landing page

where they run a web server that displays a webpage with the warning message to users or block page.

3.6.1 Differences between ISPs

All ISPs, with the exception of Multimax in the north of the island, were using DNS hijacking as the blocking to control the access of the entries in NBA's list. Comparing the network measurements from all ISPs we found multiple cases of websites (entries of the blocklist) not being blocked, providing error messages (specifically HTTP status codes 403 and 404) or were unable to connect (connection failed) to HTTPS entries instead of the blocking page or the reason (legislation) why a user cannot access the specific website in question. Additionally, we were able to detect instances where email communication to the specific websites was also blocked although the law does not imply blocking email communication but only restricting access to the website that is included in the blocklist.

Additionally, at least one ISP was found redirecting the user to the website of NBA [78], leaking the IP addresses and possible the web browser's specific user metadata.

3.6.2 Callsat ISP

Network measurements analyzed from Callsat ISP [47] (AS24672) on the entries of the NBA blocklist revealed an outdated *landing* block page with a URL that points to a non-existent web resource (HTTP status code 404). The blockpage is illustrated in Figure 3.2.

3.6.3 Cablenet ISP

Our findings from the network measurements reveal that the Cablenet ISP [46] (AS35432) was directing users trying to access the entries of the blocklist to a generic error webpage (HTTP status code 403) without providing any justification of the blocking. The user may falsely assume that the website in question experiences technical issues. The blockpage is illustrated in Figure 3.3.

3.6.4 Cyta ISP

Cyta ISP [80] (AS6866) does not point the users to a blockpage but rather redirects the users trying to access the blocked entries from the blocklist to the NBA website. The excerpt from the HTML markup code is illustrated in Listing 3.1.

https://www.wikipedia.org
https://www.torproject.org
http://www.islamdoor.com
http://www.fepproject.org
http://www.no-ip.com
https://wikileaks.org
https://psiphon.ca

Table 3.4: Multimax ISP: List of blocked websites

```
<!doctype html>
<html class="no-js">
<head>

<meta http-equiv="content-type" content="text/html; charset=utf-8">
<meta name="copyright" content="copyright 2013">
<meta name="author" content="Designed & Developed by Cyta">
<meta name="distribution" content="global">
<meta http-equiv="refresh" content="0; url=http://www.nba.gov.cy/" />

</head>
</html>
```

Listing 3.1: Cyta ISP's HTML markup landing page

3.6.5 MTN ISP

Network measurements collected from MTN ISP [187] (AS15805) on one day (02/04/2017) show no evidence of blocking.

3.6.6 Multimax ISP

Multimax [188] (AS197792) is one of the ISPs that operates in the north of Cyprus. We have not identified any block pages, however, upon closer analysis, we found many similarities to the Turkish ISPs and specifically the blocking of web resources using IP blocking. We can conclude that the websites in Table 3.4 have not been accessible for the period of time during our network measurements. Note that the list of websites in Table 3.4 is not exhaustive and there could more websites or services that may be potentially blocked by this ISP.

3.6.7 Collateral damage

```
; <<>> DiG 9.9.5-9+deb8u10-Debian <<>> MX williamhill.com @82.102.93.140
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 19519
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 1280
;; QUESTION SECTION:
;williamhill.com.        IN  MX

;; AUTHORITY SECTION:
williamhill.com.    3600    IN  SOA ns1.cablenet-as.net. noc.wavespeed.net.
1483803163 10800 3600 604800 3600

;; Query time: 97 msec
;; SERVER: 82.102.93.140#53(82.102.93.140)
;; WHEN: Tue Apr 04 02:35:07 CEST 2017
;; MSG SIZE  rcvd: 113
```

Listing 3.2: Empty (no answer) DNS MX records for williamhill.com (DiG output)

In our research we identified that the MX records are absent, and do not contain the relevant DNS records that point to the email server of the domain name in question. That is rendering email delivery to the specific domain name impossible.

In the Listing 3.2 we have requested the MX records of the domain name *williamhill.com* from the DNS server *82.102.93.140* (DNS resolver in Cyprus operated by Cablenet) compared to the Google's DNS resolver *8.8.8.8* as illustrated in Listing 3.3. Google's DNS resolver answered with 2 entries in the query (*ANSWER: 2*) for the domain name in question whereas Cablenet's DNS resolver sent no answers (*ANSWER: 0*). The DNS queries took place on 4th of April, 2017.

```
; <<>> DiG 9.9.5-9+deb8u10-Debian <<>> MX williamhill.com @8.8.8.8
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 16093
;; flags: qr rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 512
;; QUESTION SECTION:
;williamhill.com.                IN      MX

;; ANSWER SECTION:
williamhill.com.         605     IN      MX      10
mxb-0010e301.gslb.pphosted.com.
williamhill.com.         605     IN      MX      10
mxa-0010e301.gslb.pphosted.com.

;; Query time: 40 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Tue Apr 04 02:43:51 CEST 2017
;; MSG SIZE  rcvd: 116
```

Listing 3.3: Google DNS MX records for williamhill.com (DiG output)

3.6.8 Circumventing blocking

Using a block-list/allow-list model (sort of an ON/OFF model) is not granular enough and tends to fail; it will have negative impact on users and customers of the ISPs that may or may not find routes around the blocking. Furthermore, creating such an ineffective blocking gives a false sense of security as the entities that enforced such blocking would assume that the content is being blocked although blocking is easily circumvented.

3.6.9 Using alternative DNS resolvers

Circumventing the blocking enforced by the ISPs is just a tweak in the network configuration and requires no technical expertise by using a different DNS resolver such as Google DNS [82] (*8.8.8.8*) or OpenDNS [145] (*208.67.222.222*).

3.7. Conclusions and future work

Although this case study initially focused on the blocking of gambling websites specifically, it brings up interesting data regarding more general blocking practices in the north of the island, which need to be further investigated. One example is our finding that the RoC block list isn't blocked in the north of Cyprus, but that a number of other websites have been blocked there, matching the list of websites blocked in Turkey (see Table 3.4). This opens up a discussion of more than one regime of freedom of expression on the island, and also raises the question of whether there may be a third point of difference with blocking practices implemented in the British sovereign bases. Furthermore, our intention is to confirm with ISPs

regarding the technical infrastructure used to implement blocking.

As explained in the introduction, this is only the beginning of an effort to more closely study how online content-blocking is legislated and implemented, in an effort to understand the political extensions of these practices and their related dangers to internet freedom and freedom of information. For example, beyond issues of content-blocking and connected debates around censorship, the data collected here also implicate issues of privacy and personal data protection (with regard to ISP redirection practices), as well as issues of transparency (with regard to how content-blocking is implemented). We hope that the evidence this research begins to produce will come to feature in further discussion, leading to a better understanding of the dangers as well as alternative and safer technical options. More ambitiously, we hope that this research will promote a more sensitive approach guiding policy and national legal provisions that will more effectively safeguard the aforementioned freedoms.

– 4 –
Internet censorship analysis in Spain

4.1. Introduction

Surveillance and network interference infrastructures are increasingly deployed in EU member states to contain content and services that do not comply with EU legislation [19], e.g., online gambling, copyrighted material, incitement to the commission of crimes, depictions of cruel violence against humans, human death or mortal suffering, child or animal exploitation material [241]. However, despite the presumably tacit assumption that illiberal practices in the digital realm are rather likely to affect only authoritarian states, EU member states also gain attention with respect to incidence and modalities of Internet censorship [44, 234, 235, 272, 275]. Moreover, instances of "everyday acts of authoritarianism" online could be observed also in the democratic West, often with industry-state collaboration and no democratic oversight [131]. To this end, we define online censorship as any form of network interference that disrupts the normal operation of services or content in the World Wide Web to prohibit access to a specific audience. Previous research examined the presence of censorship in various countries such as China [95, 181], Thailand [116], Bangladesh [186], Pakistan [3, 189], India [122, 289], Iran [16, 28], Syria [49, 233], Turkey [246, 247], Russia [222], and Mexico [154]. There is almost no previous research about the topic of censorship in Spain, except for some clues [5, 18, 33, 177, 213]. Lundström and Xynou [177] observed that 25 sites related to the 2017 Catalan independence referendum were blocked from September 25 up to the day of the referendum (October the 1st), utilizing DNS manipulation and HTTP blocking, based on the OONI network measurements data retrieved from three local networks. A technical report by Ververis et al. [271] provides an analysis on persistent blocking of the Women On Web (WoW) website by all major ISPs in Spain from network measurements of the first quarter in 2020.

Referring to the lack of similar studies and seeking to fill the identified research gap, this article examines the practice of Internet censorship in Spain. Motivated by partial insights [5, 18, 33, 41, 177, 213, 271] and based on historical network measurements provided by the data of OONI [207], our research observes fours years (July 2016 to May 2020), including October 2017, where the Spanish referendum on Catalonia independence took place. The referendum was called by the Catalan authorities, but declared unconstitutional and suspended by the Spanish government. Held amidst repression and violence by the central government, it asked Catalan citizens: *"Do you want Catalonia to become an independent state in the form of a republic?"* The *"yes"* won with over 92 percent of popular vote [213]. Due to its highly controversial nature, the referendum represents an excellent case to observe online censorship in action. We set 2016 as the starting year for our analysis due to higher availability of OONI data. We address the following research questions:

- Which network interference techniques were in place in Spain over the past four years?
- How did the techniques evolve during the investigated time period?
- How can such an Internet censorship study be reproduced, and our method generalized to other cases?

- What are the limitations of such a long-term historical data analysis and how can we improve the measurement collection and analysis methodology?

The paper is organized as follows: Section 4.2 provides the methodology for choosing a data source, processing and validating the network measurement data used in this study, categorizing websites into categories. Then, Section 4.3 proceeds with analyzing the data and reports our findings of network blocking via HTTP blocking, DPI, DNS Manipulation, domain seizure, the case of WoW website blocking, Server Name Indication (SNI) blocking, and TLS interception, with the improvement of TLS interception testing methodology in OONI, followed by the circumvention of DPI blocking and the reproducibility of our research. Finally, the general contributions of our study and implications for research and practice are discussed in Section 4.4, followed by the ethical considerations in Section 4.4.2 and our conclusions in Section 4.5.

4.2. Methodology

We begin by introducing our methodology. This consists of four main parts: (i) choice of an appropriate data source; (ii) processing and (iii) validation of network measurement data; (iv) clustering websites into categories; (v) data analysis.

4.2.1 Data sources

We surveyed several tools that perform network measurements to detect Internet blocking and provide a repository of historical data, with a special focus on residential endpoints [4, 30, 193, 207, 220]. Specifically, we considered IClab [193], Censored Planet [220], RIPE Atlas [30] and OONI [207]. IClab mainly uses VPN endpoints for its network measurements [193]. Censored Planet tests scan the IP address space for accessible public servers excluding end-user devices and target servers, routers or embedded devices [263] and therefore do not cover residential ISP networks. RIPE Atlas is not designed to measure Internet censorship and thus HTTP measurements are not allowed to run on residential ISPs [30]. Albeit one could infer useful information by performing other available tests on residential ISPs that block websites. We abstained from using RIPE Atlas probes due to the ethical considerations and the inaccessibility of Internet blocking methodologies. Nonetheless, we queried all evaluated data repositories for any historical network measurements that could match our study's requirements. We did not find any matching data that apply to our research. Out of all the evaluated tools, only OONI provides longitudinal data of historical network measurements for our desired period (years of 2016 to 2020). We found the OONI data repository provides adequate data of over 3 million network measurements from all major residential ISPs in Spain over the last 4 years. Nevertheless and as with any software, OONI Probe software has some limitations in their TLS blocking test methodology, as we found during our data analysis. We present detailed explanations on how we overcome this limitation and implement a new testing methodology, which has been approved by the OONI developers, and is now in further development for wider adoption to the public (see Section 4.3.7). All software components of OONI's source code are released under a free and non-restrictive license and are available for everyone to download, modify and use.

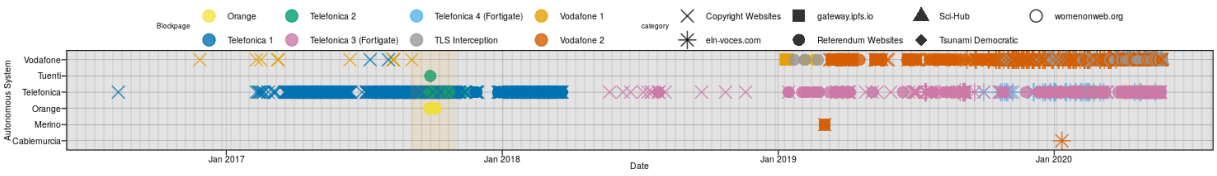

Figure 4.1: HTTP Blocking of ISP/Date per website category of OONI data in Spain

4.2.2 Data processing

OONI has been collecting network measurements from anonymous volunteers since 2012 [207] to detect evidence of possible network interference that might relate to Internet censorship or surveillance on different vantage points, primarily from anonymous end-hosts in residential networks. OONI data is made available under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license and could thus be used freely and without restrictions in our research study. We chose as the blocking methodology test the Web Connectivity test [203], that measures the reachability and possible blocking of any website, given an IP address or domain name. The test consists of the following steps: i. Performing A and AAAA DNS lookups and storing the result of the A records list, ii. Attempting to establish a TCP session on either port 80 or 443 (depending on the Uniform Resource Identifier (URI) scheme), and iii. Performing an HTTP GET request to the path specified in the URI. In all steps, the responses and possible errors are recorded in a JSON file and submitted to the OONI network measurements collector for further processing and archiving [203].

To get access to the OONI data, one may use the OONI API and OONI Explorer. However, both tools have some limitations regarding the file size of the measurements and the computational time required to get a vast amount of data. To overcome these limitations and not stress the OONI services, we setup a PostgreSQL replica of the OONI meta database. Next, we fetched the latest archived data required for a database cluster (PGDATA). A helper script was used to fetch the OONI S3 bucket data and configure the PostgreSQL server as a replica (in a hot standby configuration). It took about 10 days to sync with the master database and 800 GiB of storage capacity to accommodate the OONI meta database. The main requirement of a replica is a system with enough storage capacity and network connectivity to host a PostgreSQL database. Once the meta database was synced, we were able to run queries based on our blocking methodology heuristics and the criteria set to eliminate potential false positives. We used self-developed IPython notebooks (see Section 4.3.10) to query the database for our study data and process the retrieved data, whereas we used heuristics (see Section 4.2.4) to validate the correctness of data and ensure that there are no false positives or negatives. Here, we also categorized the data for further processing. Next, we used the R programming language to plot graphs. Finally, all the data were summarized and exported in CSV files for easier sharing and reproducibility.

4.2.3 Blockpage similarity heuristics

Based on the discovered blocked websites we built heuristics that reveal evidence of network interference. We used the simhash [180] technique to estimate the HTML body text and length of blockpage fingerprints found, allowing us to detect further blockpage fingerprints. The ISP blockpages are typically based on static pages as they are easier to configure and less computationally demanding in comparison to the dynamic blockpages. However, dynamic blockpages include more information, such as tracking bits, that can be used by customer support or other entities within an ISP for statistical purposes or legal regulation. The SHA256 checksum of the HTML body can be reliably used as a fingerprint to identify static

blockpages triggered by other websites. In the case of dynamic blockpages, the low hamming distance between the blockpage and the HTML body simhash was used to reliably discover further fingerprints and identify new blocked websites. This applies to the OONI meta database columns *body_simhash* and *body_text_simhash*. Due to the vast amount of data, we built more than a hundred fingerprints, first to eliminate any false positives, subsequently to detect new blockpages, and finally to enumerate all blocked websites. Additionally, to be confident that our methodology was correct, we manually inspected each of the detected blockpage fingerprints to eliminate any potential remaining false positives. As we created more fingerprints, we iterated our data validation process until no more false positives were left in our dataset. For the qualifying blocked websites, we defined a set of requirements and criteria, described in Section 4.2.4.

4.2.4 Data validation

Despite the presence of multiple network measurements with signs of network interference, we included only those blockpages or instances of blocking that could be verified with certainty, i.e., neither being false positives (for instance due to network connectivity errors) nor blocked due to internal network filtering regulations (such as parental controls, antivirus filtering or firewalls). Specifically, we considered only network measurements that suffice the following heuristics:

- Existence of a blockpage or any indication of blocking error (i.e. HTTP status code 403);
- Existence of DNS records that point to bogon IP addresses (such as 127.0.0.1);
- Removal of network measurements with wrong AS information (i.e. AS0);
- No blocking based on internal network filtering infrastructure (parental controls, firewall, antivirus, proxies);
- No blocking based on CDN or webserver specific filtering or security products (such as Cloudflare, Sucuri, Incapsula, Zscaler).

Due to ethical considerations, we excluded from our analysis those network measurements that may violate the anonymity of the person/entity who submitted them. We reported possible personally identifiable information in network measurements to the relevant software entities responsible for collecting these data.

4.2.5 Website categorization

Based on the finding of our data analysis, we grouped the blocked websites into five categories, regarding their content and purpose as follows:

1. **Civil Rights and Political**: This category was reserved for the websites of WoW (*womenonweb.org*) and *eln-voces.com*
2. **Sci-Hub**: Here we included the website mirrors of Sci-Hub, a file-sharing repository of research papers and books;
3. **Democratic Tsunami**: This category involves websites related to the Catalan protest group Tsunami Democràtic (in English, Democratic Tsunami);

4. **Referendum Websites**: We reserved this category for websites dealing with the Catalan referendum in 2017;
5. **Copyright Websites**: Here we included websites being blocked on grounds of copyright infringement such as video streaming, IPTV, online indexing of magnet links and torrent files.

The complete list of all the blocked websites by category type is listed in Table 4.2.

4.3. Analysis of network blocking

In this section, we analyze end-host measurements of network interference in Spain over the last 4 years (from 2016 to 2020) to spot instances of blocking and information controls, the network interference techniques being in place, and the extent of their usage. In total, we process over 3 million network measurements (3,089,892) from 17 different ASes that correspond to ISPs covering 98.45% of all broadband and 90.94% of all mobile subscribers in Spain [184]. Although much of the blocking is related to the Catalan referendum, the blocking is not limited to the autonomous community of Catalonia, but is experienced by users in all parts of Spain. The date range of network measurements during the referendum is highlighted with a color overlay visible in Figures 4.1, 4.3 and 4.4. We partition our data analysis into different sections, according to the type of network blocking methodology detected in Spanish ISPs; The list of AS network names, as well as the numbers and the dates of AS registration allocation are listed in [267].

4.3.1 HTTP blocking

The first case of HTTP blocking in the network measurements of OONI data from Spain was identified on the 8th of November 2016 with the blocking of the URL *thepiratebay.org*. The website was found systematically blocked under the ISP Telefonica (AS3352). The ISP didn't present any reasons for the blocking, which is a common practice when a website is blocked by an ISP despite the lack of transparency. Instead, users received the *ERROR 404 - File not found* error message that falsely indicates a website error [267]. The relevant measurements of this blockpage are illustrated in Figure 4.1 under the group name *Telefonica 1*. Later on, on the 26th of November 2016, we see the same URL being blocked for the first time within the Vodafone ISP (AS12430 and AS6739). In this blocking instance, users were redirected with the HTML meta refresh method (*http-equiv="refresh"*) to the blockpage URL *http://castor.vodafone.es/public/stoppages/stop.htmopt* [267]. This blockpage is represented as *Vodafone 1*, in Figure 4.1. Subsequently, after the 11th of May 2017, the same URL was found being blocked with a different blockpage in Telefonica. However, the string *PHISHING_TSOL_MENSAJE_1* in the HTML source code may indicate the Telefonica Solutions Group (TSOL) could be using the same blockpage to filter phishing websites [109]. Additionally, on the source code of that blockpage, we found the name of another authority (in Spanish) *Administrativo Ley del Juego* redirecting users elsewhere to the IP address of the (blockpage) *http://195.235.52.40* [267]. An indication that the blockpage may be used to block other websites. In Figure 4.1, this blockpage is tagged as *Telefonica 2*.

Information controls of Catalan Referendum

We identified 24 unique blocked URLs including the InterPlanetary File System (IPFS) gateway, a peer-to-peer network for storing and sharing data over a distributed filesystem. The categories of websites blocked

Civil Rights & Political	Sci-Hub	Democratic Tsunami	Referendum	Copyright
womenonweb.org	sci-hub.se	api.tsdem.org	alerta.cat	digitalplatinum.in
eln-voces.com	sci-hub.tw	app.tsdem.org	aniol.github.io	c14.xtra7.gq
		app.tsunamidemocratic.cat	cat.referendum.barcelona	digitalservices.tel
		app.tsunamidemocratic.com	garantiespelreferendum.com	elitegol.global
		democratictsunami.eu	nigeon.github.io	elitegol.tv
		tsdem.org	pedrosanchez.cat	elitegoles.com
		tsunamidemocratic.cat	ref1oct.cat	elitegoltv.me
		tsunamidemocratic.com	ref1oct.eu	elitegoltv.org
		tsunamidemocratic.github.io	ref1oct.net	futbolpirlotv.net
		tsunamidemocratic.net	ref1oct.org	gtmservices.org
			referendum.enricpineda.cat	hightquality.org
			referendum.fun	intergoles.me
			referendum.fyi	intergoles.net
			referendum.legal	iptvadur.eu
			referendum.lol	iptvesp.eu
			referendum.love	iptvld.paranosotros.ru
			referendum.ninja	iptvtool.es
			referendum.observer	landiptv.live
			referendum.party	locopelis.com
			referendum.pau.fm	mamahd.org
			referendum.pirata.cat	movspain.com
			referendum.rip	pandorapremium.ddns.net
			referendum.soy	playlist.topcam.net
			referendum.voto	pirlotv.es
			referendum.works	pirlotvhd.net
			referendum.zalo.nyc	pirlotvhd.online
			referendumcat.eu	pirlotvonline.net
			referendum.cat	playlist.topcam.me
			gateway.ipfs.io	qualitypremium.sytes.net
				realstreaming.net
				redstreamsport.online
				rojadirectaenvivo.es
				sansat.net
				sendspace.com
				todocvd.com
				thepiratebay.org
				thepiratebay.se
				veopartidos.online
				wolftm.in

Table 4.2: Blocked websites by category type

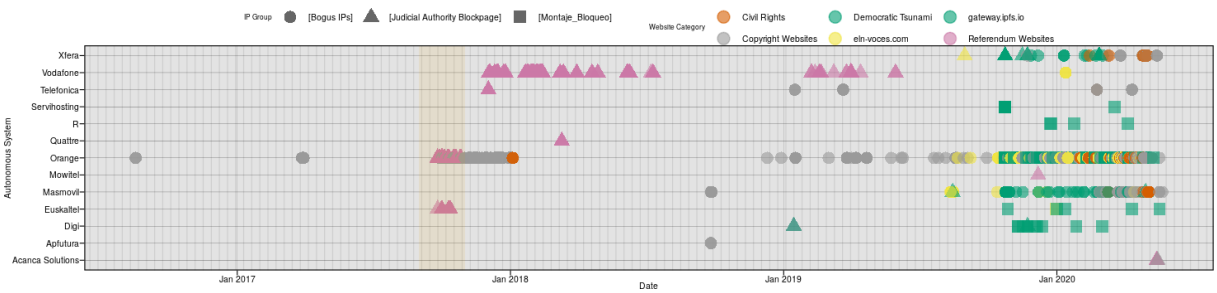

Figure 4.3: DNS Manipulation of ISP/Date per website category and IP group of OONI data in Spain. For an expanded version of this figure with layers per Website category and AS see Figure 4.4

during that period were copyright and referendum websites. The complete list of the blocked URLs can be found in Table 4.2. Furthermore, we identified seven new blockpages that contained information related to the referendum, including the names of the authorities under which the websites are blocked and which changed in later versions of the blocking from *PHISHING_TSOL_MENSAJE_1* to *Judicial_Policia_Nacional*.

Javascript switch statement for different blocking rules

The HTML body of the blockage [267] tagged in Figure 4.1 as *Telefonica 3 (Fortigate)* indicates that Telefonica may block more websites. In the code section, a switch statement evaluates the name expression, that matches the value to the case clause. In this blockpage, there are four different cases that set the HTML *h1* heading element or redirect to a URL, specified by the *replace()* method of the location interface. Specifically, the first case *PHISHING_TSOL_MENSAJE_1* sets the heading to *Error de acceso por contenido no identificado* (translated from Spanish to *Access error due to unidentified content*). The second case clause sets the heading of the page to *Administrativo_Ley_del_Juego* and redirects users to the blockpage hosted in Telefonica's network at *http://195.235.52.40*. The third case clause used by Guardia Civil sets the heading of the page to *Judicial_Guardia_Civil* and redirects the user to *http://paginaintervenida.edgesuite.net* when triggered by a blocked website related to the Catalan referendum. This blockpage is hosted in Akamai's network. Last, the default case (*id="causa"*) corresponds to the blockpage of *http://thepiratebay.org* which sets the heading of the page to *ERROR 404 - File/block not found* and which redirects users to the URL *http://webbloqueadaporpolicianacional.com* .

When further examining the blockpage's source code, we identified the URL of *Judicial_Guardia_Civil*, redirecting users to the URL *http://www.marca.com*, a Spanish national sports website owned by the company Unidad Editoria. We observed that information regarding the blocking was rather minimal or non-existent, e.g., given by an error code message at a website (HTTP 404). All source codes of the blockpages found to trigger this blocking technique (Javascript switch statements) are listed in [267]. One variation of the blockpage is illustrated in Listing 4.1.

Orange ISP was found to censor websites via HTTP blocking only during the period of the Catalan referendum. Later on, Orange switched to blocking websites via means of DNS manipulation. This finding probably suggests that Orange ISP used a different type of network blocking for censoring websites related to information on the Catalan referendum. Specifically, Orange ISP presented to users a blockpage with the exact source code used in URL *http://paginaintervenida.edgesuite.net* , however, Orange didn't redirect its users to the blockpage but rather used HTTP blocking to block access to the websites in question. The relevant measurements are grouped under the shape *Orange* as illustrated in Figure 4.1.

We proceed with studying detected blocking instances after the referendum period up to the end of our data analysis (2017-11-01 to 2020-05-21). In Vodafone networks, we identified two additional distinct blockpages. The first [267] was deployed to block a few referendum websites and the IPFS gateway, whereas the second blockpage [267] was extensively used for other website categories such as websites related to copyright, referendum (including the IPFS gateway), as well as civil and political websites. As in previous years, Telefonica ISP censored websites without any explanation of the reason for the blocking in contrast to other ISPs. Based on the network measurements, we extracted 21 different blockpage variations grouped in 4 different blockpage tags (*Telefonica 1* to *Telefonica 4*) per website category, as illustrated in Figure 4.1.

```
<!DOCTYPE html PUBLIC \"-//W3C//DTD HTML 4.01//EN\">
<!--
```

```
CATEGORY: Judicial_Guardia_Civil
DEST_IP: 107.180.40.196
FGT_HOSTNAME: RFFBTB1-02
SOURCE_IP: [REDACTED]
-->
<html>
  <head>
    <meta charset="utf-8"/>
    <title id="4">
      Judicial_Guardia_Civil
    </title>
  </head>
  <body>
    <CENTER>
      <h1 id="causa" name="Judicial_Guardia_Civil">
      </h1>
      <script type="text/javascript">
        var name = document.getElementById("causa").getAttribute('name');
        var text = "";
        var f = new Date();
        var mes=f.getMonth()+1;
        var dia=f.getDate();
        var hora=f.getHours();
        var minu=f.getMinutes();

        if (mes<10) mes="0"+mes;
        if (dia<10) dia="0"+dia;
        if (hora<10) hora="0"+hora;
        if (minu<10) minu="0"+minu;

        switch (name) {
          case "PHISHING_TSOL_MENSAJE_1":
            text = "PTM1_Judicial_Policia_Nacional";
            window.location.replace("http://a.policia.es/?url=www.eln-voces.com/");
            break;
          case "Administrativo_Ley_del_Juego":
            text = "Administrativo_Ley_del_Juego";
            window.location.replace("http://192.187.20.203");
            break;
          case "Judicial_Guardia_Civil":
            text = "Judicial_Guardia_Civil";
            window.location.replace("http://82.223.97.47");
            break;
          case "Judicial_Policia_Nacional":
            text = "Judicial_Policia_Nacional";
            window.location.replace("http://webbloqueadaporpolicianacional.com");
            break;
          case "Direccion_General_de_la_Policia":
            text = "Direccion_General_de_la_Policia";
            window.location.replace("http://a.policia.es/?url=www.eln-voces.com/");
            break;
          default:
```

```
            text = "ERROR 404 - 'Judicial_Guardia_Civil' not found";
        }
        document.getElementById("causa").innerHTML = text;
      </script>
    </CENTER>
  </body>
</html>
```

Listing 4.1: AS3352 Blockpage of byte size 2374 (Fortigate)

4.3.2 Deep Packet Inspection

Several blockpages that were found in network measurements of Telefonica confirmed the existence and usage of the DPI equipment vendor Fortinet [107]. Specifically, the blockpages with sizes of 332 to 339 bytes exposed several configuration settings of Fortinet's Fortigate DPI equipment used by Telefonica ISP [267]. The difference of 7 bytes between the blockpages is due to the different configuration options of hostnames and IPs. From the comments section revealed in the HTML source code of the blockpage, the settings of the Fortigate device can be ascertained as: *CATEGORY* for the web filter category (if any), *DEST_IP* for the destination IP of the blocked resource, *SOURCE_IP* for the source IP of the request (the source IP of the user) and the *FGT_HOSTNAME* that reveals the hostname of the Fortigate device. According to the documentation of Fortinet, the aforementioned variables (except the category variable) are used as replacement messages for the web filtering, thus the variable will change dynamically depending on the user's IP, targeted websites, and Fortigate device's hostname [108]. The *SOURCE_IP* variable is masked with the word [REDACTED]; this is done by the OONI software to protect the privacy of the users and not leak any personally identifiable information.

Further blockpages found under the Telefonica networks with sizes of 1290 and 1292 bytes reveal more configuration settings of the Fortigate devices. They expose (among other settings) the *POLICY_UUID*, which is the Universal Unique Identifier (UUID) for the policy in Fortigate's configuration. The complete blockpage with byte size 1290 [267] is illustrated in Figure 4.1 with the tag *Telefonica 3 (Fortigate)*. Few measurements from Telefonica found in this period reveal blockpages with sizes of 1514 and 1517 bytes deployed only during the referendum period, until mid January 2018 [267]. These blockpages are illustrated in Figure 4.1 under the tag *Telefonica 2*. Finally, the last blockpages identified in Telefonica target exclusively the URL *http://www.eln-voces.com/*. In this case, we see a variation of previous blockpages analysed in this section with the addition of one more entity listed as a switch case (analyzed in Section 4.3.1) in blockpage's source code, *Direccion_General_de_la_Policia* redirecting users (location replace in Javascript) to the URL *http://a.policia.es/?url=www.eln-voces.com/*. However, the category name on Fortigate's device configuration and the HTML title are set to *Judicial_Guardia_Civil* and not to the *Direccion_General_de_la_Policia* as the URL suggests, perhaps due to human error or misconfiguration. These blockpages have a size of 1989 [267] and 2186 bytes. Another finding of the blockpages with size of 2374 and 2377 bytes used in this period reveals two more cases used to block websites in Spain; *PETICION_JUDICIAL_140120* and *Administrativo_Ley_del_Juego_Temporal* both redirecting users to different URLs [267]. The blockpages are grouped under the tag *Telefonica 4 (Fortigate)* illustrated in Figure 4.1.

4.3.3 DNS manipulation

The first identified network measurement that revealed DNS manipulation was detected for the domain name *thepiratebay.org* in Orange ISP (AS12479) on date 2016-08-18 and later for the domain *thepiratebay.se*. In all measurements that displayed signs of DNS manipulation in this era, we found that the A record of the domain names in question pointed to the bogon IP address *127.0.0.1*, commonly reserved for use as the Internet host loopback address (localhost). Internet Protocol version 4 (IPv4) network standards reserve *127.0.0.1* for loopback purposes (and the complete /8 IP address block) must not appear in any network on the Internet (RFC 1700) [218]. The results are illustrated in Figure 4.3 under the point shape name *Bogon IPs*. Furthermore, we identified 24 unique blocked domains, including the IPFS gateway (*gateway.ipfs.io*) as well as 2 GitHub pages (*aniol.github.io* and *nigeon.gihub.io*) being consistently blocked during the period of the referendum (in October 2017). The blocking of the GitHub pages is evident only via DNS manipulation because of the collateral damage the HTTP blocking of GitHub could have caused (i.e. HTTP blocking would result in the complete blocking of GitHub website whereas now only specific pages of users are being targeted). This is not the case though for the IPFS gateway that is blocked employing DNS manipulation and also via HTTP blocking.

After the referendum period, new websites are still blocked on the remaining categories: copyright, Democratic Tsunami, Sci-Hub, civil rights, and political in the networks of Orange, Masmovil, Telefonica, and Vodafone ISPs. Apart from censorship of websites for copyright reasons, information controls of the referendum and attempts to silence protests from Democratic Tsunami, we found another case of political censorship concerning the website *eln-voces.com*. In the absence of any further information, we assume that the specified website was blocked because of the content from the terrorist organization National Liberation Army, as defined by the European Union council decision 2017/1426 [64]. From historical DNS data and snapshots from Wayback Machine of Internet Archive, the domain name *eln-voces.com* was expired from 2019-06-12 to 2019-08-25 and then registered by a different hosting entity with unrelated content [26]. Moreover, we found that the website *womenonweb.org*, a non-profit organization providing support to women, was blocked since 2020-01-30 and until the completion of our research (2020-05-21). The website is blocked in ISPs Orange, Masmovil, Telefonica, and Vodafone. See Section 4.3.5 for a detailed analysis of this blocking. All domain names blocked via means of DNS manipulation are listed in Table 4.2. Figure 4.1 illustrates the website category blocked and under which blockpage (IP group).

4.3.4 Domains seizure

The domains *alerta.cat*, *ref1oct.cat* and *referendum.cat* were seized and their DNS records pointed to a website hosted under the Akamai Technologies CDN network (*edgesuite.net*) with the logo of the Spanish judicial authority and the following text: *This domain name has been seized according to a seizure warrant under the Judicial Authority and is under its administration*. The website *paginaintervenida.edgesuite.net* was still accessible as to the time of this research. However, we found the aforementioned domains blocked in specific networks via means of DNS manipulation or HTTP blocking. The domains *referendum.clash.cat*, *marianorajoy.clash.cat* and *marianorajoy.cat* were not blocked but were instead seized by the Spanish Judicial Authority.

4.3.5 Blocking of Women On Web

Our data analysis revealed the persistent blocking of the WoW website by all major Spanish ISPs. OONI network measurements indicate that most Spanish ISPs had been blocking the WoW website since the end

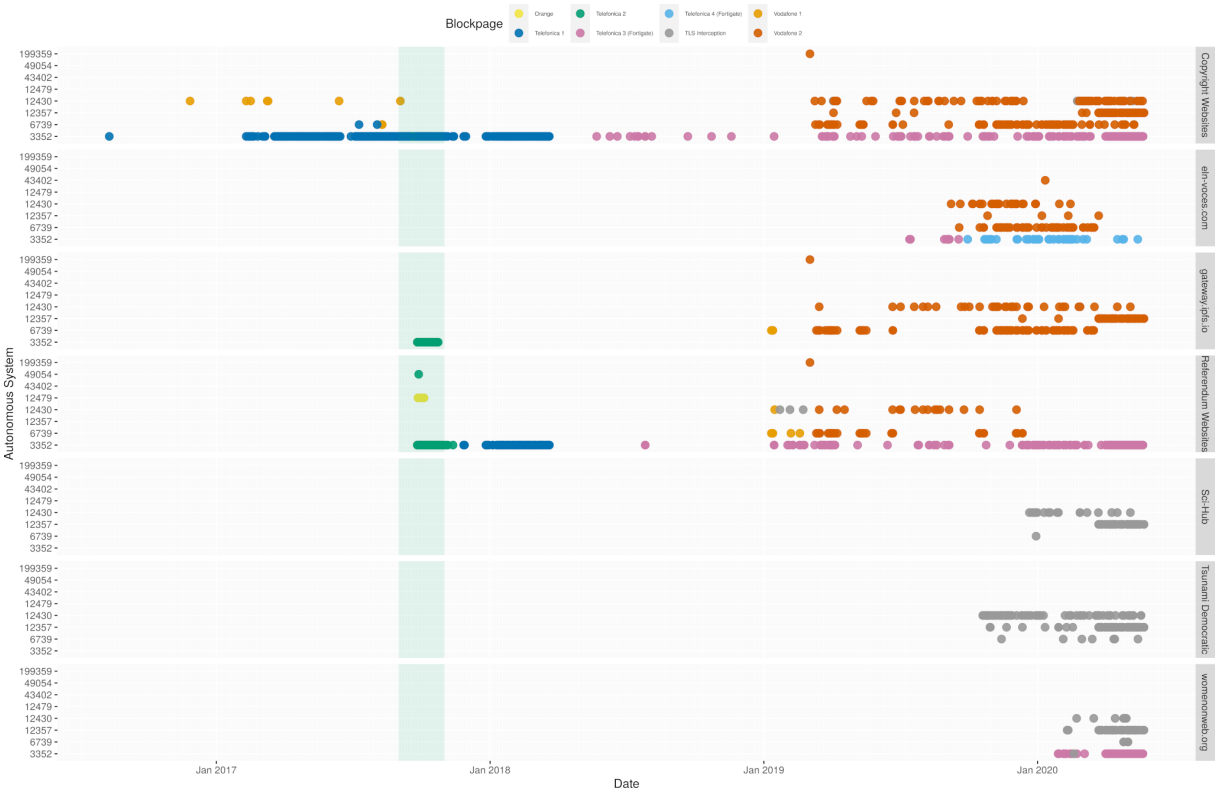

Figure 4.4: DNS Manipulation of AS/Date per website category (in layers) and IP group of OONI data in Spain

ISP	Blocking Technique	DPI
Telefonica	DNS Manipulation, HTTP Blocking	Fortinet
Vodafone	HTTP Blocking, TLS Interception	Allot
Orange	DNS Manipulation	-
Masmovil	DNS Manipulation	-

Table 4.5: Women On Web website blocking techniques per ISP

of January 2020. The blocking methodologies are similar to the other blocked websites as determined in our data analysis: DNS manipulation and HTTP blocking using DPI infrastructure. Our data analysis and reports from volunteers indicate that the following ISPs blocked the WoW website: Vodafone, Orange, Masmovil, Xfera, and Telefonica. Table 4.5 summarizes the blocking methodology as well as the DPI technology (when applicable) deployed per ISP.

Measurements from Vodafone (AS6739) show another blocking strategy, consistently over time, suggesting that between 16/03/2020 and 24/04/2020, Vodafone moved from a simpler to a more complex blocking strategy. Additionally, the identified DPI products analyzed in Section 4.3.2 are both used to block access to WoW: Fortinet in Telefonica's network and Allot in Vodafone's network. In January 2021, WoW filed a lawsuit at the Spanish National Court for the illegal and unjustified blocking of their website [279].

4.3.6 SNI blocking

Another technique detected in Vodafone networks is SNI blocking. SNI is a TLS extension used in web-servers that host multiple websites reachable under HTTPS on the same server. The SNI attribute is

transmitted in clear text and provides the website's hostname in question, thus making it easy to block. TLS protocol (version 1.3) adds experimental support for SNI encryption. However, as TLSv1.3 is a relatively new protocol and given that SNI encryption is still experimental, it may take some time to get widely deployed. As of the latest estimations, deployment of TLSv1.3 on popular domains is about 30%, and 10% across the com/net/org top-level domains [135]. In Section 4.3.7, we present further details on our TLS interception findings during our data analysis that also apply to the case of the WoW website blocking.

4.3.7 TLS interception

We discovered several measurements that present a certificate verification failure in Vodafone networks (AS12357, AS12430, and AS6739). The error (*ssl_error:error:14007086:SSLroutines:CONNECT_CR_CERT:certificate* `verify failed`) indicates that there could be TLS interception on the network. This error message from the OpenSSL library indicates that the TLS handshake is over, and the client cannot verify the certificate provided by the server. OONI's current test methodology does not capture any further details related to TLS interception. Thus, we performed further tests using the OpenSSL command-line tool. We discovered that we were connecting to a box serving a forged, invalid TLS certificate claiming to be the blocked website. This box was the same one hosting the Vodafone blockpage [267]. All websites or categories blocked utilizing TLS interception are illustrated in Figure 4.1.

4.3.8 Improvement of TLS interception testing methodology

We used the results collected by OONI's Web Connectivity experiment [203]. This experiment implements the following algorithm. It takes in input a website's URL (either using HTTP or HTTPS). It resolves the website's domain name using the system DNS resolver. Then, it attempts to connect to every resolved IP address. Next, it tries to fetch the website's URL. Finally, OONI compares its measurement results with a concurrent measurement performed by a test-helper server to detect false positives.

Issues with OONI's Web Connectivity

In the context of WoW TLS blocking, the main issue of OONI's Web Connectivity methodology is that it did not collect enough low-level information around the TLS connection. To overcome this limitation, we implemented Aladdin, a ten-step network experiment based on the OONI measurement engine that significantly extended the Web Connectivity methodology [35] to characterize the WoW censorship case.

Description of Aladdin

The input of Aladdin is a website's domain name. Aladdin assumes that the website is available over both HTTP and HTTPS. These are the Aladdin's steps:

The first step checks whether there is SNI-based blocking. We connect to an unrelated server (e.g., example.com:443) using the SNI of the target website (e.g., blocked.com) and an unrelated SNI (e.g., ok.com). If only the connection using the target website SNI is blocked, we conclude that there is probably SNI-based blocking.

The second step checks whether there is Host-header-based blocking. We connect to an unrelated server

(e.g., example.com:80) using the Host header of the target website (e.g., blocked.com) and an unrelated Host header (e.g., ok.com). If only the connection using the target website Host header is blocked, we conclude that there is probably Host-header-based blocking.

The third step checks whether there is DNS injection. It sends a DNS query to a host that we know is running no DNS server. If we get back a reply, then there is DNS injection.

The fourth step queries the default resolver (like Web Connectivity does). In addition to recording the returned addresses, this step notes whether any of them is a private address (e.g., 10.0.0.1).

The fifth step repeats the DNS query using Google's DNS over HTTPS (DoH) server. Then it checks whether the IPs returned by the default resolver are consistent with the ones returned via DoH.

The sixth step fetches the webpage over HTTPS using the IP addresses returned by the system resolver. Suppose an IP address returned by the system resolver is invalid for the domain (i.e., suppose it is a private address or just the address of an unrelated server). In that case, this step will fail because TLS would not be able to map the returned certificate to the requested domain.

The seventh step fetches the webpage using the Psiphon circumvention tool. We compare the webpage fetched using Psiphon to the one fetched in the sixth step. This step is similar to what Web Connectivity does, except that we are using the Psiphon circumvention tool instead of the test helper.

The eighth step disables TLS certificate validation and then fetches the webpage again. This step allows collecting the returned certificate and possibly fingerprinting the blocking device.

The ninth step repeats the sixth step, except that it uses the IP addresses returned by the DoH resolver.

The tenth step is like the ninth step, except that we explicitly force the code to use TLSv1.3. In TLSv1.3, the server's certificate is encrypted. This fact gives us confidence that blocking depends on the cleartext content in the Client Hello (typically, the SNI).

Findings

After repeatedly running the Aladdin experiment for WoW, we discovered the following: (1) there was no SNI-based blocking (step 1); (2) following the IP address returned by the system resolver leads to a TLS verification error (step 6); (3) disabling TLS certificate verification allows us to fetch a certificate signed by Allot (step 8); (4) the IP address returned using DoH (step 5) is the same returned by the system resolver (step 4) and used in step 6 (i.e., `67.213.76.19`). We thus confirm TLS interception of the WoW website possibly using technology developed by Allot.

4.3.9 Circumventing DPI blocking

We were able to circumvent the DPI blocking by adding the tab escape character ($\backslash t$) to the basic HTTP get request headers. Another technique to circumvent the DPI blocking is by delaying the transmission of the HTTP get requests, as mentioned in [109] where they circumvented DPI blocking websites with information related to the Catalan referendum in 2017. This is another indication that the ISPs are using the same blocking infrastructure throughout periods for blocking of different content and by different authorities.

4.3.10 Reproducibility

Our research is reproducible and can be replicated to obtain our dataset and results. All parts of our data analysis including the heuristics used to analyse the network measurements as well as the source code developed during our experimental testing methodology to overcome previous limitations of OONI Probe's software as well as the OONI meta database is made publicly available and online under a free and open source software license [34, 267].

4.4. Discussion

In this research, we observed that the websites related to the controversial Catalan referendum were blocked with the common blocking techniques. We were able to detect 16 unique blockpages, identify 2 DPI vendors (Fortigate and Allot) and a total of 78 websites being blocked. For an overview of the blocked websites and reproducibility, we compiled a matrix of all the blocked websites in Table 4.2. To the best of our knowledge, this is the only empirical study that provides a complete list of blocked websites in Spain. None of the blockpages contained any information on a law order or blocking reasons. Spanish authorities and ISPs appear to rather obfuscate the blocking information through misconfiguration of the blockpages as if the websites were not blocked but rather unreachable or erroneous. Nevertheless, being transparent about the blocked websites, also by issuing blocklists, may help to reduce over-blocking, unintended blocking or collateral damage [275] such as the blocking of an expired domain name registered from a different entity. Starting from the date 2017-09-25, we found an increase of network measurements in OONI data. The ISPs might have been preparing to block all websites related to the Catalan referendum. [177, 213] report that the Spanish court deemed the Catalan referendum of October the 1st 2017 illegal and the Spanish government attempted to stop the referendum voting by blocking access to websites, raiding the offices of the.cat Internet registrar, seizing domain name sources, and removing an application from Google Play Store. [177] also identified DNS manipulation and HTTP blocking predominantly used to censor Catalan referendum sites.

In line with these findings, we revealed the same websites in the non-DNS analysis, with more blocked websites for file sharing, video streaming, IPTV links, the gateway of IPFS (*gateway.ipfs.io*), WoW website (*womenonweb.org*) and the ex-website of the National Liberation Army in Colombia (*eln-voces.com*) that expired almost a year ago and then was registered by another entity [26, 238] hosting unrelated content. We additionally detected multiple middleboxes (DPIs) also used to block access to websites. Prior research by the Opennet Initiave in 2005 identified Burma's (Myanmar's) repressive regime to use Fortinet's Fortiguard product for censorship and information controls of websites and services in Burma's ISP networks [147] — similarly to Telefonica ISP for blocking numerous websites in Spain, as analysed in Section 4.3. In analyzing past events, our study is limited by the historical OONI network measurement data. More accurate measurements and cross-correlations could be potentially achieved by combining with other data sources, however, all data sources evaluated in Section 4.2.1 did not have relevant network measurements available that could match our research requirements (see Sections 4.1 and 4.2.4). It is also worthwhile acknowledging that our manual checks to ensure that there are no false positives might have resulted in removing some blocked websites from the data set. Although the stated limitations did not prevent us from addressing our purpose, we leave these issues to future work. Compared to other network measurement studies on a larger amount of countries in their network interference practices [193, 211, 221, 263], this present study enabled deeper technical insights in the stated field. Further, our study demonstrated the possibility of feasible, effective, and verifiable research and conclusive results based on historical network measurement data.

4.4.1 Involvement of multiple authorities

Cascading censorship and blocking that involves different stakeholders illustrate how power dynamics form a hierarchy within the sphere of control of a nation-state authority. In our research, we identified numerous entities that can force ISPs to block access to Internet resources and perform information controls. The Spanish Civil Guard (Guardia Civil), the General Directorate of Police (Direccion General de la Policia), Judicial National Police (Judicial Policia Nacional), Gambling Authority (Dirección General de Ordenación del Juego). Furthermore, the anti-phishing security group of TSOL seems to also be able to decide which websites or services can be blocked, as discussed in Section 4.3.1.

4.4.2 Ethical considerations

In our research, we used only free software tools and datasets available under a free license (Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International). We neither conducted nor asked any entities to perform network measurements. The data collected by OONI probes was sanitized to remove any personally identifiable information. The OONI team provides all data specifications and methodologies as well as the source code of their software.

4.5. Conclusion

This study analyzes OONI historical network measurements in Spain in the 2016-2020 period. We provide strong evidence that the Spanish network blocking infrastructure originally introduced for enforcing copyright and gambling regulations was also used to control political information. We documented the blocking of several websites and services, including those related to the Catalan referendum. The website of the Catalan protest group Democratic Tsunami was also blocked. We also measured the blocking of a non-profit organization's website providing support to women, WoW. We additionally found that a previously expired domain name now registered under a new entity (*eln-voces.com*) was also blocked. Furthermore, we detected and listed all network interference techniques deployed by the Spanish ISPs, which included DNS manipulation and HTTP blocking with DPI equipment. We ascertain that both blocking techniques were consistently used by each ISP, at the same time in some cases not being labeled as such in a transparent way. Our research highlighted the importance of systematic, longitudinal network measurements in a geopolitical context EU that is often under-researched. This study could help policy regulators, lawyers, civil society organizations, ISPs, and other entities to understand whether and how blocking websites and network services occur in a given country or region.

– 5 –
Mobile app store censorship detection

5.1. Introduction

The widespread adoption of smartphones over the past decade saw an extreme rise in the development, distribution and usage of mobile applications: computer programs that are specifically designed for mobile devices. For brevity, throughout the paper we will be referring to them as applications or apps. Their use cases cover a wide domain ranging from entertainment to banking, medicine, education, and communication. Mobile applications are used by billions of people on a daily basis for both personal and business purposes.

The main distribution channels for applications are a handful of centralized platforms known as app stores. The characteristics of these depend on the OS of the mobile device. Most app stores take the form of an online store, that regulates free or paid applications, and distributes them in various countries. App stores are actually common in other platforms as well, such as in Linux distributions and in game distribution [283].

The main goal of this paper is to investigate the availability of apps in several app stores across different countries. We focus on the following three app stores: Google Play (i.e. Google Play Store) operated by Google LLC, iTunes Store (i.e. Apple App Store) operated by Apple Inc., and Tencent MyApp (i.e. Tencent App Store) operated by Tencent Holdings Ltd. The first two account for a large percentage of Android and iOS mobile devices, while the latter operates primarily in China and is the largest app store in that country. For Android-supported devices, there are further app stores with significantly less market share; for iOS devices, the Apple App Store is the only available app store.

A large part of our study focuses on Russia and China, two major mobile app markets. In order to estimate the inclusiveness of our study, we have gauged the percentage of the market share that the Apple App store, Google Play Store, and Tencent App Store occupy in respective markets. In China, Apple has the largest app store by downloads with 500 million downloads in March 2017 [260]. It is followed by the Tencent App Store that recorded 250 million downloads during the same month, which amounts to 25% of the Android market share. Tencent App Store is the leading Chinese Android app market accounting for 13.71% of the market share measured in monthly average users of the second of 2018 [21]. Tencent App Store is predominantly targeted at Chinese speakers and is hence available in Chinese language [255]. Tencent App Store is differentiating its terms of service between users that reside anywhere in the world and Chinese residents (regardless of the nationality), Chinese nationals (even if they reside outside China), and Chinese companies [254].

5.2. Related work

In this section, we present a literature review and classify related work to better highlight its importance and similarity to the studies in our paper. We structure this section into the following aspects: app store regulations, app store comparisons, and app store mining.

5.2.1 App store regulation

Hestres, Luis E. analyzes [130] Apple's guidelines and approval process, discusses content based rejections of apps, and outlines the consequences of this process for developers' and consumers' freedom of expression. It also argues for principles that guarantee *app neutrality* while also guaranteeing device safety and quality control. Síthigh, Daithí Mac [178] assesses the regulation of smartphone app stores and highlights the importance of forms of regulation that are not linked to a violation of competition law. *Developer-focused issues* deals with the relationship between Apple and app developers; three themes of Apple's Guidelines are identified (content, development and payments), and the ways in which control can be challenged (through jailbreaking, web apps and regulatory intervention) are scrutinized.

5.2.2 App store comparisons

Lim, Soo Ling et al. [169] conducted one of the largest surveys to date of app users across the world, investigating user adoption of the app store concept, app needs, and rationale for selecting or abandoning an app. They collected data from more than 15 countries. The analysis of data provided by 4,824 participants showed significant differences in app user behaviors across countries. Ghazawneh, Ahmad and Ola Henfridsson [118] provide a paradigmatic analysis of different app stores that helps to understand the relationship between application marketplaces, platforms and platform ecosystems. They generate a typology that distinguishes four kinds of digital application marketplaces; closed, censored, focused and open marketplaces. Seneviratne, Suranga et al. [239] propose a method to detect spam apps solely using app metadata available at the time of publication, according to a set of checkpoint heuristics that reveal the reasons behind their removal. Their analysis suggests that approximately 35% of the apps being removed are very likely to be spam apps. They map the identified heuristics to several quantifiable features and show how distinguishing these features are for spam apps.

Peltonen, Ella et al. [212] carry out an analysis of geographic, cultural, and demographic factors in mobile usage. Their research sample is gathered from 25,323 Android users from 44 countries and 54,776 apps in 55 categories, and demographics information collected through a user survey. Their paper reveals significant differences in app category usage across countries that reflect geographic boundaries. They demonstrate that the country category gives more information about application usage than any demographic, with geographic and socio-economic subgroups in the data. Albrecht, Urs-Vito et al. [11] propose SARASA, a semiautomatic retrospective app store analysis, which provides a step-by-step filtering of apps by formal criteria. A full survey of the metadata of 103,046 apps from Apple' s German App Store in the Medicine and Health & Fitness categories was carried out.

5.2.3 App store mining

Martin, William et al. [182] studies information about applications obtained from app stores.

Group	Countries
CG1	Google Play unavailable: Syria, North Korea
CG2	Abshar, Orbot, Signal and Skype unavailable in Google Play: South Sudan, Sint Maarten, Bonaire, Sint Eustatius, Saba
CG3	96 countries where iTunes is officially not available
CG4	All of the apps in our test list unavailable in iTunes: Cocos Island, Christmas Island, Guam, Heard Island and McDonald Islands, British Indian Ocean Territory, Kiribati, Marshall Islands, Northern Mariana Islands, Norfolk Island, Nauru, United States Minor Outlying Islands

Table 5.1: Country Groups with regards to how and why they are censored

Their survey describes and compares the areas of research that have been explored thus far, drawing out common aspects, trends and directions future research should take to address open problems and challenges. Fu, Bin et al. [111] propose WisCom, a system that can analyze millions of user ratings and comments in mobile app markets at three different levels of detail: (a) discovery of inconsistencies in reviews; (b) identification of reasons why users like or dislike a given app through an interactive, zoomable view of how users' reviews evolve over time and (c) identification of users' major concerns and preferences of different types of apps. A limitation to their study is that they are analyzing apps that are only available on Google Play Store.

Chen, Ning et al. [51] propose an app review mining framework performing comprehensive analytics from raw user reviews by (i) first extracting informative user reviews and by filtering noisy and irrelevant ones, (ii) then grouping the informative reviews automatically using topic modeling, (iii) further prioritizing the informative reviews by an effective review ranking scheme, (iv) and finally presenting the group-s of most "informative" reviews via an intuitive visualization approach. Their limitation is that they have used only four Android apps to evaluate their review mining framework. Tang et al. [248] conducted an empirical study of a large-scale set of fake apps. They have collected over 150,000 samples of popular applications and performed a quantitative study of fake samples and fake authors' developing trends.

5.3. Methodology

The focus of this paper is the availability of mobile apps in major app stores (Google Play Store, Apple App Store, Tencent App Store) across different countries.

In order to understand which applications are unavailable in a country, we collected data by querying the official search engines of the app stores for a given term and for specific apps (see Section 5.3.1). So as to eliminate false positives, we identified the countries where Google, Apple and Tencent operate their app stores and the reasons why a country might be excluded (see Section 5.3.2). For that, we conducted research on the terms of service of these companies, their official online documentation, their user forums, and we got in contact with their support teams and app developers. Moreover, we referred to news sources for building a test list with candidate apps that are unavailable in specific regions, and for interpreting our findings.

Throughout the paper we cluster countries with regards to the way and the reasons why they are censored, in four Country Groups, as they appear in Table 5.1.

App Store	Search for a term	Lookup App Details
Google Play	https://play.google.com/store/search?c=apps&q={term}&gl={country}	https://play.google.com/store/apps/details?id={Google Play Id}&gl={country}
Apple App Store	https://itunes.apple.com/search?term={term}&country={country}	https://itunes.apple.com/lookup?id={iTunes Id}&country={country}
Tencent MyApp	https://android.myapp.com/myapp/searchAjax.htm?kw={term}	https://android.myapp.com/myapp/detail?apkName={MyApp Id}

Table 5.2: The URLs for querying the app stores for a term and for the details of a given application. {term}: the search term, {Google Play ID/iTunes ID/MyApp ID}: the ID of the app in each platform, {country}: the country code

5.3.1 Querying mobile app stores

For our research we made queries directly to the three app stores we are targeting. All of them can be accessed in two ways: either via their mobile application, or via the website they maintain. The respective query URLs (see Table 5.2) do not require authentication and they support searching for a term and retrieving details for a given application. To that extent, they can be used for future research on app stores, since they provide additional information including application categories, the number of installations of an app, its review ratings, price, developer and release information.

In addition, the query URLs support parameters for filtering the results. Both Google Play and iTunes provide a way to query the app stores in different countries, by including the desired ISO 3360 country code in the query parameters [155]. We take advantage of this feature to find out what the query results would be for users in a specific country.

Parsing the results of the queries can be simplified with scraper scripts. In our source code repository [237] we have uploaded wrappers for the *google-play-scraper* and *app-store-scrapper* node.js packages, together with a script that queries the app stores in different geographic areas. Also, we have uploaded the corresponding results for the term “vpn” and for apps that are reported to be censored (such as Psiphon and the Onion Browser).

5.3.2 App store operation across countries

In order to minimize false positives on app availability per app store and per country, we consulted the list of supported countries each company is operating their app store in. Currently we only take into consideration the Apple iTunes and Google Play app stores, as they are the only ones with an official country availability list on their websites, and since the Tencent MyApp store is primarily targeting users in China.

Google Play

According to Google’s documentation, the supported locations for distribution to Google Play users are listed [242] to a total of 144 countries, and a “Rest of the World” category. The actual countries included in the “Rest of the World” category, or at least their number, is not publicly available. In addition, the available information does not clarify the reason why the Google Play Store is not operating in a country; whether that is due to sanctions, regulations, or company and in-country policies.

By looking on the crawled app data for Google Play we found out that Play Store is not available in Syria and North Korea (ISO 3360 country codes Syria (SY) and North Korea (KP)). That brings us to believe that Google Play Store is most probably not operating in these countries, whereas in other countries with US sanctions we were able to retrieve app information. For these two countries (SY and KP) we are going to use the group notation Country Groups (CG) 1 throughout the paper.

App Name	Google Play Id	iTunes Id	MyApp Id	Unavailable countries
Abshar	sa.gov.moi	1004966456	N/A	CG1, CG3, CG4
I2P	net.i2p.android	N/A	+	
LinkedIn	com.linkedin.android	288429040	+	CG1, CG3, CG4, CN, RU*
New York Times	com.nytimes.android	284862083	N/A	CG1, CG3, CG4
Onion Browser	N/A	519296448	N/A	CG1, CG3, CG4, CN
Orbot	org.torproject.android	N/A	N/A	CG1
Psiphon	com.psiphon3.subscription	1276263909	N/A	CG1, CG3, CG4, CN
Signal	org.thoughtcrime.securesms	874139669	+	CG1, CG3, CG4
Skype	com.skype.raider	304878510	N/A	CG1, CG3, CG4, CN
Shadowsocks	com.github.shadowsocks	N/A	N/A	CG1
Keyword Search: VPN	CG1, CG2	CG1, CG2, CN	CN	

Table 5.3: An indicative list of queries for apps and keywords in different countries. Includes app identifiers in each app store.
+: Same app ID with Google Play
*: Unavailable in Google Play Store and Apple App Store, CG1-4: Country Groups (see Table 5.1)

Apple iTunes

The Apple App Store provides a list of the countries it is available in, which is more transparent than the Google Play Store list. We found Apple App store to not operate almost on an entire hemisphere, i.e. 96 countries [20]. We use the notation CG 3 for these countries. At first we thought that we misunderstood something or that we overlooked a website where more countries are listed. For this reason we contacted the Apple App Store customer support that directed us to the same page that we base our findings on [20], along with information on how one can change one's Apple App Store country. Upon further inquiries they were not able to provide us with more information on why these countries are blocked from operating Apple App Store.

Some restrictions may apply because of export restrictions. This is why stores are not available in Iran, North Korea, and Syria for instance. In some countries like Serbia, the Apple App Store is not available for legal or commercial reasons. This could be due to practical matters, such as Apple not having a registered legal entity in Serbia, or perhaps even because the sale would most likely be initiated in foreign currency. Maybe for that reason the Serbian Google Play Store shows prizes only in US dollar and not in the local dinar. Countries like Morocco and Rwanda (where the Apple App Store) that are not available may also fall in this category. But again we cannot be sure given the limited information available to the public.

5.4. Findings of mobile app store censorship

For our measurements we created a list comprised of free and open source censorship circumvention, anonymity and messaging mobile apps; Invisible Internet Project (I2P), Psiphon, Onion Browser, Orbot, Signal, and Shadowsocks. In order to verify claims about blocked or otherwise unavailable apps in different countries we added Skype (a widely used voice application), the New York Times news app, LinkedIn social network app and the controversial Abshar app developed by the government of Saudi Arabia.

In Table 5.3, we list the app names together with their IDs (when available) in all three app stores along with the unavailable list of countries per app. All apps on Apple App Store were not available (apart from the other 96 unavailable countries) in Cocos Islands (CC), Christmas Island (CX), Guam (GM), Heard Island and McDonald Islands (HM), British Indian Ocean Territory (IO), Kiribati (KI), Marshall Islands (MH), Northern Mariana Islands (MP), Norfolk Island (NF), Nauru (NR) and United States Minor

Outlying Islands (UM). For the aforementioned countries we are going to use the group notation CG 4. Onion Browser, Psiphon and Skype apps were found unavailable in China, whereas LinkedIn app was unavailable in both Russia and China. These findings confirm the multitude of reports and new Chinese regulations that ban unlicensed VPN providers or censorship circumvention apps to operate in China [278]. LinkedIn is unavailable in both Google Play Store and Apple App Store in Russia, because the government banned the company's app and website [229].

Apart from countries in CG1, while performing a full text search for the queries abshar, orbot, signal and skype we got an empty response for the countries South Sudan, Sint Maarten and Bonaire, Sint Eustatius and Saba, we are going to refer to these countries as CG 2.

Tencent App Store seems to have a strict policy on VPNs, since neither keyword searches nor a search of the specified app list (see Table 5.3) yielded results. We found I2P to be available; an anonymous peer to peer network and that could be confirmed by the relative high number of I2P nodes in China [132]. Similarly The LinkedIn app, which is blocked in China and Russia is available, as well as the Signal instant messaging app. All apps share the same app ID as in Google Play Store.

In our analysis we cover around 50% of the app store market share in Russia and around 30% in China. Censoring an app on Apple App store or Tencent App Store in China will have a larger effect than on Google Play Store, which is barred from the market in China [248]. Similarly censoring an app from Google Play Store in Russia has a greater effect than excluding an application from the Apple App store, due to Google's larger footprint within the country.

5.4.1 Verification

We cross-checked our results from different vantage points. We have uploaded the results of our queries for the terms "VPN", "proxy", "代理", "🄵擬私人網路", "🄵拟私人网络", "私人互联网接入", "规避", "circumvent", "专用网络", "互联网网络", "加密通讯", "翻墙", and mobile applications that are being censored in Table 5.3 in our git source repository [237].

We were able to cross-verify our methodology (presented in Section 5.3) with publicly available data from OONI; a free software project which collects and processes network measurements with the aim of detecting network anomalies, such as censorship, surveillance, and traffic manipulation [105]. OONI data show that the LinkedIn app on Google Play Store[1] was not accessible (returned an HTTP Code 404) [2].

By using OONI's API [3] we were able to obtain all anomalous measurements for the Google Play App Store URL of LinkedIn. Specifically we identified at least 52 different network vantage points of Russian autonomous systems; 12389, 12714, 12790, 12958, 15378, 16345, 20807, 21367, 24588, 25159, 25490, 25513, 28745, 28812, 29226, 29497, 31133, 31213, 31376, 31430, 3216, 3239, 3253, 34533, 35807, 39289, 41661, 41682, 41691, 41733, 42610, 42668, 43595, 44640, 47395, 48092, 48190, 48642, 49478, 51035, 51570, 51604, 5429, 5563, 56330, 56377, 6856, 8359, 8369, 8402, 8427 and 8595. The queries to the API and tools to extract the ASes can be found in our repository [237].

[1] https://play.google.com/store/apps/details?id=com.linkedin.android
[2] https://archive.fo/dWPiS
[3] https://api.ooni.io

5.5. VPN mobile app regulations in Russia and China

In the following section we compare the law on VPNs as well as availability of the top global VPN providers within China and Russia. These two case studies provide an in depth examination of two major app markets and are an addition to the large N comparison of app availability worldwide.

The law on VPNs had been quite loose for the past few years. In 2017, however, China restricted VPN usage more seriously. This is due to a *Notice of the Ministry of Industry and Information Technology on Clearing up the Market for Regulating Internet Access Services* and China's 2017 cyber security law [293, 294]. In essence, only government approved VPNs are allowed [53]. On this legal basis, China requested Apple to remove 674 VPN apps from its app store [22], [52]. In turn, VPN providers received letters from Apple saying that the content they provide is illegal in China and consequently their application had to be removed [23]. We conducted a keyword search of "VPN" on the Russian and Chinese Apple App Stores and manually verified the results. Apple's app store in China returned 54 results, which included not only the title of the app but also content that describes an app. A manual search through the results showed that less than 5 apps are actual VPN apps. It is expected that the VPNs available are government approved and surveilled ones. In most countries the number amounts to around 200 returns in results. VPN availability on the Apple App store in China is consequently low. None of the major foreign VPNs were available for download in China. It is expected that domestic alternatives are available, but these have poor privacy policies and are expected to share data with the government [185]. We also searched the Apple App Store in China for "proxy", "代理", "🄵擬私人網路", "🄵拟私人网络", "私人互联网接入", "规避", "circumvent", "专用网络", "互联网网络", "加密通讯", "翻墙", and evaluated the results one by one. These searches provided many results that matched the keyword searches. However, they contained less than five apps in total that provide VPN services.

Russia for its part instituted restrictions on VPN usage in 2017 through its amendments to the *Law on Information, Information Technologies and Information Protection* [102]. The law stipulates that VPNs are allowed but they have to make sure that no censored websites are accessed through them [284] and that they have to share user data with the Russian government [231]. The law also states that search engines have to delete VPN service related results from its services or else they will be fined [230]. In 2018 Russia went a step further and banned 50 VPNs and censorship circumvention tools [232] that allowed users to access the Telegram messaging app, which was also prohibited by the government [133].

Despite similar Russian and Chinese laws, VPN apps are still available on app stores in Russia. A keyword search for "VPN" returned 199 results, including major VPN providers. This shows that the ban is not as thoroughly implemented in Russia. This loose implementation was also reported on by Russian news and observers. The environment in Russia may be less strict because of technical challenges with the implementation, such as difficulties with distinguishing VPNs that are used by private or commercial VPNs, or problems with forcing foreign VPNs into compliance with the law. In China the environment is much harsher with almost no VPNs available in the Apple and Tencent App stores.

5.6. Third-party app stores

A significant factor that contributes to mobile app censorship is the strategy of mobile operating system developers (Google for android, and Apple for iOS) to maintain the monopoly of the app ecosystems through the lockout of third-party app stores. Centralization of app distribution to a handful of app stores makes it easier for governments to block specific apps. According to court cases against Google and Apple for allegations under anti-competitive actions legislation [97, 144], their motivation for obstructing alter-

native app stores is mainly financial, since they profit from commissions on app purchases, monetization of user analytics, and the promotion of their own services. On the other side, this strategy is justified as a security safeguard, because there can be no guarantees of the validity of the applications distributed by independent app stores.

From the point of view of app availability, third-party app stores are appealing to users because they do not enforce country-specific censorship. For example, users in Russia, where LinkedIn is not available on Apple App Store and Google Play Store (see Section 5.4), can still download the app from AppAddict [170] and Aptoide [171] respectively. However, in order to do that, they have to manually degrade their security, e.g. by modifying their settings to allow the installation of software from unknown sources, or even by jailbreaking their devices and thereby voiding their warranty. Furthermore, they have to trust third-party app stores to deliver the genuine applications and that they will make updates available in the future. In general, independent app stores bring freedom to users and developers, and contribute to a healthy software ecosystem. However, they are less regulated, do not always collaborate directly with the developers of the apps, and their revenue models are unclear [157].

5.7. Conclusion

Given the centralized nature of app stores one may think that finding out whether a specific app is available in their country is straightforward. It turns out that this question is not so easy to answer. Unfortunately, there are no transparency reports on app availability, and none of the app stores that we study provides information about which apps are censored and for what reasons. Also, it is unclear how users will receive updates on installed apps that get banned.

Our report presents a methodology for querying the major app stores – Google Play Store, Apple App Store, Tencent App Store – to find out whether (a) they are operating in a country, and (b) whether an app is available in that country. In that way, we were able to identify geographical regions where app stores are not available. Moreover, we collected evidence of unreported censorship, specifically for censorship circumvention, anonymity and messaging mobile apps. Furthermore, we took a closer look into VPN mobile app availability in China and Russia. The environment in China is much stricter than in Russia, with none of the major VPN apps being available in the Apple App Store and Tencent App Store. Russia is looser with its implementation of VPN restrictions on app stores. As a countermeasure, users can download censored applications via third-party app stores, but this can potentially degrade their security and privacy.

5.8. Recommendation

Our recommendation is that app store companies ought to launch app transparency reports. In 2018, Apple announced that future reports would include information on government removal requests of apps. However, to the best of our knowledge Apple has not yet followed up on its promise [48]. Google has no mechanism to publicly report take down requests of apps either.

Our paper highlights how opaque the global app market is and that companies such as Google, Apple, and Tencent need to become more transparent about their operations. We recommend them to introduce app transparency reports that would include the identification of apps that were removed due to government requests and the reasons for their removal. Furthermore, they should indicate why their app store is not available in certain countries, and whether this is for political, financial, or legal reasons.

– 6 –
An overview of website blocklists in European Union

6.1. Introduction

In setting up the infrastructure for monitoring and network blocking to adhere to the EU legislation for preventive dissemination measures of terrorist content, gambling regulations, copyright enforcement, tobacco and health website regulations, extremism, phishing, and hate speech [19, 291], EU member states have made it easier to block websites and services and to monitor information. Here, ISPs and network operators are (often) required to set up blocking infrastructures. Permitted practices concerning traffic management that can involve filtering by ISPs are regulated at the first stage by the Open Internet Regulation (EU) 2015/2120 [98]. In respect to Open Internet principles, network traffic interference practices such as blocking, slowing down, altering, restricting, degrading or discriminating between specific content, applications, services, or specific categories of content, applications, or services are not in principle allowed. They are subject to justified and narrowly defined exceptions in the law. Article 3(3) of the Open Internet Regulation sets the framework for such activities in the EU. In this regard, the European regulator BEREC has provided guidelines for the implementation of the Open Internet Regulation that have laid down the exceptions in which ISPs may implement such traffic management regulations [98]. However, the evidence provided in this paper demonstrates a lack of transparency in the ways in which network interference is conducted in the EU countries. Although website blocking is a current activity, regulators have not provided enough evidence on how such blocking is being conducted by the telecommunication operators. EU member states not only use blocklists as a means of blocking access to websites but also block different types and categories of websites and services that are not included in the publicly available (identified) blocklists.

Relevant studies have documented Internet censorship in non-European countries, as well as usage of such infrastructures for other political motives [159, 213]. We define Internet censorship as the practice of using any kind of hardware or software to prevent users from accessing websites or services through network interference or information control. In recent years, further studies have been conducted, which have drawn attention to online network interference and Internet blocking in individual countries of the EU [3, 44, 234, 236, 269, 272, 276]. For instance, the ”Open Net Initiative” report mentions nearly 50 countries that practice Internet censorship [3]. To the best of the authors' knowledge, no analysis of network interference in all EU countries has been performed. examines website blocking practices in the EU.

With regard to the blocking of websites as a current activity under insufficient documentation on how such blocking is carried out by the telecommunication operators, this research examines how network interference is conducted in the EU countries, to what extent EU member states use blocklists as a means of blocking access to websites and what different types and categories of websites and services are affected

by these practices that are not included in the publicly accessible (identified) blocklists.

6.1.1 Contributions

This study provides three main contributions: The study contributes by conducting a comprehensive analysis of the 27 EU countries[1], based on three different sources. These include, first, tens of millions of historical network measurements collected in 2020 by volunteers from around the world; second, the publicly available blocking lists used by EU member states; and third, all reports of all blocked websites issued by each country's network regulators.

The analysis of 27 EU countries is based on ten million historical network measurements collected during 2020 by OONI volunteers around the world [207]. OONI is an organization that develops software to perform network measurements. OONI also administers the server infrastructure to store these data in a database (see Section 6.3.1), from which data can be retrieved for further analysis, for instance to identify cases of Internet censorship or to detect surveillance network equipment. Over the years, different types of methodologies have been developed to detect filtering or blocking of network resources, tampering with communication channels, and intentional manipulation of network routes. These types of blocking methodologies can be evaluated with network measurements: data contributed to OONI gathered by anonymous volunteers from each country who use software probes [207]. These data depict a rigorous perspective of the actual network filtering or content blocking that occurs in a specific network. Network measurements are challenging to conduct as they are deployed from vantage points that either probes have access to, or are located within the underlying network being measured.

This research also lists and catalogs publicly available blocklists in the EU. The blocklists are used by EU member states to block access to websites or services. In the early 2000s, the EU issued regulations blocking access mainly to online gambling services that were not licensed by all EU member states. Contrary to other services, the EU has constrained online gambling operators to operate in each EU country by paying a licensing fee to each EU member state in which they provide online services. One of the ways to enforce this regulation was to issue website blocklists of the unlicensed gambling websites and oblige ISPs to censor them in their networks. This is one of the first instances of EU-wide website blocking that drove ISPs to create a filtering infrastructure in their networks, frequently with many inconsistencies, over-blocking and under-blocking websites [276]. Lately the censorship of websites has increased and more categories have been added to the blocklists ranging from streaming websites, subtitles, file sharing, and torrents to tobacco, health, and medicine information resources, as discussed in Section 6.5.

Finally, this paper reviews and provides a summary of the reports issued by the NRA of each EU member state with information concerning network interference such as website blocking. Other institutions than the NRAs in each country may also regulate networks there.

6.1.2 Structure

The paper is structured as follows. First, after related research is described in Section 6.2, Section 6.3 presents some essential foundations of this research, explaining the OONI architecture and network measurements in detail. Section 6.4 describes our methods for collecting and analyzing the network measurement data used in this study. Section 6.5 presents the results of our overall data analysis. We discuss

[1] Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden.

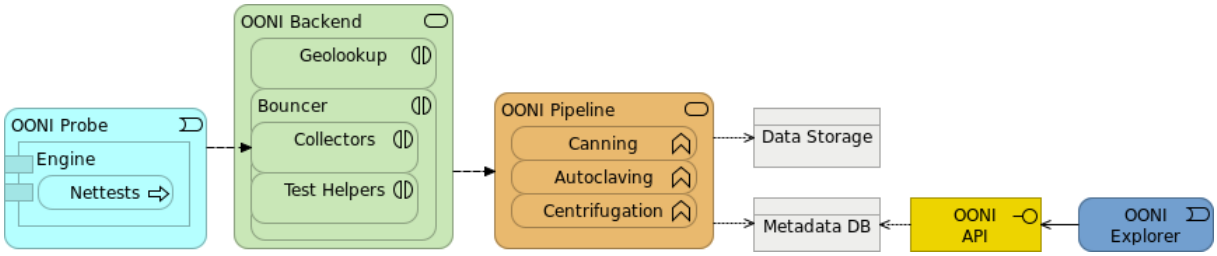

Figure 6.1: OONI high-level architecture diagram

current challenges, point out avenues for further research, derive practical implications, summarize our findings, and conclude in, Section 6.6.

6.2. Related research

Relevant research from previous academic studies has shown that censorship exists in many countries such as China [50, 54, 83, 95, 134, 136, 160, 161, 162, 163, 164, 176, 181, 192, 209, 228, 285, 286, 288], Thailand [116], Bangladesh [186], Pakistan [3, 189], India [122, 289], Iran [16, 17, 28], Syria [49, 233], Turkey [246, 247], Russia [222], and Mexico [154]. A few studies have looked at network interference and Internet blocking in the EU context [44, 234, 236, 269, 272, 276]. To the best of the authors' knowledge, there is no previous research analyzing network interference in all EU countries, specifically related to website blocking.

6.3. Foundations: OONI architecture and network measurements

OONI data are publicly released and provided as an open access dataset, available under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license [103]. OONI provides the blocking detection methodologies used by its software in the public domain for review, experimentation, and potential improvements by the community.

OONI software is released under a free license (GNU General Public License v3.0) and is publicly available for downloading, running, further distribution, modification, and improvement. Open methodologies build a capable and strong community of researchers, activists, policy advocates, hackers, data scientists, and others interested in researching Internet censorship. Having open methodologies and public access to the source code allows the community and volunteers to contribute to network measurements and make informed decisions about the potential privacy risks associated with the use of OONI software. In addition, such methodologies increase transparency regarding the validity of collected network measurements and allow a better understanding of the technical implementation and technical details of the lower level. A high-level diagram of the OONI infrastructure and software is shown in Figure 6.1.

The engine is the part of the software that runs the network measurements (nettests). OONI provides probes to perform the nettests. The probe software for mobile or desktop clients is based on different software implementations depending on the platform. Each probe (client) implementation uses a specific software architecture. The applications for mobile devices are developed in Java for Android (probe-android) and Objective-C for iOS (probe-ios). The desktop clients are developed in Go for the command-line interface (probe-cli) and JavaScript for the desktop applications of MacOS, Windows, and Linux (probe-desktop). The legacy implementation (probe-legacy) for the desktop clients (still used despite its

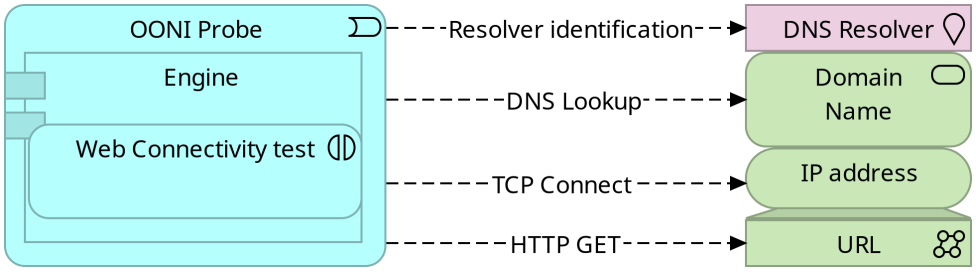

Figure 6.2: OONI Web Connectivity test diagram

legacy status) is being developed in Python.

6.3.1 OONI backend

A typical transaction of an OONI probe to the backend consists of the following steps: i. the probe requests the available collectors and test helpers from the bouncer; ii. the probe performs a geolocation lookup to find out its IP address (deduced by default for privacy considerations) and determine the AS number, the country code, and the name of the network entity owning the AS; iii. the probe opens a report for the nettest; iv. upon completion of the nettest, the probe submits the results to the collector as a JSON file.

Once the results have been submitted, they are sent to the OONI pipeline for archiving and further processing of the network measurements (reports). The pipeline aggregates the data (reports) submitted by the probes (network measurement clients) to the backend. Upon receiving the unprocessed reports, the pipeline performs the following steps: i. canning - compacts the reports to occupy less disk space and helps to reprocess the reports faster; ii. autoclaving - sanitizes and normalizes the report data, removing potential personally identified information and fixing inconsistent data formats; iii. centrifugation - aggregates the important parts of the reports and stores these metadata to a database for further processing.

Powered with data from the metadata database, the API allows analysis of data collected from OONI probes. This component is based on the Open API specification and is extensively documented. Finally, OONI Explorer [204] provides a visual representation of all OONI data and allows performing quick queries with various constraints such as (nettest, country, URL, and date) in an easy and graphically visual way without the need to download any data or use the API.

6.3.2 OONI methodology

In our research, we analyze network measurement data performed by the Web Connectivity OONI nettest [203]. This test measures the reachability and possible blocking of a website given an IP address or a domain name. The test's methodology diagram is illustrated in Figure 6.2. The Web Connectivity test consists of the following steps: i. performing an A DNS lookup and storing the results of the A records list, ii. attempting to establish a TCP session in either port 80 or 443 (depending on the URL scheme), iii. performing an HTTP GET request to the path specified in the URI. The responses and possible errors from each step are recorded in a JSON file and submitted to the OONI network measurements collector for further processing and archiving.

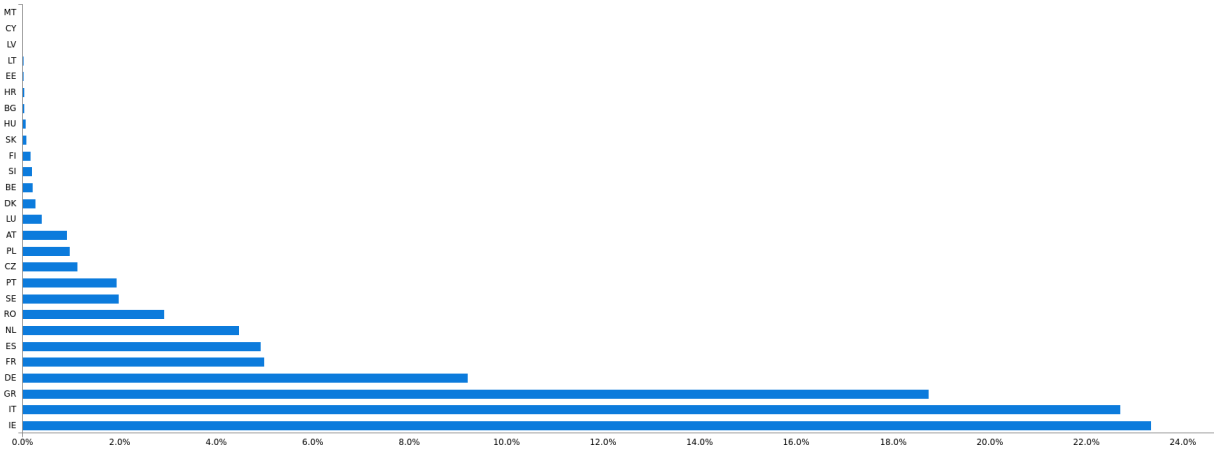

Figure 6.3: Data distribution per country code (Alpha-2 ISO 3166)

6.4. Methods of network analysis

In this study, we draw on historical OONI network measurement data. Using custom-built database queries, we were able to collect more than ten million relevant network measurements. We then created a meta database [104] in PostgreSQL to ease the workload of collecting, cleaning, and organising our dataset. We used the Jupyter Notebook software tool and the Python programming language to collect, process, clean, categorize, and analyze the OONI data. For the blocklists we performed web data scraping to extract the blocklists from the publication websites, PDF files, and documents in other file formats, as the released blocklists are not systematically distributed. Moreover we conducted interviews and requested public blocklists and information related to the blocking of websites or services from multiple authorities via email communication (see Section 6.5.2 for more details).

6.4.1 Criteria and distribution of data

In total, we analyzed almost 1 million unique network measurements (specifically 999,125) from 888 distinct ASes in 27 countries. The data distribution across networks and countries is not uniform. Some network measurements were submitted by volunteers at random intervals but many were submitted with consistent frequency. Figure 6.3 represents the distribution percentage of the analyzed network measurements per country. We use the Alpha-2 country code notation as described in the ISO 3166 international standard. Our data analysis criteria were the following: i. network measurements present in the OONI meta database; ii. data collected in the date range: 2020-01-01 to 2020-10-20; iii. data flagged as anomalous (with signs of network interference); iv. network measurements conducted from networks within the EU; v. network measurements performed with the Web Connectivity test.

6.4.2 Data collection

To access the OONI data, we set up a PostgreSQL replica of the OONI meta database [104] and we fetched the latest archived data required for a database cluster [100]. It took about 10 days to sync with the master database and required 800 GB of storage capacity to accommodate the OONI meta database. A helper script was used to fetch the OONI S3 bucket data and configure the PostgreSQL server as a replica (in a hot standby configuration). This script fetches the latest archived meta database replica instance using all the available CPUs for decompression. The main requirement of the replica is a system with enough storage

capacity and network connectivity to host a PostgreSQL database. A description of the Web Connectivity OONI test methodology is provided in Section 6.3.2, and the test diagram is illustrated in Figure 6.2.

6.4.3 Data validation

We use the term blockpage to refer to an instance of deliberate blocking. The term has been, and sometimes also still is, used to refer to the error message displayed. From among the many network measurements with signs of network interference, we only included as blockpages those cases of which it could with some certainty be verified that they were neither false positives (for instance due to network connectivity errors) nor blocked due to internal network filtering rules (such as parental controls, antivirus filtering, or firewalls). For this, we derived a set of heuristics from certainly blocked instances and excluded all network measurements unless they satisfied the specified criteria: i. existence of a blockpage or any indication of an error due to blocking (e.g. HTTP error codes); ii. existence of DNS records that point to bogus IP addresses (such as 127.0.0.1); iii. network measurements with correct AS information (i.e. if the probe's AS number is not shared, *AS0*).

6.4.4 Blockpage heuristics

Network measurements that present signs of network interference (anomalous data) are not always evidence of website blocking. In fact, it is quite common to find anomalies in network measurements due to transient network errors, website misconfigurations, geolocation blocking, or simply software issues and bugs. For this reason, we developed a number of heuristics to identify website blocking by manually looking into the dataset, and verifying that is indeed a case of website blocking. We accept that website blocking has occurred when all the criteria set during the data validation process (see Section 6.4.3) are satisfied.

In addition to the data analysis criteria of Section 6.4.1 and the data validation criteria of Section 6.4.3, there is an additional test that a network measurement must satisfy for us to consider it to be a blockpage. On the validated data set we compare if the DNS A record of the website (IP address) is on the same AS as the one in the probe's network performing the measurement. This helps to detect the blockpages hosted within the same ISP or IP address ranges of the country. This is common and usual practice as it is unlikely that a website is hosted on the same AS as the one where the network measurement has been conducted.

6.4.5 National Regulatory Authorities' monitoring and reporting on Open Internet

As an obligation imposed by art. 5(1) of the Open Internet Regulation, the NRAs should annually inform the European Commission about their activities in monitoring and enforcing the Regulation's rules. The reports would serve as summaries for the Commission on the state of affairs in national jurisdictions and would serve to provide a minimum level of transparency and comparability of the implementations across Europe. Among the things expected to this end from the reports are the overall description of the national situation regarding network neutrality, the description of the NRAs' monitoring activities, the number and types of complaints, ISPs' infringements related to the Regulation, and results of surveys, evaluations, and technical measurements implemented by the NRAs. The reports from the NRAs should present any network blocking or network neutrality issue to the European commission based on the Open

Country	Report on Blocking
Austria	Network blocking due to copyright law, Sec. 3.4 (part II) [40]
Belgium	No cases of service or application blocking, Sec. 5.117 [217]
Bulgaria	Blocking in accordance with national legislative acts, Sec. 1.2 [1]
Croatia	None mentioned [127]
Cyprus	None mentioned [60]
Czech Republic	None mentioned [198]
Denmark	42% of ISPs indicated they block access to Internet, Sec. 4.1 [8]
Estonia	None mentioned [219]
Finland	None mentioned [194]
France	None mentioned [24]
Germany	An ISP blocked certain domains via DNS due to court ruling, Sec. 3.1.2 [43]
Greece	Gambling and copyright blocklists, DNS and port blocking, Sec. 4.1.1 [250]
Hungary	None mentioned [183]
Ireland	Website blocking might be in place at a number of ISPs in April 2021, Sec. 27 [61]
Italy	None mentioned [7]
Latvia	None mentioned [59]
Lithuania	None mentioned [227]
Luxembourg	None mentioned [225]
Malta	Ongoing investigation of IP blocking, Sec. 4 [31]
Netherlands	None mentioned [62]
Poland	Blocking traffic due to obligations under Article 15f(5) on gambling, and preventing access to websites using domain names published on the blocklist maintained by Cert Polska [90]
Portugal	None mentioned [12]
Romania	ANCOM was given powers to issue decisions to block specific online content or websites presenting false news about COVID-19, and issued 15 blocking orders [14], Sec. 1.1 [15]
Slovakia	ISPs block access based upon the European or national legislation; in the event of spreading illegal content, applications or services, or gambling websites without a Slovak license, were blocked, Sec. 2 [92]
Slovenia	None mentioned [91]
Spain	Blocking of websites by request of the courts only, Sec. 3.2 [251]
Sweden	None mentioned [216]

Table 6.4: National regulatory authorities reports overview

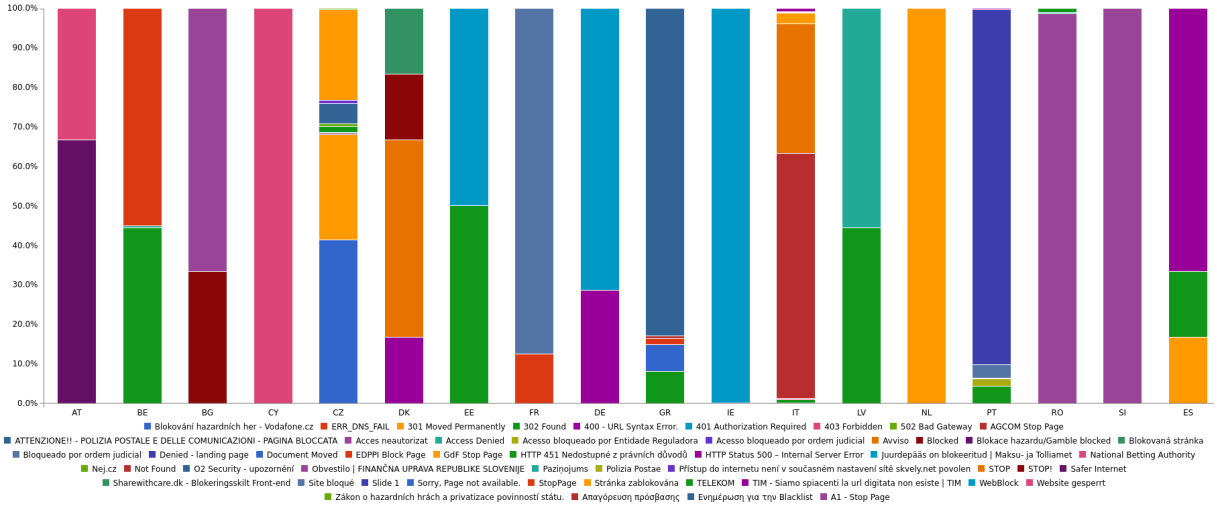

Figure 6.5: Blockpages per title and country code (Alpha-2 ISO3166)

Category Description	Code	Description
Anonymization, and circumvention tools	ANON	Used for anonymization, circumvention, proxy-services, and encryption.
Communication Tools	COMT	Sites, and tools for individual, and group communications. Includes webmail, VoIP, instant messaging, chat, and mobile messaging applications.
Control content	CTRL	Benign or innocuous content used as a control.
Culture	CULTR	Content relating to entertainment, history, literature, music, film, books, satire, and humour.
E-commerce	COMM	Websites of commercial services, and products.
Economics	ECON	General economic development, and poverty related topics, agencies, and funding opportunities.
Environment	ENV	Pollution, international environmental treaties, deforestation, environmental justice, disasters, etc.
File-sharing	FILE	Sites, and tools used to share files, including cloud-based file storage, torrents, and P2P file-sharing tools.
Gambling	GMB	Online gambling sites. Includes casino games, sports betting, etc.
Gaming	GAME	Online games, and gaming platforms, excluding gambling sites.
Government	GOVT	Government-run websites, including military sites.
Hacking Tools	HACK	Sites dedicated to computer security, including news, and tools. Includes malicious, and non-malicious content.
Hate Speech	HATE	Content that disparages particular groups or persons based on race, sex, sexuality or other characteristics.
Hosting, and Blogging Platforms	HOST	Web hosting services, blogging, and other online publishing platforms.
Human Rights Issues	HUMR	Sites dedicated to discussing human rights issues in various forms. Includes women's rights, and rights of minority ethnic groups.
LGBT	LGBT	A range of gay-lesbian-bisexual-transgender queer issues (excluding pornography).
Media sharing	MMED	Video, audio or photo sharing platforms.
News Media	NEWS	This category includes major news outlets (BBC, CNN, etc.) as well as regional news outlets, and independent media.
Online Dating	DATE	Online dating services which can be used to meet people, post profiles, chat, etc.
Pornography	PORN	Hard-core, and soft-core pornography.
Provocative Attire	PROV	Websites which show provocative attire, and portray women in a sexual manner, wearing minimal clothing.
Religion	REL	Sites devoted to discussion of religious issues, both supportive, and critical, as well as discussion of minority religious groups.
Sex Education	XED	Includes contraception, abstinence, STDs, healthy sexuality, teen pregnancy, rape prevention, abortion, sexual rights, and sexual health services.
Terrorism, and Militants	MILX	Sites promoting terrorism, violent militant or separatist movements.

Table 6.6: Categories of blocked websites illustrated in Figure 6.7 based on [166]

Internet regulation [98]. We collected, analyzed and summarized all reports issued by each EU member state's NRA from May 2020 to April 2021. Table 6.4 summarizes each country's reports and refers to any blocking of websites or services mentioned in the annual reports of NRAs.

6.5. Data analysis results

In our data analysis, we discovered several blocked websites in each country that were not listed in any public blocklist or mentioned in the annual Open Internet monitoring reports prepared by the NRA.

Our findings show a lack of transparency regarding network blocking in the EU countries. The data demonstrate that, although website blocking is a current activity, regulators have not provided enough evidence on how such blocking is being conducted by the telecommunication operators. This may result in over- or under-blocking websites, or network services being wrongfully blocked, as occurred in past incidents highlighted by some studies [269, 276].

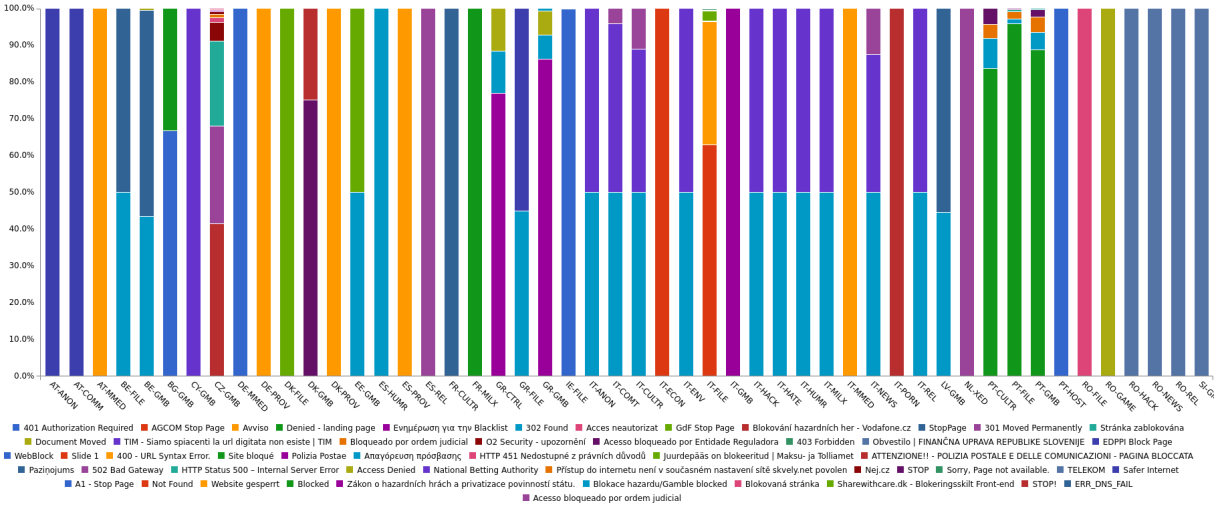

Figure 6.7: Blockpages per title and country code (Alpha-2 ISO3166) and blocked website category as depicted in Table 6.6 . The colored bars for each country code and website category illustrate the variety of blockpages found in OONI data.

6.5.1 Detected blockpages

We were able to identify 51 unique blockpages from 18 countries and 47 ASes that present a form of a blockpage or a generic error that is inconsistent with the network measurements of the control probe during the Web Connectivity test. Figure 6.5 illustrates all blockpages with the blockpage title per country code in the Alpha-2 ISO 3166 notation. Most countries present one or two blockpages while others present as many as seven. Such variation is due to network measurements performed non-uniformly by all countries, as we elaborate further in Section 6.6. Additionally, Figure 6.7 depicts all the categories of the blocked websites we detected in Figure 6.5, following the notation:

$$\{CountryCode - CategoryCode\}$$

The categories of the websites are extracted by Citizen Lab's URL test lists, the collaborative lists of websites or services curated and reviewed by community members to detect potentially blocked websites across countries [166]. The details of the category description and code of each detected blockpage are listed in Table 6.6.

6.5.2 Blocklists

A blocklist is a collection, put together by network regulators, of web addresses which may violate laws or regulations. For the sake of correct terminology and inclusive language, we use the word blocklist, instead of the word blacklist, used by almost all authorities in the EU countries that release such lists. During our research, we were able to find and identify official blocklists issued by 15 countries and one unofficial blocklist that is not issued by a country's authority, but is based on a court order. In total, we detected 23 blocklists with entries of websites from the categories regarding copyright, gambling, health, phishing, and tobacco. We developed a system that downloads, cleans, and assembles the blocklist files into Python Pandas data frames. This eases the data analysis and helps to get reproducible new versions of the blocklists in case of an update.

The majority of the blocklists are released in a PDF file format which makes them non-ideal for immediate processing. We converted these blocklists with the PyPDF2 and tabula-py Python libraries. We manually inspected the blocklist files and extracted relevant areas that include the blocked entries. Extracting

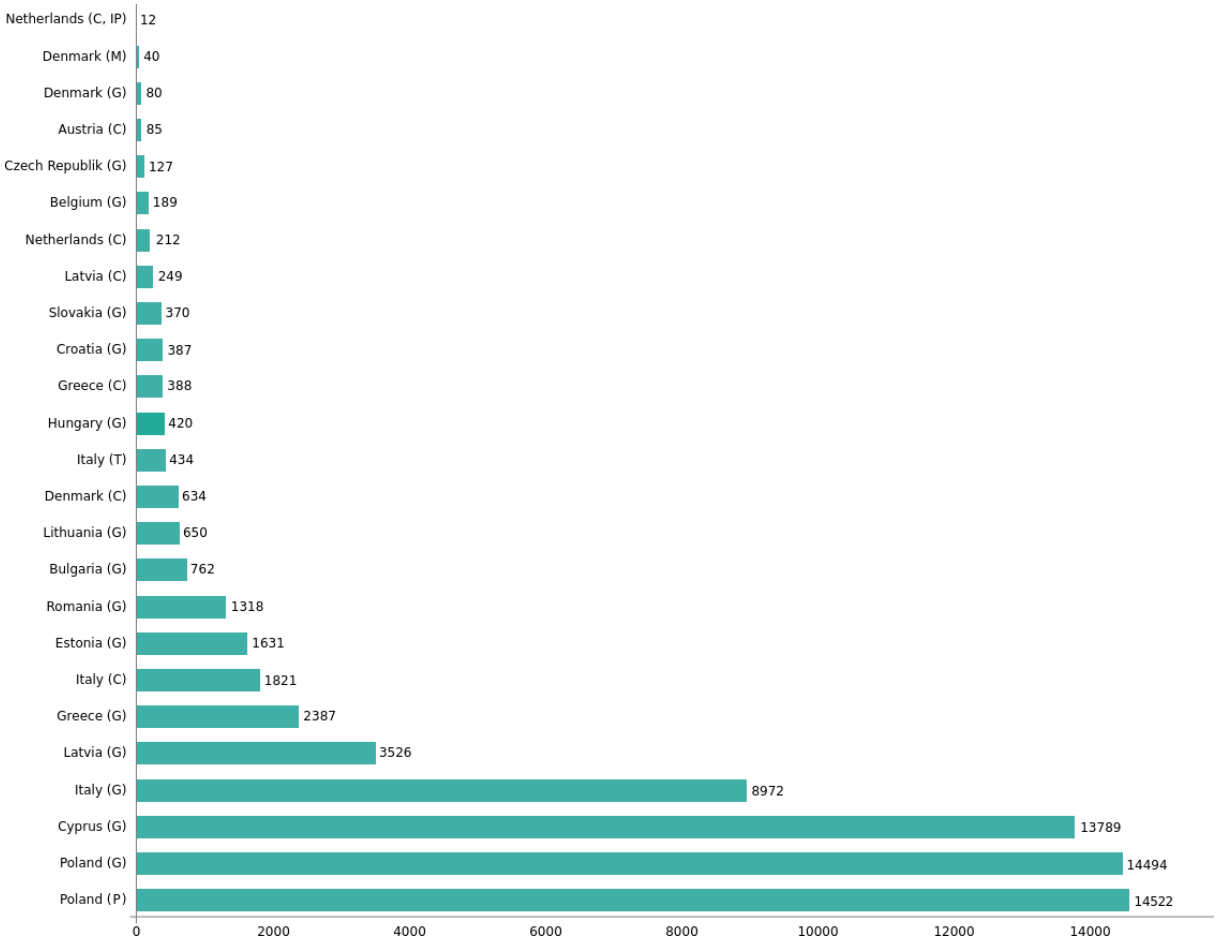

Figure 6.8: Number of entries per blocklist and blocklist type: (C) Copyright, (G) Gambling, (M) Health/Medical, (IP) IP address based, (P) Phishing, (T) Tobacco

data from PDF files in various languages and character encodings is a cumbersome process, but once the relevant areas of interest are isolated, we can then convert all the blocklists into data frames. The second most used file format, text file, is considerably easier to transform into data frames. Fewer blocklists are in HyperText Markup Language (HTML) and Comma-separated Values (CSV) file format; one is in Excel Workbook (XLSX) and another one in Extensible Markup Language (XML). All of these are simpler to extract into data frames.

We scanned them and thematically categorised the blocklists under the following categories: copyright, IP address based, health (including medical), gambling, phishing, or tobacco related websites. The results are presented in Table 6.9 and in Figure 6.8 and follow the notation:

$$\{CountryName(BlocklistType)\}$$

The data for the phishing blocklist of Poland has been omitted from Figure 6.10 because the additional 14,522 entries would make the Figure look odd. All EU countries publish a gambling blocklist, most publish a copyright blocklist, Italy publishes a blocklist for tobacco related websites, Denmark publishes a blocklist for medical related websites, and Poland a blocklist with website entries related to phishing attacks (not included in Figure 6.8). Finally, Netherlands releases an IP-based blocklist as an additional blocklist to their domain-name-based blocklist. The other 22 blocklists all contain only domain names (several include subdomains).

Because many countries publish more than one blocklist, we created Figure 6.10 to illustrate the cumu-

lative number of blocklist entries per country. The data show that Poland has almost 15 thousand entries (14,494) without including the phishing blocklist (with 14,522 entries), followed by Cyprus with 14 thousand entries (13,789) and Italy with more than 11 thousand entries (11,277). This is followed by Latvia with almost 4 thousand entries (3,775) and Greece with almost 3 thousand entries (2,775). Both Estonia (1,631) and Romania (1,318) have between two and one thousand blocklist entries. The remaining countries have significantly fewer than a thousand entries: Bulgaria (762), Denmark (754), Lithuania (650), Hungary (420), Croatia (387), Slovakia (370). The last four countries have blocklists with less than 220 entries, namely Netherlands (212), Belgium (189), Czech Republic (127), and Austria (85).

We sent an email query to all gambling authorities, as well as other agencies, for information on restricted websites for the countries for which we were unable to find any official or unofficial publicly available blocklist online. Apart from the ones published by the gambling regulators, countries typically have various blocklists. Due to ethical, legal, and humane considerations, we did not seek blocklists of websites that included or are related to Child Sexual Abuse Material (CSAM).

6.5.3 Blocklist authorities

In this section, we provide an alphabetical list of the national authorities that create blocklists and compel ISPs to block websites or services. Table 6.9 summarizes our results on the blocklist authorities, listing for each country the responsible entity that issues and publishes a blocklist of websites along with its type, the file format, and the relevant reference.

Austria Since 2016, the regulatory institution Telekom-Control-Kommission [253] has published on their website the national proceedings and decisions regarding net neutrality and the blocking of websites. The first relevant decision was published in 2018. It obliges ISPs to block access to websites due to alleged claims for injunctive relief under the copyright law [58]. There is no official blocklist and the blocked websites can be extracted from the PDF files of the decisions. Several ISPs provide a blocklist in their websites [120, 121, 173, 224], although it is unclear if the blocklists are thorough and up to date. The NRA report specifies cases of blocking based on copyright claims, typically implemented via DNS blocking (section 3.4 , part II) [40].

Belgium We detected two different blockpages in Belgium, one for gambling websites [256] directed by the Belgian Gaming Commission and another from the Belgian Entertainment Association for media content deemed illegal according to Belgian legislation. The error message for the blockpage links to a website in the source code of the blockpage that is dysfunctional (*https://onlinefairplay.info/*) and we were unable to obtain any information from the authority's website (*http://belgianentertainment.be/*) as it is also inoperative. The Belgian Institute for Postal Services and Telecommunications mentioned in their yearly report that there is no blocking of services or applications (section 5.117) [217]. The first blocklist entries date back to February 2012, as stated in the official website of the Belgian Gaming Commission [113].

Bulgaria The National Revenue Agency in Bulgaria is responsible for publishing and issuing the gambling blocklist [195]. According to the blocklist file the first released blocked entry took place in June 2013 [195]. In its annual report the communications regulation commission mentions that access to websites and content is blocked only in accordance with the national legislative acts (section 1.2) [1].

Country	Entity	Type	Format	Reference
Austria	Telekom Control Commission	Copyright (!)	PDF	[253]
Belgium	Gaming Commission	Gambling	HTML	[113]
Bulgaria	National Revenue Agency	Gambling	PDF	[195]
Croatia	Ministry of Finance and Tax Administration	Gambling	PDF	[259]
Cyprus	National Betting Authority	Gambling	TXT	[79]
Czech Republic	Ministry of Finance	Gambling	PDF	[292]
Denmark	Telecom Industry Association	Copyright	CSV	[252]
Denmark	Telecom Industry Association	Gambling	CSV	[252]
Denmark	Telecom Industry Association	Health	CSV	[252]
Estonia	Republic of Tax and Customs Board	Gambling	PDF	[37]
France	-	-	-	-
Germany	Clearinghouse Copyright on the Internet	Copyright (!)	PDF	[94]
Greece	Hellenic Copyright Association	Copyright	PDF	[128]
Greece	Hellenic Gaming Commission	Gambling	XLSX	[199]
Hungary	Supervisory Authority for Regulated Activities	Gambling	HTML	[38]
Italy	Autonomous Administration of the State Monopoly	Gambling	TXT	[9]
Italy	Autonomous Administration of the State Monopoly	Tobacco	TXT	[10]
Italy	Authority for Communications	Copyright (!)	PDF	[6]
Latvia	Lotteries and Gambling Supervisory Inspection	Gambling	TXT	[175]
Latvia	National Electronic Mass Media Council	Copyright	TXT	[190]
Lithuania	Gaming Control Authority	Gambling	TXT	[114]
Luxembourg	-	-	-	-
Malta	-	-	-	-
Netherlands	KPN ISP	Copyright (+,*)	HTML	[205]
Poland	CERT Polska	Phishing	Various	[172]
Poland	Ministry of Finance	Gambling	XML	[226]
Portugal	-	-	-	-
Romania	National Gambling Authority	Gambling	TXT	[200]
Slovakia	Gambling Regulatory Authority	Gambling	CSV, PDF	[112]
Slovenia	-	-	-	-
Spain	-	-	-	-
Sweden	-	-	-	-

Table 6.9: Detected blocklists per country (*: indicates unofficial, !: indicates assorted, +: includes IP addresses)

Croatia The Ministry of Finance Tax Administration is the responsible entity for the release and publication of the gambling blocklist. It is issued as a PDF file and contains the domain names with their subdomains along with the issue date of the blocking order for each entry in the blocklist. According to the blocklist, the first blocked entry appeared at the end of May 2019 [259]. There is no mention of Internet blocking in the country in the annual report issued by the Croatian Regulatory Authority for Network Industries [127].

Cyprus The National Betting Authority of the Republic of Cyprus is responsible for publishing and releasing the gambling blocklist in text file format [79]. It was established as an independent authority in 2012 and although the law was issued in 2012, the first public release of the blocklist was issued in February 2013 [272]. In the annual report published by the Office of the Commissioner of Electronic Communications and Postal Regulation in Cyprus there is no mention of any Internet blocking taking place [60].

Czech Republic The Ministry of Finance of the Czech Republic is responsible for issuing and publishing the blocklist of gambling websites in the country. The first blocked entry appeared in July 2017; 15 versions of the blocklist are already published, given the file name prefix (*v15*) [292]. There is no report of any blocking in the report of the Czech telecommunications authority [198].

Denmark The Telecom Industry Association of Denmark releases a number of blocklists based on Danish court orders. Three different blocklist categories exist: i. the game category contains gambling websites; ii. the health category with medical and health-related websites; and iii. the intellectual property rights category with websites related to copyright infringement. All blocklists are published in the CSV file format, and a PDF file provides the date of each entry added to the blocklist. The Danish Energy Agency sent out a questionnaire to 40 ISPs in Denmark on the grounds of the EU net neutrality regulation. 30% of the ISPs stated that they are partly blocking access to the Internet. Specifically, the ISPs mentioned blocking access to CSAM websites with extremist content, or calls for terror. Further, the ISPs mentioned blocking traffic to malicious servers related to COVID-19 crime (section 4.1) [8].

Estonia The Republic of Estonia's Tax and Customs Board is responsible for issuing and publishing the blocklist of gambling websites in the country. It is distributed as a PDF file and is publicly available to download [37]. The annual report of the Estonian consumer protection and technical regulatory authority fails to mention any Internet blocking in the country [219].

France In France, the National Commission on Informatics and Liberty publishes yearly reports on the administrative blocking of websites. The reports give an overview of the blocked websites related to terrorism and CSAM [63]. They have appointed a person to verify the validity of requests for removal of content and blocking made by the central office for combating information and communication technology crime. However they do not provide details as to which websites have been blocked, but only statistical information on the number of requests to block websites. The latest report covers the period from February to December 2019, and mentions that 18,177 blocking orders were made. Of these, 420 requests were related to blocked websites, 11,874 for content removal, and 5,883 for dereferencing of email addresses [110]. Moreover, there is no mention of Internet blocking in France's Electronic Communications, Postal and Print Media Distribution's NRA annual report [24].

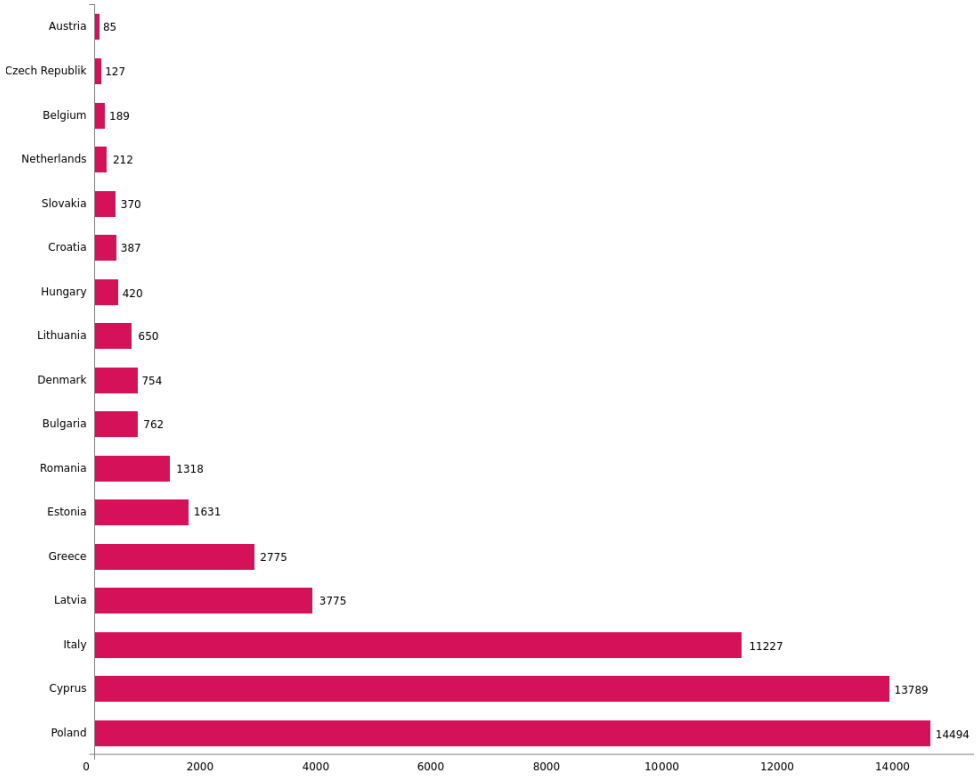

Figure 6.10: Total number of blocklist entries (cumulative) per country

Germany Clearinghouse Copyright on the Internet is an independent body established by ISPs and rights holders. Its purpose is to review and propose the blocking of websites according to certain criteria and requirements related to copyright infringement. As mentioned on its website, a review board, at the request of the copyright holder, reviews the copyright allegedly infringing website and, if the requirements are met, recommends DNS blocking. They publish the recommendations for blocking domains on their website and the first entry appeared in February 2021 [94]. According to the Federal Network Agency's annual report on net neutrality, there is no national law in Germany requiring ISPs to implement blocking in their networks. An unnamed (in the report) ISP was required to block access to some (unspecified) websites by means of DNS blocking due to a court ruling (section 3.1.2) [42].

Greece The annual report of the NRA in Greece [89] mentions that ISPs in the country block websites based on two public blocklists according to the laws related to the protection of intellectual property and blocking of gambling websites [128, 199]. Additionally, the report mentions that ISPs block domain names to protect against phishing attacks and block IP addresses to protect their internal network and defend against distributed denial of service attacks. A user who visits one of the websites listed in the blocklist gets redirected (with an HTTP 301 redirect) to the websites of the blocking authorities. The servers of the authorities may potentially collect IP addresses and further information of users trying to access the blocked websites. Previous research observed that ISPs implemented their own blocking pages without redirecting the users to the website of the gambling regulation authority when they try to access a website on the blocklist [276].

Hungary The Supervisory Authority for Regulated Activities in Hungary is responsible for issuing and releasing a public blocklist for gambling websites [38]. There is no mention of the blocklist in the annual report published by the national media and communications authority [183].

Italy In Italy, we discovered three publicly available blocklists issued by two different entities. The Autonomous Administration of the State Monopoly lists websites related to gambling [9] and tobacco products [10]. The Authority for Communications responsible for the blocklist of copyright infringement cases [6]. There is no mention of any blocking in the NRA annual report [7].

Latvia The Lotteries and Gambling Supervisory Inspection is the responsible authority for issuing and publicly releasing a gambling blocklist in Latvia. The first blocked entries appeared in August 2014 [174]. We discovered another blocklist published by the National Electronic Mass Media Council of Latvia with entries related to copyright infringement. Both blocklists are released in a text file format [190]. The annual report published by the Public Utilities Commission didn't report any blocking [59].

Lithuania The Gaming Control Authority under the Ministry of Finance is the responsible authority for issuing the gambling blocklist of websites in Lithuania. The first entries were published in January 2016 [114] in a text file format. However there is no reference to blocking in the annual report [227] published by the Communications Regulatory Authority of the Republic of Lithuania.

Malta An e-mail communication from the Malta Gaming Authority [179] revealed that they do not have the authority to block websites. However, they cooperate with Malta's police force to stop criminal gambling activity. Investigations and prosecutions are then carried out by the police, assisted by the Authority as necessary. Therefore, any repercussions (including website blocking) of illegal activities or services fall under the jurisdiction of the Malta police. The annual report of the Malta Communication Authority mentions an ongoing investigation to block specific IP addresses (undefined in the report), without saying that there was any website blocking [31].

Netherlands A blockpage [205] in OONI network measurements probed on the KPN ISP mentions that the judge for provisional legal protection in Midden-Nederland ruled in January 2018 that The Pirate Bay's website should be blocked on all KPN networks including Telfort, Simyo, and KPN Hotspots. The decision lists several IP addresses, domains and subdomains that ISPs must block. The same blockpage mentions that the judicial decision [85] was also sent to other ISPs. The unofficial blocklist extracted from the blockpage [205] lists 12 IP addresses (IPv4 and IPv6) and a list of 212 domains and subdomains that are proxies or mirrors of The Pirate Bay website. The Authority for Consumers and Markets released the annual NRA report without providing any information about the blocking of websites or services [62].

Poland The NRA report of the Office of Electronic Communications [93] mentions that ISPs are obliged to block gambling websites. The Polish Ministry of Finance releases the gambling blocklist [226], provided as a REST XML service that can be retrieved programmatically and includes the documentation of its specification. Another blocklist (called Warning List) [172] has been created to block websites related to phishing activities. An agreement was made in March 2020 with the Minister of Digital Affairs, the Office of Electronic Communications and National Research Institute, and the four largest mobile network operators, Orange, T-Mobile, P4, and Polkomtel, to block specific websites [262]. The CERT Polska team is responsible for the maintenance and release of this blocklist. On their website, they have created a form where individuals can report suspicious websites, and each report is manually verified by at least two persons. The blocklist is released in various file formats, updated every 5 minutes, and the full specification of the API is available on their website [172].

Romania The Romanian National Gambling Authority has released a gambling blocklist since 2015 [200]. It is available on their website in a text file format. They also provide a helper script (written in the PHP programming language) that replaces the A and NS DNS records of the domain (and the *www* subdomain) for all the entries found on the blocklist, compatible with the BIND DNS server configuration. According to the annual report [15] of the National Authority for Administration and Regulation in Communications, that entity issued 15 blocking orders related to COVID-19 fake news, as well as protection and prevention measures during the state of emergency that ended in May 2020 [14]. The gambling blocklist is not mentioned in the report.

Slovakia The annual report [92] published by the Regulatory Authority for Electronic Communications and Postal Services mentions that ISPs block access to applications or services in the event of illegal content as ruled by European or national legislation. Online gambling websites without a Slovak license are blocked, as well as websites that host CSAM. The Slovakian Gambling Regulatory Authority is responsible for issuing and publishing the gambling blocklist, and its first entry appeared in August 2019 [112].

6.6. Conclusion and further discussion

This study sheds light on how website blocking occurs in the European Union. The process of gathering data involved several steps and sources, sometimes not easily available. Some of the data sources were provided after e-mail communication with the regulators. The research identified blockpages and blocklists in jurisdictions across Europe. In our blocklist evaluation study (in Section 6.5.2) we detected different types of blocklist publication and distribution methods.

We identified some issues with the reporting of such blocklists. Most regulators and authorities are using PDF files, others publish the blocklists on their websites, and a few release them in a CSV or other file format. All of these approaches are cumbersome and lead to error-prone processes for the ISPs maintaining updated lists of websites and services to block. This may result in over- and under-blocking [276]. Besides, most NRAs do not describe in their reports what blocking they do. Only a few authorities publish even limited details on the resources and websites blocked, with no references to the blocklists. Details for each country are provided in Section 6.5. A well-designed system can help address a number of these problems related to Internet regulations and blocking of websites and services, albeit the issue is not just technological, but may involve political and legal questions.

We focused on the overlooked trend of EU member states deploying surveillance and network infrastructures to adhere to the EU legislation. We focused on the publication and release of EU blocklists and website blocking in 2020. Based on historical network measurements data by OONI, this paper provides evidence that countries in the EU not only use blocklists as a means of blocking access to websites but also block different types and categories of websites and services that are not included in the publicly available (identified) blocklists.

All countries publish a gambling blocklist, most publish a copyright blocklist, Italy publishes a blocklist for tobacco-related websites, Denmark publishes a blocklist for medical websites, and Poland publishes a blocklist with entries on phishing websites. Finally, Netherlands publishes IP-based blocklist as an additional blocklist to their domain-name-based blocklist. The other 22 blocklists all list only domain names or URLs.

In terms of the cumulative number of blocklist entries per country, Poland has just over 29,000 entries (including the phishing blocklist), followed by Cyprus with almost 14,000 entries and Italy with more

than 11,000 entries. Latvia, Greece, Estonia, and Romania, with between 4,000 entries and 1,000 entries in block lists, make up the midfield. The remaining countries have significantly fewer than a thousand entries.

We also analyzed data from the OONI project, a platform for detecting Internet censorship that has been actively developed since 2012. OONI network measurements are carried out on an ad-hoc basis by volunteers. The data submitted still rely on the availability and willingness of people to conduct network measurements, notwithstanding the software's ongoing improvement. Although OONI has collected and released data on network measurements from all countries worldwide, getting longitudinal network measurements is challenging. It is important that quantitative network measurements be carried out from diverse locations even for the same ISPs and ASes.

6.6.1 Regulatory sanctions and restriction to access to online resources

As the literature demonstrates, governments and state actors have used Internet censorship to influence political discourse and favor businesses under their own control [124]. Citizens can also potentially be denied access to services as a result of local regulatory laws, financial reasons, or because their country has fallen under sanctions and prohibits foreign companies from operating within their jurisdiction [273]. Some authors suggest that, having been characteristic of repressive regimes, Internet censorship could become almost ubiquitous in both democratic and authoritarian states [32].

As example, the EU has imposed a number of sanctions in response to Russia's invasion of Ukraine. In particular, the EU Council adopted Decision 2022/351, imposing new restrictive measures against the Russian state media and their subsidiaries [96]. The Council decision does not specify exact websites, domains, or URLs to be blocked, but rather says that: "It shall be prohibited for operators to broadcast or to enable, facilitate or otherwise contribute to broadcast, any content by the legal persons, entities or bodies listed in Annex XV, including through transmission or distribution by any means such as cable, satellite, IP-TV, Internet service providers, Internet video-sharing platforms or applications, whether new or pre-installed" [99]. Annex XV lists only the names of the entities or bodies, specifically: *RT- Russia Today English*, *RT- Russia Today UK*, *RT- Russia Today Germany RT- Russia Today France*, *RT- Russia Today Spanish* and *Sputnik*.

This prohibition forces ISPs to make their own decisions about which websites and services to block, which is a difficult process with many implementation gaps and the risk of under- or over-blocking [106]. The EU council also calls for blocking content distributed via cable, satellite, ISPs, and IP-TV connections on video-sharing websites in addition to the websites owned by these entities. In reality, such extensive service blocking is impractical and results in the excessive over-blocking of websites and services [106, 125, 276].

ISPs in the EU are already employing various blocking techniques to block the websites of *rt.com* and *sputniknews.com*. The majority of them make use of their current blocking infrastructure, including the same blocking pages that falsely claim the websites are blocked because of copyright infringement, gambling, or other laws [257]. These are similar to the blocking pages examined in Section 6.5.1 which have nothing to do with the blocking of these websites. The blocking infrastructure requires a significant number of labor hours and hardware infrastructure to be implemented and maintained [257]. For instance, this is the situation with smaller ISPs in the UK, where new Internet service sanctions in the country require ISPs to block access to the websites and services listed in the sanctions. Failure to do so can result in fines of up to £1,000,000 [197, 257].

6.6.2 Limitations

The conclusions of this paper have limitations which may prompt future research, especially regarding other forms of Internet censorship and further methods of network interference that may require a legal and policy analysis from the principles of network neutrality and Open Internet.

– 7 –
Conclusions

7.1. Summary

This thesis provides an overview of the Internet censorship implemented by democratic countries in the EU. During the period of the research a significant amount of network measurements has been collected with the help of OONI software and Lepidopter distribution as well as other custom developed solutions to adapt in the non homogeneity of the probed networks. Our research helps to understand how countries issue blocklists, the content of them and in some cases the problematic of using them to block websites. With respect to the implementation analyzed network measurements revealed that ISPs were often over-blocking and in few cases under-blocking websites raising awareness to the transparency of the whole blocking process.

Through out our research we described how the blocking of websites and network services in the EU has been intensified and often ISPs are blocking websites for reasons that we are not aware of. Finally in this thesis we provided all our analysis and source code available in a free license as it could be used for future work and used by other researchers, activists, human right advocates and policy regulators not only in the EU but globally. As such has been case with most of the released source code. Additionally in the App store censorship we proposed some techniques to detect the blocking or available of applications across countries for Google Play, Apple and Tencent app stores.

As Internet censorship in the EU has been increasing we hope that our results and work will be used to push back further blocking of websites and services. Fight for an open and free Internet!

7.2. Future work

Internet censorship is an arms race and although newer techniques help to circumvent the blocking and raise awareness, we are still missing historical references to network blocking incidents. To this extent a very relevant future work is the development of a glossary to collect, evaluate and catalog network disruption incident information. These incidents can vary from censorship of websites and services, network infrastructure disasters from natural causes or abrupt disruption of a network due to human error or misconfiguration.

The information from the potential incidents will be evaluated by the community, Internet censorship researchers, network engineers, policy regulators and users in the affected networks. Once evaluated and identified the network disruption incidents will be listed publicly on an open database and relevant website. Such incidents reports will include information on the nature of the blocking, accompanying network measurements, technical reports, official press releases or network maintenance documents and other information that could be used potentially to verify the cause of the disruption or censorship.

Network disruptions via means of natural disasters are not the main case cause for inaccessibility of

websites or services. However it is still useful to catalog and archive such incidences. Often governments are using natural disasters as an excuse to hide censorship that occurred due to political instability.

The network disruption glossary will be used as a baseline of communication between interested parties that are not aware of the cases where networks, websites or services available on the Internet are being intentionally blocked. In fact censors try to block them as if it was an error on the equipment of a user or ISP infrastructure. The proposed future work will potentially help to raise awareness on the matters of Internet censorship and especially in cases were complete network blackouts occur in a country.

– A –

Appendix

A.1. Published work

This section includes all of the related work published during the course of my Ph.D. It contains peer-reviewed and journal articles.

P.1 Vasilis Ververis, George Kargiotakis, Arturo Filastò, Benjamin Fabian, and Afentoulis Alexandros. *"Understanding Internet Censorship Policy: The Case of Greece"*. In: *Free and Open Communications on the Internet*. Usenix, 2015. https://www.usenix.org/system/files/conference/foci15/foci15-paper-ververis-updated-2.pdf

P.2 Vasilis Ververis, Marios Isaakidis, Chrystalleni Loizidou, and Benjamin Fabian. *"Internet Censorship Capabilities in Cyprus: An Investigation of Online Gambling Blocklisting"*. In: *E-Democracy*. Springer, 2017. https://censorbib.nymity.ch/pdf/Ververis2017a.pdf

P.3 Vasilis Ververis, Tatiana Ermakova, Marios Isaakidis, Simone Basso, Benjamin Fabian, and Stefania Milan. *"Understanding Internet Censorship in Europe: The Case of Spain"*. In: *13th ACM Web Science Conference 2021*. WebSci '21. Virtual Event, United Kingdom: Association for Computing Machinery, 2021, pp. 319–328. ISBN: 9781450383301. DOI: 10.1145/3447535.3462638. https://doi.org/10.1145/3447535.3462638

P.4 Vasilis Ververis, Marios Isaakidis, Valentin Weber, and Benjamin Fabian. *"Shedding Light on Mobile App Store Censorship"*. In: *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*. UMAP'19 Adjunct. Larnaca, Cyprus: Association for Computing Machinery, 2019, pp. 193–198. ISBN: 9781450367110. DOI: 10.1145/3314183.3324965. https://doi.org/10.1145/3314183.3324965

P.5 Vasilis Ververis, Tatiana Ermakova, Lucas Lasota, and Benjamin Fabian. *"Website Blocking in the European Union: Network Interference from the Perspective of Open Internet"*. In: *Policy and Internet* (forthcoming 2023). DOI: 10.1002/poi3.367

P.6 Ververis Vasilis, Marguel Sophia, and Fabian Benjamin. *"Cross-Country Comparison of Internet Censorship: A Literature Review"*. In: *Policy & Internet* 12.4 (2020), pp. 450–473. https://onlinelibrary.wiley.com/doi/full/10.1002/poi3.228

P.7 Vasilis Ververis, Olga Khrustaleva, and Eliana Quiroz. *"Network interference in Latin America: Evaluating network measurements to detect information controls and Internet censorship"*. In: (Nov. 2017). ISSN:

2175-9596. https://lavits.org/wp-content/uploads/2018/04/32-Vasilis-Ververis-Olga-Khrustaleva-e-Eliana-Quiroz.pdf

P.8 Abbas Razaghpanah, Anke Li, Arturo Filastò, Rishab Nithyanand, Vasilis Ververis, Will Scott, and Phillipa Gill. *"Exploring the Design Space of Longitudinal Censorship Measurement Platforms"*. In: *CoRR* abs/1606.01979 (2016). arXiv: 1606.01979. http://arxiv.org/abs/1606.01979

P.9 Benjamin Fabian, Annika Baumann, Mathias Ehlert, Vasilis Ververis, and Tatiana Ermakova. *"CORIA —Analyzing internet connectivity risks using network graphs"*. In: *2017 IEEE International Conference on Communications (ICC)*. Ieee. 2017, pp. 1–6

P.10 Hee-Eun Lee, Tatiana Ermakova, Vasilis Ververis, and Benjamin Fabian. *"Detecting child sexual abuse material: A comprehensive survey"*. In: *Forensic Science International: Digital Investigation* 34 (2020), p. 301022

– A –

Bibliography

[1] Communications Regulation Commission (CRC). *Annual report on the implementation of the regulation (EC) 2015/2120 for 2020*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78849 (cit. on pp. 75, 79).

[2] Nicholas Aase, R. Crandall Jedidiah, Álvaro Díaz, Jeffrey Knockel, Jorge Oca Molinero, Jared Saia, Dan Wallach, and Tao Zhu. *"Whiskey, Weed, and Wukan on the World Wide Web: On Measuring Censors' Resources and Motivations"*. In: (2012). https://www.usenix.org/system/files/conference/foci12/foci12-final17.pdf (cit. on p. 18).

[3] Giuseppe Aceto, Alessio Botta, Antonio Pescapè, M. Faheem Awan, Tahir Ahmad, and Saad Qaisar. *"Analyzing Internet Censorship in Pakistan"*. In: *Research and Technologies for Society and Industry*. Ieee, 2016. http://wpage.unina.it/giuseppe.aceto/pub/aceto2016analyzing.pdf (cit. on pp. 45, 69, 71).

[4] Giuseppe Aceto and Antonio Pescapè. *"Internet Censorship detection: A survey"*. In: *Computer Networks* 83 (2015), pp. 381–421. https://censorbib.nymity.ch/pdf/Aceto2015b.pdf (cit. on p. 46).

[5] Acn. *Spain passes decree to shut down websites and social media over public order threats. Online tools have been key in the organization of the Catalan independence movement*. 2019. https://www.catalannews.com/politics/item/spain-passes-decree-to-shut-down-websites-and-social-media-over-public-order-threats (cit. on p. 45).

[6] *Agcom*. June 2021. https://www.agcom.it (cit. on pp. 80, 83).

[7] Agcom. *Report on the activities carried out by the Authority in the field of Open Internet*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78867 (cit. on pp. 75, 83).

[8] Danish Energy Agency. *The Danish Energy Agency's supervision of the EU Regulation on access to the open internet*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78855 (cit. on pp. 75, 81).

[9] *Agenzia delle dogane e dei Monopoli*. June 2021. https://www1.agenziadoganemonopoli.gov.it/files%5C%5Fsiti%5C%5Finibiti%5C%5Ftabacchi/elenco%5C%5Fsiti%5C%5Finibiti.txt (cit. on pp. 80, 83).

[10] *Agenzia delle dogane e dei Monopoli*. June 2021. https://www1.agenziadoganemonopoli.gov.it/files_siti_inibiti_tabacchi/elenco_siti_inibiti.txt (cit. on pp. 80, 83).

[11] Urs-Vito Albrecht, Uta Hillebrand, and Ute von Jan. *"Relevance of trust marks and CE labels in German-language store descriptions of health apps: analysis"*. In: *JMIR mHealth and uHealth* 6.4 (2018) (cit. on p. 62).

[12] Anacom. *Report on net neutrality - May 2020 to April 2021*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78917 (cit. on p. 75).

[13] The Anatomy. *of Web Censorship in Pakistan*. Zubair Nabi In: Free and Open Communications on the Internet (Washington, DC, USA), USENIX (cit. on p. 18).

[14] Ancom. *ANCOM Decisions for the implementation of Decree no. 195*. https://ec.europa.eu/newsroom/dae/redirection/document/78877 (cit. on pp. 75, 84).

[15] Ancom. *Monitoring compliance with Regulation (EU) No 2015/2120 on open Internet access 01 May 2020 - 30 April 2021*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78877 (cit. on pp. 75, 84).

[16] Collin Anderson. *Dimming the Internet: Detecting Throttling as a Mechanism of Censorship in Iran*. Tech. rep. University of Pennsylvania, 2013. https://arxiv.org/pdf/1306.4361v1.pdf (cit. on pp. 45, 71).

[17] Collin Anderson. *The Hidden Internet of Iran: Private Address Allocations on a National Network*. Tech. rep. 2012. https://arxiv.org/pdf/1209.6398v1.pdf (cit. on p. 71).

[18] Andy. *Block Pirate Bay in 72 Hours, Spanish Court Tells ISPs*. 2015. https://torrentfreak.com/block-pirate-bay-in-72-hours-spanish-court-tells-isps-150327/ (cit. on p. 45).

[19] Christina Angelopoulos. *"Filtering the Internet for Copyrighted Content in Europe"*. In: *IRIS plus 2009-4, European Audiovisual Observatory* 2009.4 (2009) (cit. on pp. 45, 69).

[20] *App Store, iTunes Store, and Apple Books availability*. Mar. 2019. https://support.apple.com/en-us/HT204411 (cit. on p. 65).

[21] *AppInChina Chinese App Store Rankings - AppInChina | The Market*. Mar. 2019. https://www.appinchina.co/market/app-stores (cit. on p. 61).

[22] *Apple drops hundreds of VPN apps at Beijing' s request*. Nov. 2017. https://www.ft.com/content/ad42e536-cf36-11e7-b781-794ce08b24dc (cit. on p. 67).

[23] *Apple removes VPN Apps from China App Store | ExpressVPN*. July 2017. https://www.expressvpn.com/blog/china-ios-app-store-removes-vpns (cit. on p. 67).

[24] Arcep. *The state of the internet in France*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78869 (cit. on pp. 75, 81).

[25] Internet Archive. *The Wayback Machine*. http://web.archive.org/ (visited on 06/05/2015) (cit. on p. 37).

[26] Web Archive. *eln-voces.com - This Domain Has Expired*. July 2019. https://web.archive.org/web/20190716230422/http://www.eln-voces.com/ (visited on 07/16/2019) (cit. on pp. 54, 58).

[27] Archive.is. *Webpage Capture*. http://archive.is (visited on 06/04/2015) (cit. on p. 37).

[28] Simurgh Aryan, Homa Aryan, and J. Alex Halderman. *"Internet Censorship in Iran: A First Look"*. In: *Free and Open Communications on the Internet*. Usenix, 2013. https://censorbib.nymity.ch/pdf/Aryan2013a.pdf (cit. on pp. 45, 71).

[29] D. Atkins and R. Austein. *Threat analysis of the domain name system (DNS)*. Aug. 2004. http://web.archive.org/web/20140826081656/http://www.ietf.org/rfc/rfc3833.txt (cit. on pp. 22, 32, 39).

[30] RIPE Atlas. *Ripe*. 2020. https://atlas.ripe.net (cit. on p. 46).

[31] Malta Communications Authority. *Report of the Malta Communications Authority on its monitoring and findings in accordance with Article 5 of Regulation (EU) 2015/2120 concerning the European Net Neutrality Rules*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78866 (cit. on pp. 75, 83).

[32] Derek E. Bambauer. *"Censorship v3.1."* In: *IEEE Internet Computing* (2013), pp. 26–33. https://ieeexplore.ieee.org/document/6415890 (cit. on p. 85).

[33] Alex Barrera. *¿Que? A judge has ordered the blockage of Uber' s website in Spain*. 2014. https://tech.eu/news/court-shutdown-uber-spain/ (cit. on p. 45).

[34] Simone Basso. *Aladdin: Experimental Web Connectivity implementation*. 2021. https://github.com/bassosimone/aladdin (cit. on p. 58).

[35] Simone Basso. *"Evaluating OONI's New Measurement Engine"*. In: (May 2020). https://ooni.org/post/2020-engine-evaluation-spain/ (cit. on p. 56).

[36] Tim Berners-Lee. *Universal Resource Identifiers in WWW*. http://web.archive.org/web/20140829223110/http://www.w3.org/Addressing/URL/uri-spec.html (cit. on p. 31).

[37] *Blocked illegal remote gambling sites: Estonian Tax and Customs Board*. Jan. 2021. https://www.emta.ee/eng/private-client/land-vehicle-forest-gambling/remote-gambling-sites (cit. on pp. 80, 81).

[38] *Blokkolt honlapok – Szerencsejtk Felgyelet*. June 2021. https://szf.gov.hu/hatosag/blokkolt-honlapok (cit. on pp. 80, 82).

[39] Yana Breindl and Joss Wright. *"Internet Filtering in Liberal Democracies"*. In: (2012). https://www.usenix.org/conference/foci12/workshop-program/presentation/Breindl (cit. on p. 18).

[40] Austrian Regulatory Authority for Broadcasting and Telecommunications. *RTR Net Neutrality Report 2021*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78845 (cit. on pp. 75, 79).

[41] Matthias Brugger. *Internet censorship in the Catalan referendum*. Dec. 2017. https://mirror.netcologne.de/CCC/congress/2017/slides-pdf/34c3-9028-internet%5C%5Fcensorship%5C%5Fin%5C%5Fthe%5C%5Fcatalan%5C%5Freferendum.pdf (cit. on p. 45).

[42] Bundesnetzagentur. *Net Neutrality in Germany Annual Report 2019/2020*. Apr. 2020. https://ec.europa.eu/newsroom/dae/document.cfm?doc%5C%5Fid=68751 (cit. on p. 82).

[43] Bundesnetzagentur. *Net Neutrality in Germany Annual Report 2020/2021*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78865 (cit. on p. 75).

[44] Andreas Busch, Patrick Theiner, and Yana Breindl. *"Internet Censorship in Liberal Democracies: Learning from Autocracies?"* en. In: *Managing Democracy in the Digital Age*. Ed. by Julia Schwanholz, Todd Graham, and Peter-Tobias Stoll. Cham: Springer International Publishing, 2018, pp. 11–28. DOI: 10.1007/978-3-319-61708-4_2. http://link.springer.com/10.1007/978-3-319-61708-4%5C%5F2 (visited on 04/05/2020) (cit. on pp. 45, 69, 71).

[45] Christophorou C. *Cyprus: Media Pluralism Monitor 2015 [European University Institute, Robert Schuman Centre for Advanced Studies]*. https://web.archive.org/web/20170606205318/http://monitor.cmpf.eui.eu/mpm2015/results/cyprus/ (visited on 06/06/2017) (cit. on p. 36).

[46] Cablenet. *Cablenet ISP official website*. http://archive.is/UBxqc (visited on 06/05/2017) (cit. on p. 40).

[47] Callsat. *Callsat ISP official website*. http://archive.is/CAuFL (visited on 06/04/2015) (cit. on p. 40).

[48] Ashley Carman. *"Apple will start reporting government requests to remove apps from the App Store"*. In: *Verge* (May 2018). https://www.theverge.com/2018/5/25/17396512/apple-transparency-report-app-takedown-requests (cit. on p. 68).

[49] Abdelberi Chaabane, Terence Chen, Mathieu Cunche, Emiliano De Cristofaro, Arik Friedman, and Mohamed Ali Kaafar. *"Censorship in the Wild: Analyzing Internet Filtering in Syria"*. In: *Internet Measurement Conference*. Acm, 2014. http://conferences2.sigcomm.org/imc/2014/papers/p285.pdf (cit. on pp. 45, 71).

[50] Le Chen, Chi Zhang, and Christo Wilson. *"Tweeting Under Pressure: Analyzing Trending Topics and Evolving Word Choice on Sina Weibo"*. In: *Conference on Online Social Networks*. Acm, 2013. https://cbw.sh/static/pdf/weibo-cosn13.pdf (cit. on p. 71).

[51] Ning Chen, Jialiu Lin, Steven C. H. Hoi, Xiaokui Xiao, and Boshen Zhang. *"AR-miner: Mining Informative Reviews for Developers from Mobile App Marketplace"*. In: *Proceedings of the 36th International Conference on Software Engineering*. Icse 2014. Hyderabad, India: Acm, 2014, pp. 767–778. ISBN: 978-1-4503-2756-5. DOI: 10.1145/2568225.2568263. http://doi.acm.org/10.1145/2568225.2568263 (cit. on p. 63).

[52] *China will block all non-approved VPNs from next month*. Apr. 2019. https://www.techradar.com/news/china-will-block-all-non-approved-vpns-from-next-month (cit. on p. 67).

[53] *China's Cybersecurity Law: An Intro for Foreign Businesses*. Mar. 2018. https://www.china-briefing.com/news/chinas-cybersecurity-law-an-introduction-for-foreign-businesspeople (cit. on p. 67).

[54] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson. *"Ignoring the Great Firewall of China"*. In: *Privacy Enhancing Technologies*. Springer, 2006, pp. 20–35. https://www.cl.cam.ac.uk/~rnc1/ignoring.pdf (cit. on p. 71).

[55] Greek Gaming Commision. *"EEEP Blocklist First Version"*. June 2013. http://www.gamingcommission.gov.gr/images/apofaseis/lists/black%20list0001.pdf (cit. on p. 31).

[56] Greek Gaming Commision. *"EEEP Blocklist Second Version"*. Nov. 2013. http://www.gamingcommission.gov.gr/images/Anakoinoseis/BlackList%5C%5FEEEP%5C%5F%2022112013.pdf (cit. on p. 31).

[57] Greek Gaming Commision. *"EEEP Blocklist Third Version"*. Feb. 2014. http://www.gamingcommission.gov.gr/images/Anakoinoseis/BlackList%5C%5FEEEP%5C%5F%2021022014.pdf (cit. on p. 31).

[58] Telekom Control Commission. *Decisions by the regulator on net neutrality*. Jan. 2023. https://www.rtr.at/TKP/was_wir_tun/telekommunikation/weitere-regulierungsthemen/netzneutralitaet/nn_procedures.en.html (cit. on p. 79).

[59] The Public Utilities Commission. *Report on Compliance with the Regulation of Open Internet Access*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78862 (cit. on pp. 75, 83).

[60] Office of the Commissioner of Electronic Communications and Postal Regulation. *Annual Report 2021 on Open Internet*. Aug. 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78851 (cit. on pp. 75, 81).

[61] Commission for Communications Regulation. *Implementation of EU Open Internet Access Regulations in Ireland*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78859 (cit. on p. 75).

[62] Authority for Consumers & Markets. *2020-2021 Annual Report on Net Neutrality*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78871 (cit. on pp. 75, 83).

[63] *Contrôle du blocage administratif des sites : la personnalité qualifiée présente son 5ème rapport d' activité*. June 2021. https://www.cnil.fr/fr/controle-du-blocage-administratif-des-sites-la-personnalite-qualifiee-presente-son-5eme-rapport (cit. on p. 81).

[64] Aug. 2017. https://eur-lex.europa.eu/legal-content/en/TXT/PDF/?uri=CELEX%3A32017D1426&from=EN (cit. on p. 54).

[65] *"Court Ordered ISPs to Block IP Addresses"*. May 2012. http://web.archive.org/web/20141114193131/https://www.void.gr/kargig/blog/wp-content/4658%5C%5F2012.pdf (cit. on p. 17).

[66] National Betting Authority of Cyprus. *Betting Law 2012*. http://web.archive.org/web/20170605132235/http://nba.gov.cy/wp-content/uploads/TheBettingLawof2012.pdf (visited on 06/05/2015) (cit. on pp. 36, 37).

[67] National Betting Authority of Cyprus. *Blocklist*. Feb. 14, 2013. https://web.archive.org/web/20130217021102/http://blocking.nba.com.cy:80/ (visited on 02/17/2013) (cit. on p. 37).

[68] National Betting Authority of Cyprus. *Blocklist*. Apr. 19, 2013. https://web.archive.org/web/20130906231633/http://blocking.nba.com.cy:80/ (visited on 09/06/2013) (cit. on p. 37).

[69] National Betting Authority of Cyprus. *Blocklist*. Nov. 12, 2013. https://web.archive.org/web/20131124123355/http://blocking.nba.com.cy:80/ (visited on 11/24/2013) (cit. on p. 37).

[70] National Betting Authority of Cyprus. *Blocklist*. Jan. 29, 2016. https://web.archive.org/web/20160201084135/http://blocking.nba.com.cy:80/ (visited on 02/20/2016) (cit. on p. 37).

[71] National Betting Authority of Cyprus. *Blocklist*. Feb. 15, 2016. https://web.archive.org/web/20160303044805/http://blocking.nba.com.cy:80/ (visited on 03/03/2016) (cit. on p. 37).

[72] National Betting Authority of Cyprus. *Blocklist*. Nov. 4, 2016. https://web.archive.org/web/20161106114742/http://blocking.nba.com.cy:80/ (visited on 11/06/2016) (cit. on p. 37).

[73] National Betting Authority of Cyprus. *Blocklist*. Feb. 17, 2016. https://archive.fo/Wdb9n (visited on 02/24/2017) (cit. on p. 37).

[74] National Betting Authority of Cyprus. *Blocklist*. Mar. 13, 2017. https://archive.fo/Z7WtK (visited on 03/19/2017) (cit. on p. 37).

[75] National Betting Authority of Cyprus. *Blocklist*. As it appeared in Google cache on Apr 27, 2017 05:09:03 GMT. Apr. 27, 2017. (Visited on 04/27/2017) (cit. on p. 37).

[76] National Betting Authority of Cyprus. *Blocklist*. May 25, 2017. https://web.archive.org/web/20170526021718/http://blocking.nba.com.cy (visited on 05/26/2017) (cit. on p. 37).

[77] National Betting Authority of Cyprus. *Blocklist website*. https://web.archive.org/web/20170605133936/http://blocking.nba.com.cy (visited on 06/05/2015) (cit. on pp. 37, 38).

[78] National Betting Authority of Cyprus. *Official website*. https://web.archive.org/web/20170605132348/http://nba.gov.cy (visited on 06/05/2015) (cit. on pp. 37, 40).

[79] *Cyprus National Betting Authority, Blocklist*. June 2021. https://nba.gov.cy/wp-content/uploads/BlockingListLatest.txt (cit. on pp. 80, 81).

[80] Cyta. *Cyta ISP official website*. http://archive.is/NBDpH (visited on 06/05/2017) (cit. on p. 40).

[81] D. Dagon, N. Provos, C. Lee, and W. Lee. *Corrupted DNS Resolution Paths: The Rise of a Malicious Resolution Authority*. in Proceedings of the Network and Distributed Security Symposium (NDSS), 2008 (cit. on p. 32).

[82] Google Developers. *Using Google Public DNS*. http://web.archive.org/web/20140829223153/https://developers.google.com/speed/public-dns/docs/using (visited on 08/29/2014) (cit. on p. 43).

[83] Arun Dunna, Ciarán O'Brien, and Phillipa Gill. *"Analyzing China's Blocking of Unpublished Tor Bridges"*. In: *Free and Open Communications on the Internet*. Usenix, 2018. https://www.usenix.org/system/files/conference/foci18/foci18-paper-dunna.pdf (cit. on p. 71).

[84] Aben E., Evdokimov L., and Xynou M. *Internet Access Disruption in Turkey*. 2016. https://web.archive.org/web/20170606190316/https://ooni.torproject.org/post/turkey-internet-access-disruption/ (visited on 06/06/2017) (cit. on p. 36).

[85] *ECLI:NL:RBMNE:2018:114, Rechtbank Midden-Nederland, C/16/448423 / KG ZA 17-382*. June 2021. https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBMNE:2018:114 (cit. on p. 83).

[86] Eeep. *"EEEP Blocklist Press Release"*. Retrieved July 2014. Aug. 2013. https://www.gamingcommission.gov.gr/images/apofaseis/lists/anakoinosi.pdf (cit. on p. 26).

[87] Eeep. *"Eeep Blocklist press release"*. July 2013. http://www.gamingcommission.gov.gr/images/deltia%5C%5Ftipou/dt2.pdf (cit. on pp. 19, 31).

[88] Eeep. *Press Release: "Responsible Play"*. http://web.archive.org/web/20140829223227/https://www.gamingcommission.gov.gr/index.php/el/ypefthino-paixnidi-ypmenu-im (cit. on p. 21).

[89] Eett. *Open Internet Report 2019-2020*. June 2020. https://ec.europa.eu/newsroom/dae/document.cfm?doc%5C%5Fid=68329 (cit. on p. 82).

[90] Office of Electronic Communications. *Report of the President of the Office of Electronic Communications on compliance in the Polish market with Regulation 2015/2120 on open internet access*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78874 (cit. on p. 75).

[91] Regulatory Authority for Electronic Communications and Postal Services. *Annual Report on Monitoring the Regulation (EU) 2015/2120 of the European Parliament and of the Council*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78879 (cit. on p. 75).

[92] Slovak Republic Regulatory Authority for Electronic Communications and Postal Services. *Annual Report on Monitoring the Regulation (EU) 2015/2120 of the European Parliament and of the Council*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78879 (cit. on pp. 75, 84).

[93] Urz d Komunikacji Elektronicznej. *Report on compliance in the Polish market with Regulation 2015/2120 on open internet access*. June 2020. https://ec.europa.eu/newsroom/dae/document.cfm?doc%5C%5Fid=68330 (cit. on p. 83).

[94] *Empfehlungen Clearingstelle Urheberrecht im Internet*. Jan. 2023. https://cuii.info/empfehlungen (cit. on pp. 80, 82).

[95] Roya Ensafi, Philipp Winter, Abdullah Mueen, and Jedidiah R. Crandall. *"Analyzing the Great Firewall of China Over Space and Time"*. In: *Privacy Enhancing Technologies* 2015.1 (2015). https://censorbib.nymity.ch/pdf/Ensafi2015a.pdf (cit. on pp. 45, 71).

[96] *EU sanctions against Russia following the invasion of Ukraine*. Mar. 2022. https://ec.europa.eu/info/strategy/priorities-2019-2024/stronger-europe-world/eu-solidarity-ukraine/eu-sanctions-against-russia-following-invasion-ukraine%5C%5Fen (cit. on p. 85).

[97] *European Commission - PRESS RELEASES - Press release - Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine*. July 2018. http://europa.eu/rapid/press-release%5C%5FIP-18-4581%5C%5Fen.htm (cit. on p. 67).

[98] Council of the European Union. *Regulation (EU) 2015/2120 of the European Parliament and of the Council*. Nov. 2015. https://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/9277-berec-guidelines-on-the-implementation-o_0.pdf (cit. on pp. 69, 76).

[99] The Council of the European Union. *Council regulation (EU) 2022/350 of 1 March 2022 amending Regulation (EU) No 833/2014 concerning restrictive measures in view of Russia's actions destabilising the situation in Ukraine*. Mar. 2022. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L:2022:065:FULL&from=EN (cit. on p. 85).

[100] Leonid Evdokimov. *metadb s3 tarx: fetch public OONI metadb backup from AWS S3 Open Data*. https://github.com/ooni/sysadmin/blob/master/scripts/metadb_s3_tarx. 2019 (cit. on p. 73).

[101] Benjamin Fabian, Annika Baumann, Mathias Ehlert, Vasilis Ververis, and Tatiana Ermakova. *"CORIA—Analyzing internet connectivity risks using network graphs"*. In: *2017 IEEE International Conference on Communications (ICC)*. Ieee. 2017, pp. 1–6 (cit. on p. 90).

[102] *Federal Law of 29.07.2017 276-FZ "On Amendments to the Federal Law" On Information, Information Technology and Information Protection*. Apr. 2019. http://publication.pravo.gov.ru/Document/View/0001201707300002?index=14&rangeSize=1 (cit. on p. 67).

[103] Arturo Filastò. *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International*. https://github.com/ooni/license/blob/master/data/CC4.0-BY-NC-SA.md. 2018 (cit. on p. 71).

[104] Arturo Filastò. *OONI MetaDB Sharing*. https://github.com/ooni/sysadmin/blob/master/docs/metadb-sharing.md. 2019 (cit. on p. 73).

[105] Arturo Filastò and Jacob Appelbaum. *"OONI: Open Observatory of Network Interference"*. In: *Free and Open Communications on the Internet*. Usenix, 2012. https://www.usenix.org/system/files/conference/foci12/foci12-final12.pdf (cit. on pp. 19, 38, 66).

[106] *Formal Internet Censorship: Copyright blocking injunctions*. Feb. 2019. https://www.openrightsgroup.org/blog/formal-internet-censorship-copyright-blocking-injunctions (cit. on p. 85).

[107] Fortinet. *Deep inspection*. https://docs.fortinet.com/document/fortigate/6.2.0/cookbook/122078/deep-inspection. 2019 (cit. on p. 53).

[108] *FortiOS 5.2: Security Profiles*. 2020. https://help.fortinet.com/fos50hlp/52data/Content/FortiOS/fortigate-whats-new-52/securityprofiles.htm (visited on 05/31/2020) (cit. on p. 53).

[109] Qurium Media Foundation. *Blocking Techniques Catalunya*. https://www.qurium.org/alerts/spain/blocking-techniques-catalunya. Oct. 2020 (cit. on pp. 49, 57).

[110] *France: Freedom on the Net 2020 Country Report: Freedom House*. June 2021. https://freedomhouse.org/country/france/freedom-net/2020 (cit. on p. 81).

[111] Bin Fu, Jialiu Lin, Lei Li, Christos Faloutsos, Jason Hong, and Norman Sadeh. *"Why people hate your app: Making sense of user feedback in a mobile app store"*. In: *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. Acm. 2013, pp. 1276–1284 (cit. on p. 63).

[112] *Gambling Regulatory Authority of Slovak Republic*. June 2021. https://www.urhh.sk/documents/20127/74354/Zoznam%5C%20zak%5C%C3%5C%A1zan%5C%C3%5C%BDch%5C%20webov%5C%C3%5C%BDch%5C%20s%5C%C3%5C%ADdiel%5C%20k%5C%2028.12.2020.pdf (cit. on pp. 80, 84).

[113] *Gaming Commission*. June 2021. https://www.gamingcommission.be/opencms/opencms/jhksweb%5C%5Fen/establishments/Online/blacklist/index.html (cit. on pp. 79, 80).

[114] *Gaming Control Authority under the Ministry of Finance of the Republic of Lithuania*. June 2021. https://lpt.lrv.lt/uploads/lpt/documents/files/neleg.txt (cit. on pp. 80, 83).

[115] Dead Gaudet. *"Tracking Without Cookies"*. Feb. 2003. http://web.archive.org/web/20140829223305/https://en.wikipedia.org/wiki/HTTP%5C%5FETag (cit. on p. 31).

[116] Genevieve Gebhart, Anonymous Author, and Tadayoshi Kohno. *"Internet Censorship in Thailand: User Practices and Potential Threats"*. In: *European Symposium on Security & Privacy*. Ieee, 2017. https://homes.cs.washington.edu/~yoshi/papers/GebhartEtAl-IEEEEuroSP.pdf (cit. on pp. 45, 71).

[117] John Geddes, Max Schuchard, and Nicholas Hopper. *"Cover your ACKs: pitfalls of covert channel censorship circumvention"*. In: *13). ACM, New York, NY, USA*. Proceedings of the ACM SIGSAC conference on Computer & Communications Security (CCS), 2013, pp. 361–372 (cit. on p. 18).

[118] Ahmad Ghazawneh and Ola Henfridsson. *"A paradigmatic analysis of digital application marketplaces"*. In: *Journal of Information Technology* 30.3 (2015), pp. 198–208 (cit. on p. 62).

[119] Digineo GmbH. *Public DNS Server List*. https://web.archive.org/web/20170606195759/https://public-dns.info/ (visited on 06/06/2017) (cit. on p. 38).

[120] kabelplus GmbH. *Gesperrte Websites wegen Urheberrechtsverletzungen*. Jan. 2023. https://www.kabelplus.at/specialpages/gesperrte-websites-wegen-urheberrechtsverletzungen-aspx (cit. on p. 79).

[121] kabelplus GmbH. *Update - bersicht der gesperrten Webseiten (VO 350/2022)*. Jan. 2023. https://www.kabelplus.at/privat/service/neuigkeiten/gesperrte-websites (cit. on p. 79).

[122] Devashish Gosain, Anshika Agarwal, Sahil Shekhawat, H. B. Acharya, and Sambuddho Chakravarty. *"Mending Wall: On the Implementation of Censorship in India"*. In: *SecureComm*. Springer, 2017. https://censorbib.nymity.ch/pdf/Gosain2017a.pdf (cit. on pp. 45, 71).

[123] *"Greek law, Law No. 4002/2011, Article 52"*. Apr. 2012. http://web.archive.org/web/20140829223000/http://nomoi.info/ (cit. on pp. 17, 19, 30).

[124] Samuel Greengard. *"Censored!"* In: *Communications of the ACM* (2010), pp. 16–18. https://dl.acm.org/doi/10.1145/1785414.1785423 (cit. on p. 85).

[125] Open Rights Groups. *Mobile Internet censorship: what's happening and what we can do about it*. May 2012. http://web.archive.org/web/20140411142104/https://www.openrightsgroup.org/ourwork/reports/mobile-internet-censorship:-whats-happening-and-what-we-can-do-about-it (cit. on p. 85).

[126] *Hadara Palestine*. 2012. http://web.archive.org/web/20141115030108/https://ooni.torproject.org/hadara-palestine.html (cit. on p. 18).

[127] Hakom. *Annual Report on the National Implementation of the Regulation (EU) 2015/2120 (period from 1th of May 2020 – 30th of April 2021)*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78852 (cit. on pp. 75, 81).

[128] *Hellenic Copyright Organization*. June 2021. https://opi.gr/images/epitropi/edppi%5C%5Flist%5C%5Fv8.pdf (cit. on pp. 80, 82).

[129] Noman Helmi. *Tunisian journalist sues government agency for blocking Facebook, claims damage for the use of 404 error message instead of 403*. Open Net Initiative, Sept. 2008. http://web.archive.org/web/20140829222940/http://opennet.net/node/950 (cit. on p. 24).

[130] Luis Hestres. *"App neutrality: Apple' s app store and freedom of expression online"*. In: (2013) (cit. on p. 62).

[131] Arne Hintz and Stefania Milan. *"'Through a Glass, Darkly" : Everyday Acts of Authoritarianism in the Liberal West"*. In: *International Journal of Communication* 12 (2018), pp. 3939–3959 (cit. on p. 45).

[132] Nguyen Phong Hoang, Panagiotis Kintis, Manos Antonakakis, and Michalis Polychronakis. *"An Empirical Study of the I2P Anonymity Network and Its Censorship Resistance"*. In: *Proceedings of the Internet Measurement Conference 2018*. Imc '18. Boston, MA, USA: Acm, 2018, pp. 379–392. ISBN: 978-1-4503-5619-0. DOI: 10.1145/3278532.3278565. http://doi.acm.org/10.1145/3278532.3278565 (cit. on p. 66).

[133] Sean Hollister. *"Russia bans Telegram encrypted messaging app"*. In: *Cnet* (Apr. 2018). https://www.cnet.com/news/russia-bans-telegram-encrypted-messaging-app (cit. on p. 67).

[134] John Holowczak and Amir Houmansadr. *"CacheBrowser: Bypassing Chinese Censorship without Proxies Using Cached Content"*. In: *Computer and Communications Security*. Acm, 2015. https://people.cs.umass.edu/~amir/papers/CacheBrowser.pdf (cit. on p. 71).

[135] Ralph Holz, Jens Hiller, Johanna Amann, Abbas Razaghpanah, Thomas Jost, Narseo Vallina-Rodriguez, and Oliver Hohlfeld. *"Tracking the Deployment of TLS 1.3 on the Web: A Story of Experimentation and Centralization"*. In: *SIGCOMM Comput. Commun. Rev.* 50.3 (July 2020), p. 315. ISSN: 0146-4833. DOI: 10.1145/3411740.3411742. https://doi.org/10.1145/3411740.3411742 (cit. on p. 56).

[136] Austin Hounsel, Prateek Mittal, and Nick Feamster. *"Automatically Generating a Large, Culture-Specific Blocklist for China"*. In: *Free and Open Communications on the Internet*. Usenix, 2018. https://www.usenix.org/system/files/conference/foci18/foci18-paper-hounsel.pdf (cit. on p. 71).

[137] Freedom of House. *Freedom of the Press, 2006, Cyprus Country report*. https://web.archive.org/web/20170605134525/https://freedomhouse.org/report/freedom-press/2006/cyprus (visited on 06/05/2015) (cit. on p. 36).

[138] Freedom of House. *Freedom of the Press, 2007, Cyprus Country report*. https://web.archive.org/web/20170605134610/https://freedomhouse.org/report/freedom-press/2007/cyprus (visited on 06/05/2015) (cit. on p. 36).

[139] Freedom of House. *Freedom of the Press, 2008, Cyprus Country report*. https://web.archive.org/web/20170605134650/https://freedomhouse.org/report/freedom-press/2008/cyprus (visited on 06/05/2015) (cit. on p. 36).

[140] Freedom of House. *Freedom of the Press, 2011, Cyprus Country report*. https://web.archive.org/web/20170605134800/https://freedomhouse.org/report/freedom-press/2011/cyprus (visited on 06/05/2015) (cit. on p. 36).

[141] Freedom of House. *Freedom of the Press, 2012, Cyprus Country report*. https://web.archive.org/web/20170605134838/https://freedomhouse.org/report/freedom-press/2012/cyprus (visited on 06/05/2015) (cit. on p. 36).

[142] Freedom of House. *Freedom of the Press, 2013, Cyprus Country report*. https://web.archive.org/web/20170605134916/https://freedomhouse.org/report/freedom-press/2013/cyprus (visited on 06/05/2015) (cit. on p. 36).

[143] Freedom of House. *Freedom of the Press, 2014, Cyprus Country report*. https://web.archive.org/web/20170605135021/https://freedomhouse.org/report/freedom-press/2014/cyprus (visited on 06/05/2015) (cit. on p. 36).

[144] *In re Apple iPhone Antitrust Litig., 846 F.3d 313 (9th Cir. 2017) (No. 17-204)*. 2018 (cit. on p. 67).

[145] OpenDNS Inc. *OpenDNS IP Addresses*. http://web.archive.org/web/20140829223158/http://www.opendns.com/opendns-ip-addresses/ (visited on 08/29/2014) (cit. on p. 43).

[146] Wikimedia Foundation Inc. *"Greek blog aggregation service administrator jailed"*. In: (Oct. 2006). http://web.archive.org/web/20140829223116/https://en.wikinews.org/wiki/Greek%5C%5Fblog%5C%5Faggregation%5C%5Fservice%5C%5Fadministrator%5C%5Fjailed (cit. on p. 17).

[147] OpenNet Initiative. *Internet Filtering in Burma in 2005: A Country Study*. https://opennet.net/studies/burma. Oct. 2020. https://opennet.net/studies/burma (cit. on p. 58).

[148] ooniprobe Installation Instructions and. https://github.com/thetorproject/ooni-probe#installation (cit. on p. 20).

[149] Berkman Klein Center for Internet and Society at Harvard University. *Lumen*. http://archive.is/BwyBv (visited on 06/05/2015) (cit. on p. 38).

[150] IPduh. *Internet Service Providers - Greece*. 2015. http://ipduh.com/macro/gr/isp/ (cit. on p. 17).

[151] Internet Censorship in Iran: A. First Look. *"Simurgh Aryan, Homa Aryan and J. Alex Halderman"*. In: *Free and Open Communications on the Internet (, DC, USA), USENIX* (cit. on p. 18).

[152] Forthnet ISP. *"Market Results"*. In: (Mar. 2014). http://web.archive.org/web/20140829223235/http://www.forthnet.gr/media/Company/anakinoseis/2014/FR,March (cit. on p. 24).

[153] Forthnet Isp. *Press Release: EEEP Blocklist*. 2013. http://web.archive.org/web/20131011002638/http://www.forthnet.gr/News.aspx?a%5C%5Fid=6787 (cit. on p. 33).

[154] Gunnar Eyal Wolf Iszaevich. *"Distributed Detection of Tor Directory Authorities Censorship in Mexico"*. In: *International Conference on Networks*. Iaria, 2019. https://tics.site/proceedings/2019a/icn%5C%5F2019%5C%5F6%5C%5F20%5C%5F38010.pdf (cit. on pp. 45, 71).

[155] *iTunes Search API - Affiliate Resources*. Mar. 2019. https://affiliate.itunes.apple.com/resources/documentation/itunes-store-web-service-search-api/#searchexamples (cit. on p. 64).

[156] Ity. *I. T. U. 2016. Measuring the Information Society Report*. https://web.archive.org/web/20170605134129/http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2016/MISR2016-w4.pdf (visited on 06/05/2015) (cit. on p. 36).

[157] Slinger Jansen and Ewoud Bloemendal. *"Defining App Stores: The Role of Curated Marketplaces in Software Ecosystems"*. In: *Software Business. From Physical Products to Software Services and Solutions*. Ed. by Georg Herzwurm and Tiziana Margaria. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 195–206. ISBN: 978-3-642-39336-5 (cit. on p. 68).

[158] Ben Jones, Tzu-Wen Lee, Nick Feamster, and Phillipa Gill. *Automated Detection and Fingerprinting of Censorship Block Pages*. ACM Internet Measurement Conference, 2014. http://conferences2.sigcomm.org/imc/2014/papers/p299.pdf (cit. on p. 20).

[159] Farid Shirazi abd Kathleen Greenaway. *"Examining Validity Claims for Internet Filtering in Islamic Middle Eastern Countries: A Critical Discourse Analysis"*. In: *AMCIS 2009 Proceedings*. 2009. http://aisel.aisnet.org/amcis2009/794 (cit. on p. 69).

[160] Gary King, Jennifer Pan, and Margaret E. Roberts. *"How Censorship in China Allows Government Criticism but Silences Collective Expression"*. In: *American Political Science Review* (2012). https://gking.harvard.edu/files/censored.pdf (cit. on p. 71).

[161] Gary King, Jennifer Pan, and Margaret E. Roberts. *"Reverse-engineering censorship in China: Randomized experimentation and participant observation"*. In: *Science* 345.6199 (2014). http://cryptome.org/2014/08/reverse-eng-cn-censorship.pdf (cit. on p. 71).

[162] Jeffrey Knockel, Masashi Crete-Nishihata, Jason Q. Ng, Adam Senft, and Jedidiah R. Crandall. *"Every Rose Has Its Thorn: Censorship and Surveillance on Social Video Platforms in China"*. In: *Free and Open Communications on the Internet*. Usenix, 2015. https://www.usenix.org/system/files/conference/foci15/foci15-paper-knockel.pdf (cit. on p. 71).

[163] Jeffrey Knockel, Lotus Ruan, and Masashi Crete-Nishihata. *"An analysis of automatic image filtering on WeChat Moments"*. In: *Free and Open Communications on the Internet*. Usenix, 2018. https://www.usenix.org/system/files/conference/foci18/foci18-paper-knockel.pdf (cit. on p. 71).

[164] Jeffrey Knockel, Lotus Ruan, and Masashi Crete-Nishihata. *"Measuring Decentralization of Chinese Keyword Censorship via Mobile Games"*. In: *Free and Open Communications on the Internet*. Usenix, 2017. https://www.usenix.org/system/files/conference/foci17/foci17-paper-knockel.pdf (cit. on p. 71).

[165] Citizen Lab et al. *URL testing lists intended for discovering website censorship*. https://github.com/citizenlab/test-lists. 2014. https://github.com/citizenlab/test-lists (cit. on p. 38).

[166] Citizen Lab et al. *URL testing lists intended for discovering website censorship*. 2014. https://github.com/citizenlab/test-lists (cit. on pp. 76, 77).

[167] C. S. Leberknight, M. Chiang, H. V. Poor, and F. Wong. *"A taxonomy of Internet censorship and anti-censorship"*. In: 2010 () (cit. on p. 18).

[168] Hee-Eun Lee, Tatiana Ermakova, Vasilis Ververis, and Benjamin Fabian. *"Detecting child sexual abuse material: A comprehensive survey"*. In: *Forensic Science International: Digital Investigation* 34 (2020), p. 301022 (cit. on p. 90).

[169] Soo Ling Lim, Peter J Bentley, Natalie Kanakam, Fuyuki Ishikawa, and Shinichi Honiden. *"Investigating country differences in mobile app user behavior and challenges for software engineering"*. In: *IEEE Transactions on Software Engineering* 41.1 (2015), pp. 40–64 (cit. on p. 62).

[170] *LinkedIn - AppAddict*. Mar. 2019. https://www.appaddict.org/view.php?trackid=288429040 (cit. on p. 68).

[171] *LinkedIn 4.1.249 Download APK for Android - Aptoide*. Mar. 2019. https://linkedin-android.en.aptoide.com (cit. on p. 68).

[172] *Lista ostrzee przed niebezpiecznymi stronami*. June 2021. https://www.cert.pl/posts/2020/03/ostrzezenia%5C%5Fphishing/%5C#files (cit. on pp. 80, 83).

[173] *LIWEST Netzsperren*. Mar. 2020. http://netzsperre.liwest.at (cit. on p. 79).

[174] Lotteries and Gambling Supervisory Inspection Of Latvia. *Unlicensed interactive gambling websites blocked*. June 2021. https://www.iaui.gov.lv/images/Blokesana/domeni.txt (cit. on p. 83).

[175] *Lotteries And Gambling Supervisory Inspection Of Latvia*. June 2021. https://www.iaui.gov.lv/images/Blokesana/domeni.txt (cit. on p. 80).

[176] Graham Lowe, Patrick Winters, and Michael L. Marcus. *The Great DNS Wall of China*. Tech. rep. New York University, 2007. https://censorbib.nymity.ch/pdf/Lowe2007a.pdf (cit. on p. 71).

[177] Tord Lundström and Maria Xynou. *Evidence of Internet censorship during Catalonia' s independence referendum*. Open Observatory of Network Interference [OONI]. 2017. https://ooni.torproject.org/post/internet-censorship-catalonia-independence-referendum/ (cit. on pp. 45, 58).

[178] Daithı Mac Sıthigh. *"App law within: rights and regulation in the smartphone age"*. In: *International Journal of Law and Information Technology* 21.2 (2013), pp. 154–186 (cit. on p. 62).

[179] *Malta Gaming Authority*. Feb. 2021. https://www.mga.org.mt (cit. on p. 83).

[180] Gurmeet Singh Manku, Arvind Jain, and Anish Das Sarma. *"Detecting near-duplicates for web crawling"*. In: *Proceedings of the 16th international conference on World Wide Web - WWW '07*. ACM Press, 2007 (cit. on p. 47).

[181] Bill Marczak, Nicholas Weaver, Jakub Dalek, Roya Ensafi, David Fifield, Sarah McKune, Arn Rey, John Scott-Railton, Ron Deibert, and Vern Paxson. *"An Analysis of China's "Great Cannon""*. In: *Free and Open Communications on the Internet*. Usenix, 2015. https://www.usenix.org/system/files/conference/foci15/foci15-paper-marczak.pdf (cit. on pp. 45, 71).

[182] William Martin, Federica Sarro, Yue Jia, Yuanyuan Zhang, and Mark Harman. *"A survey of app store analysis for software engineering"*. In: *IEEE transactions on software engineering* 43.9 (2017), pp. 817–847 (cit. on p. 62).

[183] National Media and Communications Authority. *The state of open Internet in Hungary in 2021*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78861 (cit. on pp. 75, 82).

[184] Comisión Nacional de los Mercados y la Competencia. *CNMCData - Informe Trimestral*. Oct. 2020. http://data.cnmc.es/datagraph/jsp/inf%5C%5Ftrim.jsp (cit. on p. 49).

[185] Simon Migliano. *"Free VPN Apps: Chinese Ownership, Secretive Companies & Weak Privacy"*. In: *Top10VPN* (Dec. 2018). https://www.top10vpn.com/free-vpn-app-investigation (cit. on p. 67).

[186] Mehrab Bin Morshed, Michaelanne Dye, Syed Ishtiaque Ahmed, and Neha Kumar. *"When the Internet Goes Down in Bangladesh"*. In: *Computer-Supported Cooperative Work and Social Computing*. Acm, 2017. https://nehakumardotorg.files.wordpress.com/2014/03/p1591-bin-morshed.pdf (cit. on pp. 45, 71).

[187] Mtn. *MTN ISP official website*. http://web.archive.org/web/20170605020143/http://www.mtn.com.cy/ (visited on 06/05/2017) (cit. on p. 41).

[188] Multimax. *Multimax ISP official website*. http://web.archive.org/web/20170605015656/http://www.mmcyp.com (visited on 06/05/2017) (cit. on p. 41).

[189] Zubair Nabi. *"The Anatomy of Web Censorship in Pakistan"*. In: *Free and Open Communications on the Internet*. Usenix, 2013. https://censorbib.nymity.ch/pdf/Nabi2013a.pdf (cit. on pp. 45, 71).

[190] *Neplp*. June 2021. https://www.neplpadome.lv/lv/sakums/mediju-lietotajiem/ierobezoto-domenu-vardu-saraksts/ (cit. on pp. 80, 83).

[191] Linx Public Affairs News. *BT Cleanfeed: the facts*. Sept. 2014. http://web.archive.org/web/20140829222922/https://publicaffairs.linx.net/news/?p=154 (cit. on p. 24).

[192] Kei Yin Ng, Anna Feldman, and Chris Leberknight. *"Detecting Censorable Content on Sina Weibo: A Pilot Study"*. In: *Hellenic Conference on Artificial Intelligence*. Acm, 2018. https://censorbib.nymity.ch/pdf/Ng2018a.pdf (cit. on p. 71).

[193] Arian Akhavan Niaki, Shinyoung Cho, Zachary Weinberg, Nguyen Phong Hoang, Abbas Razaghpanah, Nicolas Christin, and Phillipa Gill. *"ICLab: A Global, Longitudinal Internet Censorship Measurement Platform"*. In: *Proceedings of the 41st IEEE Symposium on Security and Privacy*. May 2020 (cit. on pp. 46, 58).

[194] Aarnio Niko, Nieminen Klaus, Pallas Elina, and Priiki Marko. *Annual Net Neutrality Report 2021*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/79071 (cit. on p. 75).

[195] *NRA Gambling Authority blocklist*. June 2021. https://nra.bg/wps/portal/nra/gambling/Online-hazat (cit. on pp. 79, 80).

[196] Hack66 Observatory. *A custom set of tools to perform ooniprobe network measurements*. https://github.com/hack66/bet2512. 2017. https://github.com/hack66/bet2512 (cit. on p. 38).

[197] Ofcom. *Open letter to industry about new restrictions on the provision of certain internet services to, or for the benefit of, "designated persons"*. Apr. 2022. https://www.revk.uk/2022/04/the-latest-crazy-law.html (cit. on p. 85).

[198] Czech Telecommunication Office. *Report of the Czech telecommunications authority (for the period from 1st of May 2020 to 30th of April 2021)*. Aug. 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78853 (cit. on pp. 75, 81).

[199] *Official Webpage Of Hellenic Gaming Commission (Hgc)*. June 2021. https://www.gamingcommission.gov.gr/index.php/en (cit. on pp. 80, 82).

[200] *Oficiul National pentru Jocuri de Noroc*. June 2021. http://onjn.gov.ro/wp-content/uploads/Onjn.gov.ro/Acasa/BlackList/Lista-neagra.txt (cit. on pp. 80, 84).

[201] Ooni. *measurements files repository*. https://web.archive.org/web/20170606210652/https://measurements.ooni.torproject.org/ (visited on 06/06/2017) (cit. on pp. 38, 39).

[202] Ooni. *"ooniprobe Http Request Test Class"*. In: (2014). http://web.archive.org/web/20140829223145/https://gitweb.torproject.org/ooni-probe.git/blob/HEAD:/ooni/nettests/blocking/http%5C%5Frequests.py (cit. on p. 19).

[203] Ooni. *Web Connectivity test specification*. https://github.com/ooni/spec/blob/master/nettests/ts-017-web-connectivity.md. 2019 (cit. on pp. 47, 56, 72).

[204] *OONI Explorer - Open Data on Internet Censorship Worldwide*. May 2022. https://explorer.ooni.org (cit. on p. 72).

[205] *OONI Explorer - Open Data on Internet Censorship Worldwide: KPN Blockpage*. June 2020. https://explorer.ooni.org/m/01202006164675ba341d04b443650f0e8dd3c5b9 (cit. on pp. 80, 83).

[206] *"OONI Url repository and EEEP Blocklist"*. 2015. https://web.archive.org/web/20150514103452/https://github.com/hellais/ooni-inputs/blob/master/processed/bycountry/GR/urls/EEEP%5C%5FBlacklist.txt (cit. on pp. 20, 25, 31).

[207] *Open Observatory of Network Interference*. 2020. https://ooni.org (cit. on pp. 45–47, 70).

[208] Block Landing Page. Cyta ISP. http://web.archive.org/web/20140829223216/http://www.cyta.gr/misc/eeep%5C%5Fnoaccess (cit. on p. 24).

[209] Jong Chun Park and Jedidiah R. Crandall. *"Empirical Study of a National-Scale Distributed Intrusion Detection System: Backbone-Level Filtering of HTML Responses in China"*. In: *Distributed Computing Systems*. Ieee, 2010, pp. 315–326. https://www.cs.unm.edu/~crandall/icdcs2010.pdf (cit. on p. 71).

[210] Michelle Paulson and Wikimedia Foundation Inc. *"Wikimedia Foundation supports Wikipedia user subject to defamation lawsuit in Greece"*. In: (Feb. 2014). http://web.archive.org/web/20140829223131/https://blog.wikimedia.org/2014/02/14/wikimedia-foundation-supports-wikipedia-user-subject-to-defamation-lawsuit-in-greece/ (cit. on p. 17).

[211] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nick Weaver, and Vern Paxson. *"Global Measurement of DNS Manipulation"*. In: *USENIX Security Symposium*. Usenix, 2017. https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-pearce.pdf (cit. on p. 58).

[212] Ella Peltonen, Eemil Lagerspetz, Jonatan Hamberg, Abhinav Mehrotra, Mirco Musolesi, Petteri Nurmi, and Sasu Tarkoma. *"The hidden image of mobile apps: geographic, demographic, and cultural factors in mobile usage"*. In: *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. Acm. 2018, p. 10 (cit. on p. 62).

[213] Marta Poblet. *"Distributed, privacy-enhancing technologies in the 2017 Catalan referendum on independence: New tactics and models of participatory democracy"*. In: *First Monday* 23.12 (2018) (cit. on pp. 45, 58, 69).

[214] Euro2day News Portal. *Forthnet Market Result First Quarter 2014*. May 2014. http://web.archive.org/web/20140829223245/http://www.euro2day.gr/news/enterprises/article/1219910/forthnet-afxhsh-esodon-lianikhs-kai-syndromhton-s.html (cit. on p. 24).

[215] Torrentfreak news portal. *""Greek Court Orders ISP Blockades of Pirate Music Sites""*. May 2012. http://web.archive.org/web/20140829222910/https://torrentfreak.com/greek-court-orders-isp-blockades-of-pirate-music-sites-120521 (cit. on p. 17).

[216] Market Regulation Department Swedish Post and Telecom Authority. *Open internet - annual reporting 2020/2021*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78878 (cit. on p. 75).

[217] Belgian Institute for Postal Services and Telecommunications. *Report regarding the monitoring of net neutrality in Belgium (period from 1 May 2020 - 30 April 2021)*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78849 (cit. on pp. 75, 79).

[218] J. Postel and J. Reynolds. *Assigned Numbers*. Rfc 1700. RFC Editor, Oct. 1994. https://www.rfc-editor.org/rfc/rfc1700.txt (cit. on p. 54).

[219] Estonian Consumer Protection and Technical Regulatory Authority. *Report on the Estonian Consumer Protection and Technical Regulatory Authority' s work on the implementation of the EU Net Neutrality Regulation*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78857 (cit. on pp. 75, 81).

[220] Ram Sundara Raman, Prerana Shenoy, Katharina Kohls, and Roya Ensafi. *"Censored Planet: An Internet-wide, Longitudinal Censorship Observatory"*. In: *Computer and Communications Security*. Acm, 2020 (cit. on p. 46).

[221] Ram Sundara Raman, Adrian Stoll, Jakub Dalek, Reethika Ramesh, Will Scott, and Roya Ensafi. *"Measuring the Deployment of Network Censorship Filters at Global Scale"*. In: *Network and Distributed System Security*. The Internet Society, 2020. https://censorbib.nymity.ch/pdf/Raman2020a.pdf (cit. on p. 58).

[222] Reethika Ramesh, Ram Sundara Raman, Matthew Bernhard, Victor Ongkowijaya, Leonid Evdokimov, Anne Edmundson, Steven Sprecher, Muhammad Ikram, and Roya Ensafi. *"Decentralized Control: A Case Study of Russia"*. In: *Network and Distributed System Security*. The Internet Society, 2020. https://www.ndss-symposium.org/wp-content/uploads/2020/02/23098.pdf (cit. on pp. 45, 71).

[223] Abbas Razaghpanah, Anke Li, Arturo Filastò, Rishab Nithyanand, Vasilis Ververis, Will Scott, and Phillipa Gill. *"Exploring the Design Space of Longitudinal Censorship Measurement Platforms"*. In: *CoRR* abs/1606.01979 (2016). arXiv: 1606.01979. http://arxiv.org/abs/1606.01979 (cit. on p. 90).

[224] Magenta Redaktion. *"Netzsperre: Was bedeutet "Diese Seite ist gesperrt"?"* In: *Magenta* (Aug. 2022). https://blog.magenta.at/internet/sicherheit/netzsperre (cit. on p. 79).

[225] Luxembourg Institute of Regulation. *Access to an open internet in Luxembourg - activity report*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78864 (cit. on p. 75).

[226] *Rejestr domen - Rejestr Domen Słu cych do Oferowania Gier Hazardowych Niezgodnie z Ustaw* . June 2021. https://hazard.mf.gov.pl/api/Register (cit. on pp. 80, 83).

[227] Communications Regulatory Authority of the Republic of Lithuania. *Open internet and implementation of the regulation (EU) 2015/2120 in Lithuania Report to the European Commission*. June 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78860 (cit. on pp. 75, 83).

[228] David Robinson, Harlan Yu, and Anne An. *Collateral Freedom: A Snapshot of Chinese Internet Users Circumventing Censorship*. Tech. rep. OpenITP, 2013. https://www.upturn.org/static/files/CollateralFreedom.pdf (cit. on p. 71).

[229] Goople Play Store in Russia. *LinkedIn application not found*. Mar. 2019. http://web.archive.org/web/20190313183244/https://play.google.com/store/apps/details?id=com.linkedin.android&gl=ru (cit. on p. 66).

[230] *Russia to Fine Search Engines for Linking to Banned VPN services*. Apr. 2019. https://thehackernews.com/2018/06/russian-vpn-services.html (cit. on p. 67).

[231] *Russia Will Penalize For Using and Promoting VPN Services, So Now What? | VPNBase*. July 2018. https://vpnbase.com/blog/russia-will-penalize-for-using-and-promoting-vpn-services (cit. on p. 67).

[232] *Russian telecom watchdog blocks 50 VPN services and anonymizers for Telegram access*. Apr. 2019. http://tass.com/economy/1002762 (cit. on p. 67).

[233] Walid Al-Saqaf. *"Internet Censorship Circumvention Tools: Escaping the Control of the Syrian Regime"*. In: *Media and Communication* 4.1 (2016). http://www.cogitatiopress.com/ojs/index.php/mediaandcommunication/article/download/357/357 (cit. on pp. 45, 71).

[234] Pekka Savola. *"Internet Connectivity Providers as Involuntary Copyright Enforcers: Blocking Websites in Particular"*. en. In: (Mar. 2018), p. 300 (cit. on pp. 45, 69, 71).

[235] Maria José Schmidt-Kessen, Julia Hörnle, and Alan Littler. *"Preventing Risks from Illegal Online Gambling Using Effective Legal Design on Landing Pages"*. en. In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3474296. https://www.ssrn.com/abstract=3474296 (visited on 04/05/2020) (cit. on p. 45).

[236] Maria José Schmidt-Kessen, Julia Hörnle, and Alan Littler. *"Preventing Risks from Illegal Online Gambling Using Effective Legal Design on Landing Pages"*. en. In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3474296. https://www.ssrn.com/abstract=3474296 (visited on 04/05/2020) (cit. on pp. 69, 71).

[237] *Scripts for querying mobile app stores (Google Play store, Apple iTunes, Tencent MyApp) and results from queries in different countries*. Mar. 2019. https://github.com/hack66/appavail (cit. on pp. 64, 66).

[238] SecurityTrails. *Domain Security, DNS Trails and IP Tools*. 2020. https://securitytrails.com/ (visited on 06/02/2020) (cit. on p. 58).

[239] Suranga Seneviratne, Aruna Seneviratne, Mohamed Ali Kaafar, Anirban Mahanti, and Prasant Mohapatra. *"Early detection of spam mobile apps"*. In: *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee. 2015, pp. 949–959 (cit. on p. 62).

[240] StatCounter. *GlobalStats*. http://gs.statcounter.com/ (visited on 06/05/2015) (cit. on p. 39).

[241] Paul Sturges. *"Access Denied: The Practice and Policy of Global Internet Filtering"*. In: *The Electronic Library* 26.6 (Nov. 2008), pp. 924–925 (cit. on p. 45).

[242] *Supported locations for distribution to Google Play users - Play Console Help*. Mar. 2019. https://support.google.com/googleplay/android-developer/table/3541286 (cit. on p. 64).

[243] Christos Syllas. *Free speech takes a beating in Greece*. Index on Censorship Organization, Mar. 2013. http://web.archive.org/web/20140829223121/http://www.indexoncensorship.org/2013/03/free-speech-takes-a-beating-in-greece/ (cit. on p. 17).

[244] *T-Mobile Usa Web Guard*. Mar. 2012. https://web.archive.org/web/20141115030554/https://ooni.torproject.org/t-mobile-usa-web-guard.html (cit. on p. 18).

[245] *Tab-Tab, Come in! Bypassing Internet blocking to categorize DPI devices*. May 2013. https://web.archive.org/web/20130926190044/https://ooni.torproject.org/tab-tab-come-in-bypassing-internet-blocking-to-categorize-dpi-devices.html (cit. on p. 18).

[246] Rima Tanash, Zhouhan Chen, Dan Wallach, and Melissa Marschall. *"The Decline of Social Media Censorship and the Rise of Self-Censorship after the 2016 Failed Turkish Coup"*. In: *Free and Open Communications on the Internet*. Usenix, 2017. https://www.usenix.org/system/files/conference/foci17/foci17-paper-tanash.pdf (cit. on pp. 45, 71).

[247] Rima S. Tanash, Zhouhan Chen, Tanmay Thakur, Dan S. Wallach, and Devika Subramanian. *"Known Unknowns: An Analysis of Twitter Censorship in Turkey"*. In: *Workshop on Privacy in the Electronic Society*. Acm, 2015. https://censorbib.nymity.ch/pdf/Tanash2015a.pdf (cit. on pp. 45, 71).

[248] Chongbin Tang, Sen Chen, Lingling Fan, Lihua Xu, Yang Liu, Zhushou Tang, and Liang Dou. *"A Large-Scale Empirical Study on Industrial Fake Apps"*. In: *CoRR* abs/1902.00647 (2019). arXiv: 1902.00647. http://arxiv.org/abs/1902.00647 (cit. on pp. 63, 66).

[249] adslgr.com Technological Forum. *"Pirate Bay blocking"*. Feb. 2010. http://web.archive.org/web/20120112215459/http://www.adslgr.com/forum/showpost.php?p=3325943&postcount=14 (cit. on p. 17).

[250] National Telecommunications and Postal Commission. *Open Internet report 2020-2021*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78858 (cit. on p. 75).

[251] State Secretariat for Telecommunications, Digital Infrastructures of the Ministry of Economic Affairs, and Digital Transformation. *Report on Spain's Supervision of European Regulations on the Open Internet Access (Net neutrality)*. July 2021. https://ec.europa.eu/newsroom/dae/redirection/document/78882 (cit. on p. 75).

[252] *Teleindustrien*. June 2021. https://www.teleindu.dk/brancheholdninger/blokeringer-pa-nettet (cit. on p. 80).

[253] *Telekom Control Kommission (TKK)*. Jan. 2023. https://www.rtr.at/TKP/wer_wir_sind/tkk/TKK.de.html (cit. on pp. 79, 80).

[254] *Tencent - Terms of Service*. Jan. 2019. https://www.tencent.com/en-us/zc/termsofservice.shtml (cit. on p. 61).

[255] *Tencent App Store*. Mar. 2019. https://android.myapp.com (cit. on p. 61).

[256] *The Gaming Commission - The regulator of the gambling sector in Belgium*. Jan. 2023. https://www.gamingcommission.be/en (cit. on p. 79).

[257] *The latest crazy law*. Apr. 2022. https://www.revk.uk/2022/04/the-latest-crazy-law.html (cit. on p. 85).

[258] *The Raspberry Pi Foundation*. http://web.archive.org/web/20141111081017/http://www.raspberrypi.org/ (cit. on p. 14).

[259] *The Republic of Croatia, Ministry of Finance, Tax Administration - Blocklist*. June 2021. https://www.porezna-uprava.hr/Dokumenti%20razno/Nedozvoljeno%20obavljanje%20djelatnosti%20igara%20na%20sre%C4%87u%20putem%20interneta/Popis%20web%20adresa%20prire%C4%91iva%C4%8Da%20igara%20na%20sre%C4%87u%20za%20koje%20je%20izdan%20nalog%20o%20zabrani%20rada.pdf (cit. on pp. 80, 81).

[260] *Top App Stores & mobile apps in China for Q1 2017*. Apr. 2017. https://www.chinainternetwatch.com/20410/top-app-stores-mobile-apps-q1-2017 (cit. on p. 61).

[261] Point Topic. *"Internet service provider market share in Greece fourth quarter 2013"*. Aug. 2014. https://web.archive.org/web/20150514213631/http://point-topic.com/free-analysis/greece-broadband-overview/ (cit. on p. 17).

[262] *UKE przyst pił do porozumienia chroni cego abonentw - Urz d Komunikacji Elektronicznej*. June 2021. https://www.uke.gov.pl/akt/uke-przystapil-do-porozumienia-chroniacego-abonentow,300.html (cit. on p. 83).

[263] Benjamin VanderSloot, Allison McDonald, Will Scott, J. Alex Halderman, and Roya Ensafi. *"Quack: Scalable Remote Measurement of Application-Layer Censorship"*. In: *USENIX Security Symposium*. Usenix, 2018. https://www.usenix.org/system/files/conference/usenixsecurity18/sec18-vandersloot.pdf (cit. on pp. 46, 58).

[264] Ververis Vasilis, Marguel Sophia, and Fabian Benjamin. *"Cross-Country Comparison of Internet Censorship: A Literature Review"*. In: *Policy & Internet* 12.4 (2020), pp. 450–473. https://onlinelibrary.wiley.com/doi/full/10.1002/poi3.228 (cit. on p. 89).

[265] EEEP Blocklist Fourth Version. July 2014. https://www.gamingcommission.gov.gr/images/Anakoinoseis/BlackListVersion4%5C%5F11072014.pdf (cit. on pp. 31, 38).

[266] Vasilis Ververis. *Lepidopter, OONI powered Raspberry Pi image*. https://github.com/TheTorProject/lepidopter (cit. on pp. 14, 20).

[267] Vasilis Ververis. *"Network measurements analysis of Spanish ISPs"*. en. In: (2021). DOI: 10.5281/zenodo.4743905. https://zenodo.org/record/4743905 (cit. on pp. 49, 51, 53, 56, 58).

[268] Vasilis Ververis. *OONI releases Lepidopter Raspberry Pi distribution*. Jan. 2016. https://ooni.org/post/lepidopter (cit. on p. 14).

[269] Vasilis Ververis, Tatiana Ermakova, Marios Isaakidis, Simone Basso, Benjamin Fabian, and Stefania Milan. *"Understanding Internet Censorship in Europe: The Case of Spain"*. In: *13th ACM Web Science Conference 2021*. WebSci '21. Virtual Event, United Kingdom: Association for Computing Machinery, 2021, pp. 319–328. ISBN: 9781450383301. DOI: 10.1145/3447535.3462638. https://doi.org/10.1145/3447535.3462638 (cit. on pp. 69, 71, 76, 89).

[270] Vasilis Ververis, Tatiana Ermakova, Lucas Lasota, and Benjamin Fabian. *"Website Blocking in the European Union: Network Interference from the Perspective of Open Internet"*. In: *Policy and Internet* (forthcoming 2023). DOI: 10.1002/poi3.367 (cit. on p. 89).

[271] Vasilis Ververis, Fadelkon, Ana, Bita, and Samba. *Women on Web website censored in Spain*. Magma. May 2020. https://blog.magma.lavafeld.org/post/women-on-web-blocking/ (cit. on p. 45).

[272] Vasilis Ververis, Marios Isaakidis, Chrystalleni Loizidou, and Benjamin Fabian. *"Internet Censorship Capabilities in Cyprus: An Investigation of Online Gambling Blocklisting"*. In: *E-Democracy*. Springer, 2017. https://censorbib.nymity.ch/pdf/Ververis2017a.pdf (cit. on pp. 45, 69, 71, 81, 89).

[273] Vasilis Ververis, Marios Isaakidis, Valentin Weber, and Benjamin Fabian. *"Shedding Light on Mobile App Store Censorship"*. In: *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization, UMAP' 19 Adjunct*. ACM, 2019. https://dl.acm.org/doi/10.1145/3314183.3324965 (cit. on p. 85).

[274] Vasilis Ververis, Marios Isaakidis, Valentin Weber, and Benjamin Fabian. *"Shedding Light on Mobile App Store Censorship"*. In: *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*. UMAP'19 Adjunct. Larnaca, Cyprus: Association for Computing Machinery, 2019, pp. 193–198. ISBN: 9781450367110. DOI: 10.1145/3314183.3324965. https://doi.org/10.1145/3314183.3324965 (cit. on p. 89).

[275] Vasilis Ververis, George Kargiotakis, Arturo Filastò, Benjamin Fabian, and Afentoulis Alexandros. *"Understanding Internet Censorship Policy: The Case of Greece"*. In: *Free and Open Communications on the Internet*. Usenix, 2015. https://www.usenix.org/system/files/conference/foci15/foci15-paper-ververis-updated-2.pdf (cit. on pp. 36, 37, 45, 58, 89).

[276] Vasilis Ververis, George Kargiotakis, Arturo Filastò, Benjamin Fabian, and Afentoulis Alexandros. *"Understanding Internet Censorship Policy: The Case of Greece"*. In: *Free and Open Communications on the Internet*. Usenix, 2015. https://www.usenix.org/system/files/conference/foci15/foci15-paper-ververis-updated-2.pdf (cit. on pp. 69–71, 76, 82, 84, 85).

[277] Vasilis Ververis, Olga Khrustaleva, and Eliana Quiroz. *"Network interference in Latin America: Evaluating network measurements to detect information controls and Internet censorship"*. In: (Nov. 2017). ISSN: 2175-9596. https://lavits.org/wp-content/uploads/2018/04/32-Vasilis-Ververis-Olga-Khrustaleva-e-Eliana-Quiroz.pdf (cit. on p. 89).

[278] Haoyu Wang, Zhe Liu, Jingyue Liang, Narseo Vallina-Rodriguez, Yao Guo, Li Li, Juan Tapiador, Jingcun Cao, and Guoai Xu. *"Beyond Google Play: A Large-Scale Comparative Study of Chinese Android App Markets"*. In: *Proceedings of the Internet Measurement Conference 2018*. Imc '18. Boston, MA, USA: Acm, 2018, pp. 293–307. ISBN: 978-1-4503-5619-0. DOI: 10.1145/3278532.3278558. http://doi.acm.org/10.1145/3278532.3278558 (cit. on p. 66).

[279] Women on Web. *"Courtcase against Spanish government for blocking abortion website during COVID19 @ Women on Web"*. In: (2021). https://www.womenonweb.org/en/page/20678/courtcase-against-spanish-government-for-blocking-abortion-website (cit. on p. 55).

[280] EEEP Official Website. https://www.gamingcommission.gov.gr/ (cit. on p. 30).

[281] Tor Project Website. https://www.torproject.org (cit. on p. 32).

[282] Florian Weimer. *"Passive dns replication"*. In: *17th Annual FIRST Conference on Computer Security Incident Handling (FIRST '05)*. 2005 (cit. on p. 32).

[283] *Welcome to Steam*. Apr. 2019. https://store.steampowered.com (cit. on p. 61).

[284] *What is Russia's new VPN law all about?* Apr. 2019. https://www.bbc.com/news/technology-41829726 (cit. on p. 67).

[285] Philipp Winter and Stefan Lindskog. *"How the Great Firewall of China is Blocking Tor"*. In: *Free and Open Communications on the Internet*. Usenix, 2012. https://www.usenix.org/system/files/conference/foci12/foci12-final2.pdf (cit. on p. 71).

[286] Joss Wright. *Regional Variation in Chinese Internet Filtering*. Tech. rep. University of Oxford, 2012. https://papers.ssrn.com/sol3/Delivery.cfm/SSRN%5C%5FID2265775%5C%5Fcode1448244.pdf?abstractid=2265775&mirid=3 (cit. on p. 71).

[287] Joss Wright, Tulio Souza, and Ian Brown. *"Fine-Grained Censorship Mapping: Information Sources, Legality and Ethics"*. In: (). http://static.usenix.org/event/foci11/tech/final%5C%5Ffiles/Wright.pdf (cit. on p. 18).

[288] Xueyang Xu, Z. Morley Mao, and J. Alex Halderman. *"Internet Censorship in China: Where Does the Filtering Occur?"* In: *Passive and Active Measurement Conference*. Springer, 2011, pp. 133–142. https://web.eecs.umich.edu/~zmao/Papers/china-censorship-pam11.pdf (cit. on p. 71).

[289] Tarun Kumar Yadav, Akshat Sinha, Devashish Gosain, Piyush Kumar Sharma, and Sambuddho Chakravarty. *"Where The Light Gets In: Analyzing Web Censorship Mechanisms in India"*. In: *Internet Measurement Conference*. Acm, 2018. http://delivery.acm.org/10.1145/3280000/3278555/p252-Yadav.pdf (cit. on pp. 45, 71).

[290] *"Zambia, a country under Deep Packet Inspection"*. July 2013. http://web.archive.org/web/20141005212653/https://ooni.torproject.org/zambia-a-country-under-deep-packet-inspection.html (cit. on p. 18).

[291] Jonathan L. Zittrain and John G. Jr. Palfrey. *Access Denied: The Practice and Policy of Global Internet Filtering*. Oxford Internet Institute, Research Report No. 14. June 2007. https://www.oii.ox.ac.uk/archive/downloads/publications/RR14.pdf (cit. on p. 69).

[292] *Zveejovan daje ze Seznamu nepovolench internetovch her k 29.6.2021*. June 2021. https://www.mfcr.cz/cs/soukromy-sektor/hazardni-hry/seznam-nepovolenych-internetovych-her/2021/zverejnovane-udaje-ze-seznamu-nepovoleny-42322 (cit. on pp. 80, 81).

[293] **中□人民共和国网□安全法**. Apr. 2019. http://web.archive.org/web/20190331080839/http://www.npc.gov.cn/npc/xinwen/2016-11/07/content%5C%5F2001605.htm (cit. on p. 67).

[294] **工□和信息化部关于清理□范互□网网□接入服□市□的通知**. Apr. 2019. http://www.miit.gov.cn/n1146290/n4388791/c5471946/content.html (cit. on p. 67).

– A –

Index

– A –

Glossary

API Application Programming Interface. 30, 72, 83

AS Autonomous System. 38, 48–50, 55, 66, 72–74, 77, 85

CC Cocos Islands. 65

CG Country Groups. 64–66

CSAM Child Sexual Abuse Material. 79, 81, 84

CSV Comma-separated Values. 78, 81, 84

CX Christmas Island. 65

DNS Domain Name System. 20, 22–27, 30–32, 35, 38–40, 42, 43, 45–48, 51, 54, 56–59, 72, 74, 79, 82, 84

DoH DNS over HTTPS. 57

DPI Deep Packet Inspection. 22, 25, 26, 31, 32, 37, 46, 53, 55, 57–59

EEEP Hellenic Gaming Commission. 19–21, 25, 26, 30–33, 36

EU European Union. 8, 9, 11–13, 15, 35, 59, 69–71, 73, 76–78, 81, 84, 85, 87

GM Guam. 65

HM Heard Island and McDonald Islands. 65

HTML HyperText Markup Language. 78

HTTP Hypertext Transfer Protocol. 19, 22–25, 30–32, 37, 39, 40, 45–49, 51, 54–59, 66, 72, 74, 82

HTTPS Hypertext Transfer Protocol Secure. 25, 30–32, 37, 40, 55–57

I2P Invisible Internet Project. 65, 66

IO British Indian Ocean Territory. 65

IPFS InterPlanetary File System. 49, 51, 54, 58

IPv4 Internet Protocol version 4. 54

IPv6 Internet Protocol version 6. 26

ISP Internet Service Provider. 11, 13–15, 17–28, 30–33, 35–41, 43–47, 49, 51, 53–55, 57–59, 69, 70, 74, 79, 81–85

KI Kiribati. 65

KP North Korea. 64

MH Marshall Islands. 65

MP Northern Mariana Islands. 65

MX Mail Exchanger. 25–28, 42, 43

NBA National Betting Authority. 36–40

NF Norfolk Island. 65

NR Nauru. 65

NRA National Regulatory Authority. 13, 15, 70, 74, 76, 79, 81–84

NS Name Server. 26, 32

OONI Open Observatory of Network Interference. 8, 9, 13, 14, 18, 19, 35, 36, 38, 39, 45–50, 53, 54, 56, 58, 59, 66, 70–74, 83–85, 87

OS Operating System. 14, 61

PDF Portable Document Format. 30, 73, 77–79, 81, 84

RoC Republic of Cyprus. 35–38, 43

RSS Rich Site Summary. 17

SMTP Simple Mail Transfer Protocol. 26, 27

SNI Server Name Indication. 46, 55–57

SY Syria. 64

TCP/IP Transmission Control Protocol/Internet Protocol. 12

TLS Transport Layer Security. 13, 46, 55–57

TSOL Telefonica Solutions Group. 49, 59

UM United States Minor Outlying Islands. 65

URI Uniform Resource Identifier. 47, 72

URL Uniform Resource Locator. 11, 19–26, 30–32, 37, 38, 49, 51, 53, 56, 66, 72, 77, 84, 85

UUID Universal Unique Identifier. 53

VPN Virtual Private Network. 33

WoW Women On Web. 45, 46, 48, 54–59

WWW World Wide Web. 12

XLSX Excel Workbook. 78

XML Extensible Markup Language. 78, 83