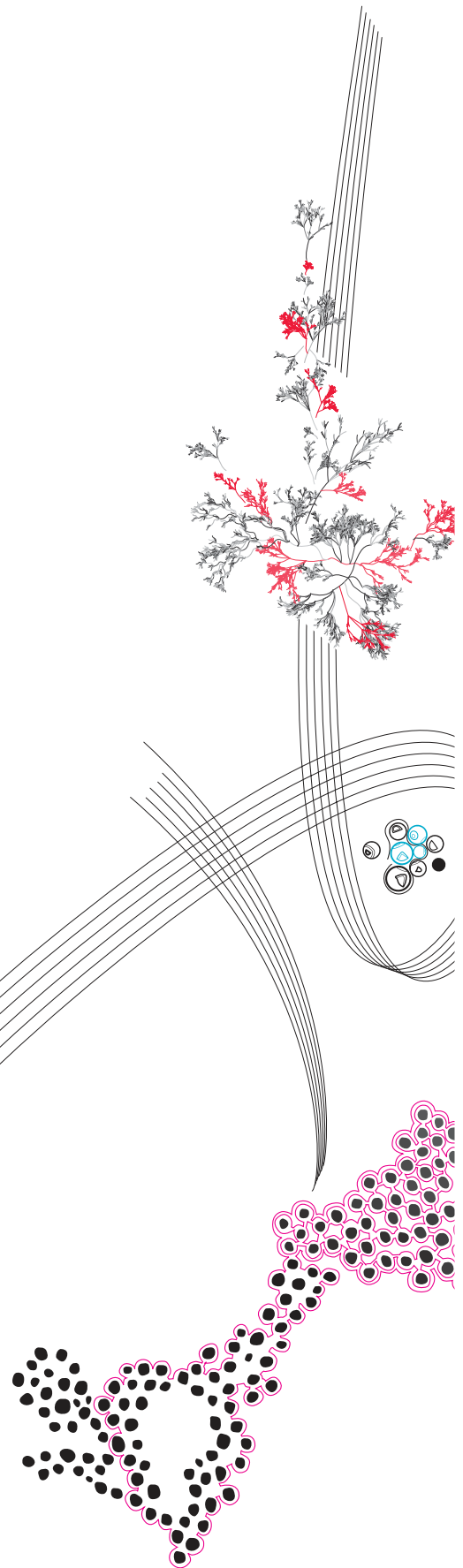




MSc Computer Science
Final Project

Parental control effectiveness: Evaluating filtering solutions of the leading Italian ISPs

Chiara Rizzato

Supervisor: Antonia Affinito, Matteo Liberato

June, 2025

Department of Computer Science
Faculty of Electrical Engineering,
Mathematics and Computer Science,
University of Twente

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Research questions	2
1.3	Contribution	3
1.4	Structure	3
2	Background	4
2.1	Parental control	4
2.1.1	Parental control solutions	4
2.2	European Regulatory approaches to protect children online	5
2.3	AGCOM resolution	6
2.4	Italian ISPs compliance with the AGCOM resolution	7
2.5	Related work	8
2.5.1	Parental control vulnerabilities	8
2.5.2	Innovative frameworks	9
2.5.3	Censorship and filtering techniques	9
3	Methodology	11
3.1	Dataset	11
3.1.1	Categorization and filtering	12
3.2	Setup	12
3.2.1	Internet traffic measurements	13
3.2.2	Output analysis	14
3.2.3	Wireshark	14
3.3	Use cases	14
3.3.1	Research question 1	15
3.3.2	Research question 2	16
3.3.3	Research question 3	19
4	Results	21
4.1	Network-Level Blocking Techniques	21
4.1.1	Technical filtering mechanism	21
4.1.2	User experience and customization	26
4.2	Effectiveness of Parental Control Filtering	27
4.2.1	Underblocking	27
4.2.2	Category analysis	29
4.3	Bypassing Techniques	30
4.3.1	VPNs results	31
4.3.2	HTTP requests results	31

4.3.3	Google DNS results	32
5	Discussion	34
5.1	Limitations and future work	35
6	Conclusions	36
6.1	Research answers	36

Abstract

The protection of minors' online activities is a critical concern that requires careful attention. Today, children and teenagers spend a considerable amount of time online for educational or recreational purposes throughout the day. However, the more time they spend online, the greater the risk of being exposed to harmful content such as violence, pornography, or tasteless content. One possible way to protect minors is to monitor and filter their online activities through parental control solutions. In some cases, this issue was tackled at the institutional level, for instance, in Italy, the AGCOM resolution 'Protecting minors in cyberspace' requires all ISPs to offer a basic, free-of-charge parental control mechanism that filters web content based on categories to their customers. Motivated by the lack of investigation into those solutions and the issues regarding the effectiveness of parental control tools in the existing literature, our aim is to provide a comprehensive overview of the parental control systems provided by the main Italian ISPs. Through the use of a curated blocklist, we seek to address our research questions related to the technical aspects of the solutions implemented, their effectiveness, and their robustness against various bypassing techniques. By sending DNS queries and HTTPS requests to the domains in the blocklist, we determined the network-level filtering mechanism in place and the presence of underblocking, i.e., harmful domains that remain accessible. Our results reveal two different implementation approaches among the three ISPs analyzed: DNS-based and TLS-based filters. Moreover, we identified a notable amount of harmful websites remaining unblocked, indicating a significant level of underblocking that ranges from 20,5% to 34,2%. Furthermore, the most common harmful categories - pornography, adult content, gambling - are also the ones most affected by underblocking. The results further reveal inconsistencies among providers, with underblocking ranging from 2,9% to 50%. Finally, the response of the filtering mechanisms to different bypassing techniques reflects the type of solution in place and raises concerns regarding security vulnerabilities. In fact, from our results it is evident that when a DNS-based filter is in place, changing the DNS server is an effective method to bypass it. On the other hand, in the case of TLS-based filtering, by sending HTTP requests instead of HTTPS, allows users to circumvent the filtering block.

Keywords: Parental control, ISP, filtering, underblocking, security

Chapter 1

Introduction

The safety of children and teenagers accessing the Internet is an important issue that requires careful consideration and investigation. The young generation spends a significant amount of time in front of screens for various purposes throughout the day. In fact, according to the EU Kids Online report [21], over 80% of European minors between the ages of 9 and 16 access the Internet at least once a day using a smartphone, which is the most commonly used device for connecting to the Internet, followed by computers, wearable devices, and connected toys. Regarding young children, a survey conducted by D. Ribner et al. [32] found that, during the COVID-19 pandemic, children aged 3 to 7 years old have spent an average of 50 minutes more per day in front of screens than they did before the pandemic. Consequently, along with factors such as gender and age, the risk of developing problematic Internet use (PIU) among youth has increased since the onset of the pandemic [33]. While Internet access can be beneficial for minors by providing digital resources for educational purposes, entertainment, and social connection, it also poses significant security and privacy risks. The more time minors spend using the Internet, the more they are exposed to harmful content such as violence, bullying, gambling, and pornography [28]. Based on the aforementioned report [21], despite being considered 'digital natives', i.e., people who are immersed in technology on a daily basis, young people's navigation skills, such as the ability to assess the reliability of information on a web page, vary significantly across European countries, ranging from 36% to 75%. For these reasons, parental control solutions are essential to protect children from harmful content and educate them about safe online behavior. Since protecting minors from online risks is not solely the responsibility of parents or guardians, but also involves governments, legislators have taken steps to address this challenge. In the European Union, the GDPR (General Data Protection Regulation) addresses this issue through its 2018 data protection reform package. It states that children deserve specific protection regarding the use and collection of personal data for marketing purposes and introduces the right to be forgotten, which allows individuals to request the removal of their personal information from the web if desired [31]. Similarly, in the U.S., COPPA (Children's Online Privacy Protection), in effect since 2000, restricts the types of data gathered from children and grants parents more control.

At the national level, Italy introduced more specific measures to protect young people in the online environment, which involve Internet Service Providers (ISPs). An ISP is a company or organization that provides end-users access to the public Internet and different ISPs can cooperate with each other to provide global connectivity. They provide access to the Internet through various means, including broadband services such as fiber or cable, as well as mobile data. The AGCOM (Italian Communications Regulatory Authority),

in January 2024, required all ISPs to implement basic parental control systems that filter inappropriate content for minors and block content reserved for audiences over the age of 18. This system must be provided free of charge to the customers and not in bundle with other additional services for a fee [2]. To the best of our knowledge, all Italian ISPs have implemented measures to comply with this resolution in collaboration with third-party solutions.

1.1 Motivation

The adoption of mobile parental control tools in Europe has decreased on average from 28% in 2010 to 22% in 2020 [17], despite the growing use of online resources and the associated risks to which children are exposed. According to Gnanasekaran et al. [17], this decline is mainly due to parental concerns about these solutions, which often lack robust privacy and security controls. Furthermore, parental control tools typically do not account for age and cultural differences, protect sensitive information, and encourage digital independence in children. Several studies have highlighted issues with parental control in terms of security, privacy, and efficiency, as we will explain in section 2. In this research, driven by the issues highlighted in the literature regarding parental control features and the lack of research specifically on tools provided by Italian ISPs, we aim to assess the effectiveness of the filtering mechanisms offered free of charge by the main Italian ISPs. Specifically, our goal is to understand which filtering mechanism each ISP has implemented and how effective they are. To determine effectiveness and robustness, we are interested in the analysis of the blocked and accessible domains and if there are any techniques that disable or bypass the block.

1.2 Research questions

We propose the following research questions that this work aims to answer:

- RQ1: *Which network techniques do the three targeted ISPs, i.e., Tim, Vodafone and WindTre use for their parental control blocking mechanism?*
- RQ2: *Is there any evidence of underblocking, i.e., domains that should be blocked but are accessible despite the parental control solutions provided by the targeted ISPs?*
 - RQ2.a: *How does underblocking affects the categories identified by the AGCOM?*
- RQ3: *Can the blocking mechanisms of the targeted ISPs' parental control solutions be bypassed?*

To address the **RQ1**, we evaluated the DNS and HTTPS response analysis and compared the results for the same set of domains with parental control enabled versus a device without parental control. Based on these results, we then determined whether a domain was blocked by the filtering mechanism or accessible. To approach the **RQ2**, we first investigated whether each ISP exhibits any underblocking, i.e., if there were accessible domains within the harmful domains dataset. Since the answer was affirmative, we proceeded to analyze the distribution of accessible domains over time by conducting multiple measurements and examining how this distribution varies across different content categories. We tackled the **RQ3** by employing various techniques such as using VPNs, switching to a DNS service different from the provider's, and changing the URL from https:// to http://. The goal was to investigate whether these techniques could bypass the parental control solutions in place for each ISP.

1.3 Contribution

Our research contributes to the academic literature in three ways:

- Providing an overview of how the filtering mechanisms of the studied parental control solutions operate at a network level. Currently, there is limited information available in the literature or from the providers themselves regarding these functionalities, as well as the third parties they collaborate with. Furthermore, one of the requirements of the AGCOM resolution is to offer users the ability to customize and monitor blocking mechanisms. However, we found that this requirement is not consistently met by the solutions under study.
- Evaluating the effectiveness of the filtering mechanisms by reporting the non-negligible percentage of accessible domains, as well as analyzing how underblocking varies across different content categories.
- Assessing the functionalities and security of parental control solutions by demonstrating techniques that can bypass the blocking mechanisms, thereby providing insights for potential improvements to these solutions.

1.4 Structure

The thesis is structured as follows: Chapter 2 provides an overview of the background on parental control functionalities, the measures adopted across Europe to protect minors online, the AGCOM resolution in detail, and the state of the art regarding ISP compliance with the resolution. Additionally, Section 2.5 discusses the related works. Chapter 3 outlines the methodology used to address our research questions, while the results are presented in Chapter 4. A discussion of the results, including limitations and future work, is provided in Chapter 5. Finally, Chapter 6 concludes our work.

Chapter 2

Background

In this section, we will provide an overview of the concept of parental control as discussed in the literature and explore the various types of parental control solutions. Furthermore, we will explore the different measures adopted by the Governments and Institutions across Europe to address the problem of the potential risks that children online are exposed to. Finally, we present an overview of the regulation in Italy, exploring the AGCOM resolution and the measures taken by the ISP in implementing the parental control solutions.

2.1 Parental control

Blancaflor et al.[7] describe parental control solutions as applications or software integrated into the operating systems of various devices such as mobile phones, computers or televisions, designed to restrict access to inappropriate websites or software not suitable for children. In the literature, we can find the distinction between three related terms: parental monitoring, parental mediation, and cyber parental control. Parental monitoring is defined as "a set of actions involving attention to and tracking of the child's whereabouts, activities, and adaptations" [12], which implies supervision and safeguarding of the child. Parental mediation, on the other hand, refers to the interaction between parents and children with regard to media use. This can be either restrictive, where parents control and limit children's online behavior, or evaluative, where constructive discussions about Internet use occur. Finally, cyber parental control encompasses both of these definitions and refers to the control, monitoring, and restriction of children's online activity [19] [6].

2.1.1 Parental control solutions

The primary goal of parental control solutions is to detect harmful content for minors online. In general, their core functionalities include: filtering, usage monitoring, time control and search engines [3] [6].

Filtering approaches There are several approaches to blocking content on the network, such as IP-based, content-based, and DNS filtering. In the first approach, routers filter and drop IP data packets exchanged through the network based on their headers if they are destined for IP addresses on a blocklist. The literature highlights the correlation between the use of IP-based filtering and the risk of overblocking, due to the use of the same IP address by both unobjectionable and harmful websites [29]. A more efficient approach is to filter packets based on their IP/TCP headers. By doing so, not all services on a blocklisted host are blocked, but only the targeted services identified by IP address and port number. The second approach requires not only the

headers to be inspected, but the content of the packets as well. In this case, additional equipment may be needed since routers typically inspect only the packet's header. Finally, in DNS filtering, web pages are filtered based on a comparison between the domain name list to be blocked and the domain name of the requested IP address. Specifically, DNS filtering observes every DNS query that passes through its devices and can return response packets to clients or DNS resolvers [9].

Usage monitoring Monitoring software records the activities of children online so that parents can track and observe them. However, this approach does not block minors from navigating harmful websites, but when used together with a filtering system, it can be an efficient method for parental control.

Time control Similar to usage monitoring, time control is a feature that most parental control solutions provide to limit the time children can spend online. For instance, parents can block access to the Internet at night or during homework time.

Search engines They are designed to be child-friendly and with a safe interface and are either integrated into parental control solutions or provided as features by browsers such as Chrome. They allow access to educational and school websites but could block access to some potentially useful websites or social media. However, this parental control feature can be easily bypassed by children using a different browser.

There exist different types of parental control tools that implement these functionalities in various ways. Network-based devices monitor the encrypted network traffic, usually without inspecting the content. They can control the traffic by using Address Resolution Protocol (ARP) spoofing or by creating a different access point for the children devices. In doing so, they act as a man-in-the-middle between user devices and the router. On the other hand, Android applications for mobile devices and Windows applications for computers employ several techniques for monitoring and filtering. These include device administration features, capturing and retrieving window content, using Android VPNs to filter web content, employing TLS proxies to analyze and monitor HTTPS content. Finally, Chrome extensions for parental control use Chrome APIs to monitor the user-requested URLs, including: intercepting and redirecting traffic, modifying page content, and meta-data such as cookies. [3]

2.2 European Regulatory approaches to protect children online

The measures taken to address the potential risks for children online varie significantly across Europe and can change depending on the country. In October 2011, the four main UK ISPs—TalkTalk, BT, Sky, and Virgin Media— encouraged by the government and the UK Council for Child Internet Safety, agreed to a code of practice concerning adult content filters. In this code of practice, they committed to making parental controls free of charge [21].

TalkTalk ISP launched HomeSafe software in May 2011, the first solution among all UK ISPs to offer router-level blocks. This software allows parents to prevent children from reaching inappropriate websites, control the time they spend online, filter websites mis-used by scammers, and protect devices connected to the home network with a virus alert that blocks infected websites [25].

BT launched its network-level parental control in 2013. The features include blocking

harmful websites with three different filter levels: light, moderate, and strict. The tool allows parents to set up homework time and schedule when filters are active. Additionally, the filter works for devices connected to the home network and also when they are connected to a BT Wi-Fi hotspot [8].

Sky’s network-level parental control, called Sky Broadband Shield, was released in 2013. It protects children’s online activities from phishing and viruses and includes a Safe Search feature to block harmful websites [24].

Virgin Media released their network-level parental control solution, Web Safe, in 2014, and the Child Safe settings have been turned on by default for new customers since 2018. Virgin Media allows parents to customize Web Safe by blocking certain websites and adding timing restrictions [18].

In the European Union, different policies, whether regulatory or self-regulatory, are implemented across Member States concerning parental controls, content rating, and reporting harmful and illegal content. In 2022, the Polish government put forth a draft law requiring internet service providers to offer a way to prevent free access to pornographic content online. However, since the government was charged with censorship and privacy leaks, there is still ongoing debate over the validity of this proposed plan. In France, in January 2022, the French National Assembly’s Committee on Cultural Affairs and Education approved a bill requiring device manufacturers of products that allow Internet access to build in parental control tools free of charge. Similarly, in Spain, in 2024, the Parliament approved a law for the protection of minors online, which requires mobile phones, tablets, computers, and televisions to include an effective and free parental control system. Therefore, French and Spanish laws require built-in parental control solutions in devices free of charge, while only the UK required them mandatory for national ISPs. In other countries, such provisions may not exist. While these measures can have a positive impact, it is evident that not all children across Europe have equal opportunities for empowerment and protection online [20].

2.3 AGCOM resolution

The AGCOM is an entity for regulation and supervision in the area of telecommunications, television, print media and postal services in Italy. It is also responsible for the regulations of the country’s ISPs, by ensuring compliance with the industries standards, consumer protection and fair competition. In November 2023, the Italian government has released a document called ‘Protecting minors in cyberspace, AGCOM resolution’ which obligates ISPs to implement a content filtering solution in order to protect children [2]. Specifically, the resolution requires that:

1. ISPs must provide at least one parental control system based on DNS or another network-level filter, or an app downloadable by the consumer.
2. ISPS must filter a minimum set of categories, as shown in the table 2.1.
3. ISPs must offer to the customers the possibility to customize the categories of websites they want to filter.
4. Operators must notify the Authority of the category used for the filtering system, which must be offered by ISPs free of charge to customers, without costs for activation, deactivation, or configuration.
5. Compliance with the resolution and communication of the technology partner must occur within 9 months from its publication.

6. Reimburse users within 60 days in case of inadequate protection.

Moreover, in September 2024, the AGCOM approved a document outlining regulation regarding the technical methods for age verification [1]. Specifically, it stipulates that access to certain harmful categories, such as pornography and gambling, should only be granted through identification and authentication, placing the responsibility on major online platform providers. These providers must ensure that individuals attempting to access the content are at least 18 years old, using age verification systems supplied by third parties. Additionally, the proposed regulation specifies that age verification providers must not know:

- For which service the age verification is being conducted.
- That two age verification proofs must originate from the same source.
- That a user must have previously utilized the age verification system.

However, at the time of writing, the implementation of the verification mechanisms has not yet taken place.

Name	Description
Adult content	Websites restricted to an adult audience, sites that display full or partial nudity in a sexual or pornographic context, sexual accessories, and sex-oriented activities. Websites that support the online purchase of such goods and services
Gambling & betting	Websites that provide information or promote gambling or support online gambling and/or betting
Weapons	Websites that provide information, promote, or support the sale of weapons and related items
Violence	Websites that promote violence or personal injury, including self-harm, suicide, or that display violence.
Hate and discrimination	Websites that promote or support hatred or intolerance toward any individual or group
Health Damaging Practices	websites that promote or support anorexia and/or bulimia, the use of illegal drugs, alcohol, or tobacco
Anonymizer	Websites that provide tools and methods to make online activity untraceable
Cults	Websites that promote methods, offer instruction, or other resources to influence real-world events through the use of spells, curses, magical powers, or supernatural beings

TABLE 2.1: Categories to be blocked identified by the AGCOM [2]

2.4 Italian ISPs compliance with the AGCOM resolution

In Italy, all ISPs have provided free filtering solutions to comply with the AGCOM resolution. However, each provider offers different solutions for broadband and cellular networks. The leading ISP based on the number of allocated IPv4 addresses is **TIM**. Since November

2023, TIM has provided 'TIM navigazione sicura' free of charge for children. This parental control service allows users to automatically block Internet browsing on sites unsuitable for minors. If a user attempts to access blocked content, the browser will display a block page and the block cannot be bypassed. In cases where redirecting navigation to a blocked page is not technically feasible, the browser itself will show an error message based on the browser type. The service is already active for SIM cards registered in the name of a minor. Additionally, TIM, in collaboration with F-Secure, developed 'TIM Navigazione Sicura APP' to protect smartphones, tablets, and PCs. Rule management can be configured remotely, and protection remains effective even when the user is connected to a Wi-Fi network [26].

The second ISP in the ranking is **WindTre**. It has automatically and free of charge activated the 'Minor Protection Service' on SIMs registered in the name of a minor to block online access to content that is prohibited or unsuitable. A parental control service is available on all broadband services of WindTre customers. To utilize the service, changing DNS configurations is necessary [33].

Regarding the third ISP in the ranking, **Vodafone**, parental control services are already active on offers dedicated to minors, with the filter pre-activated and only able to be deactivated by parents or guardians. Customers using broadband telephony offers can also access parental control for free, which in this case is provided through the F-Secure app [32].

2.5 Related work

In the literature, parental control solutions have been studied from various perspectives, as outlined in the following section. For example, a bibliometric study conducted by Altarturi et al. [6], analyzed research on cyber parental control from 2000 to 2009 and found a 30% increase in publications over the period, reflecting growing attention to the field. However, they underlined that researches have largely focused on the psychological and social aspects of parental control, with less emphasis on the technological features, despite an increasing number of publications addressing these areas.

2.5.1 Parental control vulnerabilities

Parental control solutions, while useful for monitoring children's online activities, have significant functional limitations. Ali et al. [3] conducted static and dynamic analyses of both network-based and software-based tools, revealing vulnerabilities in network devices and software. These included misconfigured, outdated, or unauthenticated server software and API endpoints, as well as improper access control. Of the Android and Windows solutions studied, 18 out of 22 exhibited suspicious activities, such as children creating new accounts or changing passwords without alerting parents. Additionally, 20 out of 22 solutions were vulnerable to brute-force attacks on online passwords, and many exposed Personally Identifiable Information (PII) to third parties.

Blancaflor et al. [7] found that popular parental control apps like FamilyTime, OurPact, and Family Link did not meet OWASP's Mobile Application Security Verification Standard (MASVS), with 10 common vulnerabilities, mostly of moderate to high severity. Another study of 46 Android parental control apps revealed that 11% transmitted PII in plain text, and 34% shared PII without explicit consent, violating privacy policies [15]. Furthermore, many apps requested unnecessary permissions and shared personal data with third parties.

Altarturi and Badrul Anuar [5] explored different methods for filtering content, including URL-based, keyword-based, and content-based approaches. They emphasized that relying on a single filtering method introduces weaknesses, suggesting that a multi-layered approach would be more effective. They also advocated combining textual and visual filtering techniques to improve parental controls. The study highlighted the absence of a publicly available and comprehensive dataset to assess web content filtering, limiting the effectiveness of solution comparisons and evaluations.

Moctezuma et al. [19] focused on the usability of parental control solutions for streaming services, highlighting that, despite users recognizing the importance of parental controls, adoption rates remained low. This was attributed to complex interfaces, varying implementation across platforms, and a lack of understanding of the tools' functionalities, which restricted user experience and security.

2.5.2 Innovative frameworks

In terms of technological advancement, some researchers have proposed new frameworks to improve parental controls. One study by Dhivvya et al.[22], addressed the limitations of traditional parental controls in identifying harmful new applications for children by developing a network traffic classification system using machine learning and deep learning techniques, which proved to be effective in categorizing new apps. Fuertes et al. [16] designed an application that not only blocks inappropriate websites but also tracks and monitors children's Internet activity, providing insights into which categories are most frequently accessed.

2.5.3 Censorship and filtering techniques

Studies specifically focusing on filtering and blocking efficiency have noted significant issues. For instance, a study conducted by Rowe et al.[23], investigating UK ISPs and third-party companies that supports clients' systems and IT infrastructures, known as Managed Service Providers (MSPs), revealed substantial evidence of both overblocking and underblocking in censorship. Overblocking refers to the blocking of websites that should not be blocked, while underblocking allows users to access harmful content that should be restricted. They used the data provided by the Open Rights Group and the blocked.org.uk project that allows to check whether a website is blocked by IPS filters in the UK. The study showed that underblocking rates ranged from 30% to 86% among ISPs, and from 2% to 6% for MSPs. These discrepancies highlight the inconsistency in filtering effectiveness across different providers.

On the other hand, a study conducted by Dalek et al.[10], examined the use of URL filtering methods for surveillance and censorship in countries like the UAE and Yemen. URL filtering, in fact, can potentially be used by authoritarian regimes to block Web content. The study presented a method to identify the installation of URL filtering products and evaluated whether their presence indicated censorship by checking if they accepted user-submitted sites for blocking.

Another study focusing on the censorship and DNS filtering is the study of Cheng et al. [9]. They evaluated the impact of DNS filtering on three different types of DNS resolvers: public, ISP and open, in the censored networks in China. The study demonstrates that many open DNS resolvers are unaffected by the DNS filtering and that in some cases they fail and answer with the correct IP address even for blocked domains. Moreover, they

demonstrate that the effectiveness of the national-level DNS filtering mechanism is influenced by three factors: DNS resolver, client location and blocked domain name. Finally, a paper by Ververis et al. in 2021 [30], investigated the Internet censorship in Europe, specifically the role of the ISP. The results show the use of new and advanced network interference techniques by the ISP to filter Internet content. They found, among other things, that Vodafone ISPs used deep packet inspection system and TLS blocking in Spain during the period 2016-2020.

Therefore, much of the literature focuses on the security and privacy limitations of parental control solutions, with several studies identifying critical vulnerabilities in current tools. Although some propose alternative solutions, few papers examine the efficacy of filtering and blocking mechanisms specifically. Only one study evaluates the filtering effectiveness of UK parental control solutions, while others focus on filtering methods used for censorship. There is a notable gap in research evaluating the functionality of parental control solutions provided by Italian ISPs, especially concerning the efficiency of their filtering mechanisms. This study aims to fill that gap by assessing how effectively Italian ISPs' parental control tools block harmful content and ensure children's online safety.

Chapter 3

Methodology

In this chapter, we present the methodology used to address our research questions. First, we acquired eSIMs and phone plans from the three targeted providers—Tim, WindTre, and Vodafone. These eSIMs were installed on a cellular device to conduct our measurements. For each provider, we enabled the parental control filtering mechanism by following the official activation procedure. Since our eSIMs were registered to an adult account - where parental control is not activated by default - the process required calling the provider’s service number and explicitly requesting activation. The activation itself took up to 24 hours to complete. All measurements were performed on a computer connected to the cellular network via USB tethering to ensure a direct and stable connection. In the following sections, we introduce the datasets used, the domain categorization process, and the measurement procedures. We then outline the implementation of these measurements across different use cases to address our research questions.

3.1 Dataset

For our measurements, we utilized two domain datasets: Tranco [27] and DNS for Family [14].

- **Tranco dataset:** Tranco is a publicly accessible list of the top 1 million most visited websites. It is generated by combining multiple web rankings from Cisco Umbrella, Majestic, and Quantcast, providing a more comprehensive ranking compared to individual sources. Le Pochat, Victor, et al. research [27] has shown minimal overlap between these sources, making Tranco a more exhaustive dataset. For our analysis, we used the Tranco dataset from October 11, 2024.
- **DNS for Family dataset:** DNS for Family is a service that offers free DNS-based content filtering, blocking websites deemed inappropriate for minors and enforcing safe search on major browsers and online content platforms. The service maintains an automatically updated blocklist of restricted domains, which we used as a reference for blocked websites. From this list, we randomly sampled 50,000 domains from the December 16, 2024 snapshot of the dataset for our study.

These datasets provide a balanced mix of popular domains and domains that a parental control should filter, allowing us to effectively evaluate the behavior of parental control mechanisms.

3.1.1 Categorization and filtering

We conducted a categorization of all the domains in the datasets by using Cisco Umbrella Investigate content categorization API [11]. Specifically, we utilized the Get Domain Status and Categorization endpoint, which returns the status, content categories, and security category IDs for each domain. If a domain is malicious, the status returned is -1, if it is benign, the status returned is 1. Finally, if it is unclassified, the status returned is 0. Each domain can belong to up to 11 categories, but, for each domain, we considered only the first 7 categories, as they cover 99.996% of domains.

In the second stage, we filtered the datasets. In particular, we examined the Cisco categories and determined whether they fall under the AGCOM primary categories based on the categories’ names and descriptions; if not, we eliminated the domains from our analysis. The filtering was necessary since, as previously stated, our Tranco dataset covers the top 1 million most visited websites, which include both harmless and inappropriate content for minors. Similarly, while our DNS for family dataset contains domains that are inappropriate for children, it may also include domains that are not included in the resolution’s categories. After filtering, the Tranco dataset contains 62,822 domains, while the DNS for Family blocklist includes 17,081 domains. To facilitate further analysis, we grouped the Cisco Umbrella Investigate categories into broader AGCOM-defined categories based on Cisco’s category descriptions. This categorization is summarized in Table 3.1

AGCOM harmful categories	Cisco categories
Adult content	Adult, Adult Themes, Dating, Nudity, Pornography, Sexuality, Tasteless
Gambling & betting	Gambling
Weapons	Weapons
Violence	Tasteless, Terrorism, Child Abuse Content, Extreme, Terrorism and Violent Extremism
Hate and discrimination	Hate Speech, Hate/Discrimination
Health Damaging Practices	Alcohol, Drugs, Tobacco, Cannabis, Illegal Drugs
Anonymizer	Filter Avoidance, Illegal Activities, Illegal Downloads, Hacking, Proxy/Anonymizer
Sects	Paranormal

TABLE 3.1: Mapping of the Cisco harmful content categories into the main categories identified by the AGCOM resolution

3.2 Setup

In this section, we describe the setup configuration used to conduct our measurements. A MacBook Pro was used as the primary testing device, while a mobile device was equipped with eSIMs from the targeted ISPs (Tim, Vodafone, and WindTre). Each provider’s official application was installed on the mobile device to monitor Internet traffic usage. This monitoring was necessary because the tests were conducted in the Netherlands, and due to provider policies, despite having approximately 200 GB of data available per month, roaming restrictions limited usage to only 10 GB per month. To ensure that all tests were conducted over the targeted ISP’s data network, the MacBook Pro accessed the internet exclusively through the mobile phone’s cellular data connection. The connection between the computer and the mobile device was established via a USB cable to provide a more

stable and fast connection.

3.2.1 Internet traffic measurements

The objective of this research was to analyze how parental control solutions implement website blocking mechanisms and to verify whether access to harmful websites is effectively restricted. To achieve this, we developed a script that evaluates domain accessibility using HTTPS and DNS queries. Specifically, the script performs two main checks:

HTTPS check It sends HTTPS GET requests to the domains in our dataset using the Requests Python library to retrieve the server's response status code.

DNS query It retrieves DNS resolution information, including IP addresses, status codes, response times, TTL (Time-to-Live), and message size, using the pydig Python library.

By performing DNS queries, we checked whether a domain name could be resolved to an IP address, and through HTTPS requests, we verified whether the resolved domain could be accessed via a secure connection. The Python script allowed us to conduct the tests in a scalable way for all the domains in the datasets. The results were saved in a CSV file in the following format **Domain | HTTPS status | DNS return code | IP address | size | TTL | response time | exception**

To enhance efficiency, the input data was processed in batches using multithreading. Concurrency was handled with Python's ThreadPoolExecutor, running 150 concurrent threads to speed up execution. Based on the measurement results, we could analyze how website blocking is enforced. Specifically, the possible outcomes of the HTTPS status code response are:

- 1xx** Informational responses, these indicate that the request is still being processed. They may provide information about protocol switching or indicate that the final response is not yet available.
- 2xx** These confirm that the request was successfully processed. In particular, 200 OK indicates that the request was delivered successfully and the content is fully accessible. Other 2xx codes may indicate that the request was accepted but only partial content was retrieved or that no content is available to return.
- 3xx** Redirection responses, these indicate that the requested content has been moved to a different URL, requiring redirection.
- 4xx** Client error, these signify that the request was incorrect or unauthorized, typically due to client-side issues.
- 5xx** Server error, these indicate that the request failed due to a problem on the server side.

Additionally, the GET request for a given domain may raise an Error or Exception. These can occur at the client-side, server-side, or during communication between client and server, often at the transport layer (e.g., TCP/IP, SSL/TLS).

Regarding DNS queries, the outcome depends on the status code returned:

Status 0 : The query was successful, and the response contains an IP addresses, response size, TTL (Time-To-Live), and response time, with no exceptions.

Status 1 : The query was invalid or incorrectly formatted, leading to an exception. In this case, the IP address and TTL are missing, while the response size and time are minimal.

Status 2 : The server was unable to retrieve the requested domain information. This status could raise an exception and no IP addresses and TTL are returned.

Status 3 : The queried domain does not exist (NXDOMAIN). In some cases, the TTL for such a response may be set to a negative value.

3.2.2 Output analysis

We analyzed the measurement outputs to determine whether blocking is DNS-based or it occurs at another network level. Based on our preliminary findings, we established the following baseline for interpreting filtering behavior:

- If a domain is blocked at the DNS level, we expected to observe either a DNS status code different from 0, indicating a failure in resolving the domain, or a status 0 with an IP address that redirects to a blocking page instead of the actual destination. Additionally, the TTL (Time-To-Live) value for blocked domains is often significantly lower compared to allowed domains, ensuring frequent lookups and preventing users from caching a bypassed DNS resolution.
- On the other hand, if blocking occurs beyond the DNS resolution stage, we inferred it from the HTTPS response. In such cases, we expected to see client-side errors such as 403 Forbidden, indicating restricted access, redirection responses (3xx) that lead to an alternative page instead of the requested content, or timeouts and connection failures, which may suggest filtering mechanisms at the ISP or firewall level.

With the examination of both DNS responses and HTTPS status codes, we gained insight into the mechanisms used to restrict access to certain websites.

3.2.3 Wireshark

Parallel to our measurements, we captured the traffic between the computer and the access point using the Wireshark tool, an open-source network protocol analyzer that enables real-time network traffic capture through the pcap library [34]. Additionally, it provides various analysis functions such as stream reassembly, decoding, and protocol statistics. This allowed us to examine not only the outcomes of HTTPS requests and DNS queries but also the entire communication between our client and the server. For a deeper packet inspection, we applied packet filters based on packet header fields, which enabled us to focus on packets directed to specific destinations or using particular port numbers. Another useful filtering method in our analysis was selecting packets based on DNS or TLS streams, allowing us to observe the entire communication flow for a specific destination. One limitation of our HTTPS measurements was that we only recorded the HTTPS response for a GET request. However, by using Wireshark, we could inspect the full TCP/TLS communication, including the TLS handshake and the exchanged packet contents, providing a more comprehensive view of the interactions between the client and the server.

3.3 Use cases

In this section, we present the comprehensive tests conducted in our analysis. Each test involved running the script described in Section 3.2.1 on the domains in our datasets while

using the network data of the ISP under analysis at that moment. The test outputs were saved as CSV files for further examination. Additionally, during each test run, network traffic was captured using Wireshark and stored in pcap files. To address our research questions, we conducted different tests with varying setups, tailored to the specific aspects we aimed to investigate. In the following sections, we provide a detailed explanation of the approach used to answer each research question.

3.3.1 Research question 1

The objective of the first research question is to investigate the filtering mechanisms of parental control solutions provided by the main ISPs in Italy: Tim, Vodafone and WindTre. As discussed in Section 2.4, each ISP has implemented its own parental control solution; however, there is no publicly available information regarding the technical aspects of website filtering. This leads us to our research question: *Which network techniques do the three targeted ISPs, i.e., Tim, WindTre and Vodafone use for their parental control blocking mechanism?*

General approach

Our approach is described in figure 3.1 and involves conducting measurements for each ISP with parental control enabled, then repeating the same measurements with parental control disabled, and comparing the results. The rationale behind this reverse-engineering approach is that if a filtering mechanism is applied, the behavior of the network should change accordingly. By comparing the two sets of results, we aim to eliminate ISP-specific biases caused by network policies or traffic management mechanisms that might otherwise be mistakenly attributed to parental control filtering. Differences in the number of accessible domains or variations in DNS and HTTPS outputs could indicate the presence of a filtering mechanism, leading to further investigation. Additionally, comparing the measurements across different ISPs with parental control enabled may provide further insights into their filtering techniques. To ensure accuracy, consistency, and a fair comparison, all measurements were conducted on the same day, December 16, 2024, minimizing the impact of network fluctuations or variations in filtering policies. We used a refined version of our Tranco dataset containing 1,000 domains and performed measurements on the networks of Tim, WindTre, and Vodafone using mobile data connections from eSIMs with parental control enabled. The same tests were then repeated on three other devices—one for each ISP—where parental control was disabled. Simultaneously, we captured network traffic using Wireshark. For analysis, we first examined DNS responses to identify variations in status codes. If a response had a status code of 0, we analyzed the resolved IP addresses to check for localhost IPs, non-routable addresses (e.g., 0.0.0.0), or redirections to blocking pages. Subsequently, we compared HTTPS responses obtained with and without parental control enabled to identify potential discrepancies. By the end of our analysis, we were able to describe the technical aspects of the filtering solutions employed by each ISP and classify domains based on whether they are blocked by parental control or remain accessible.

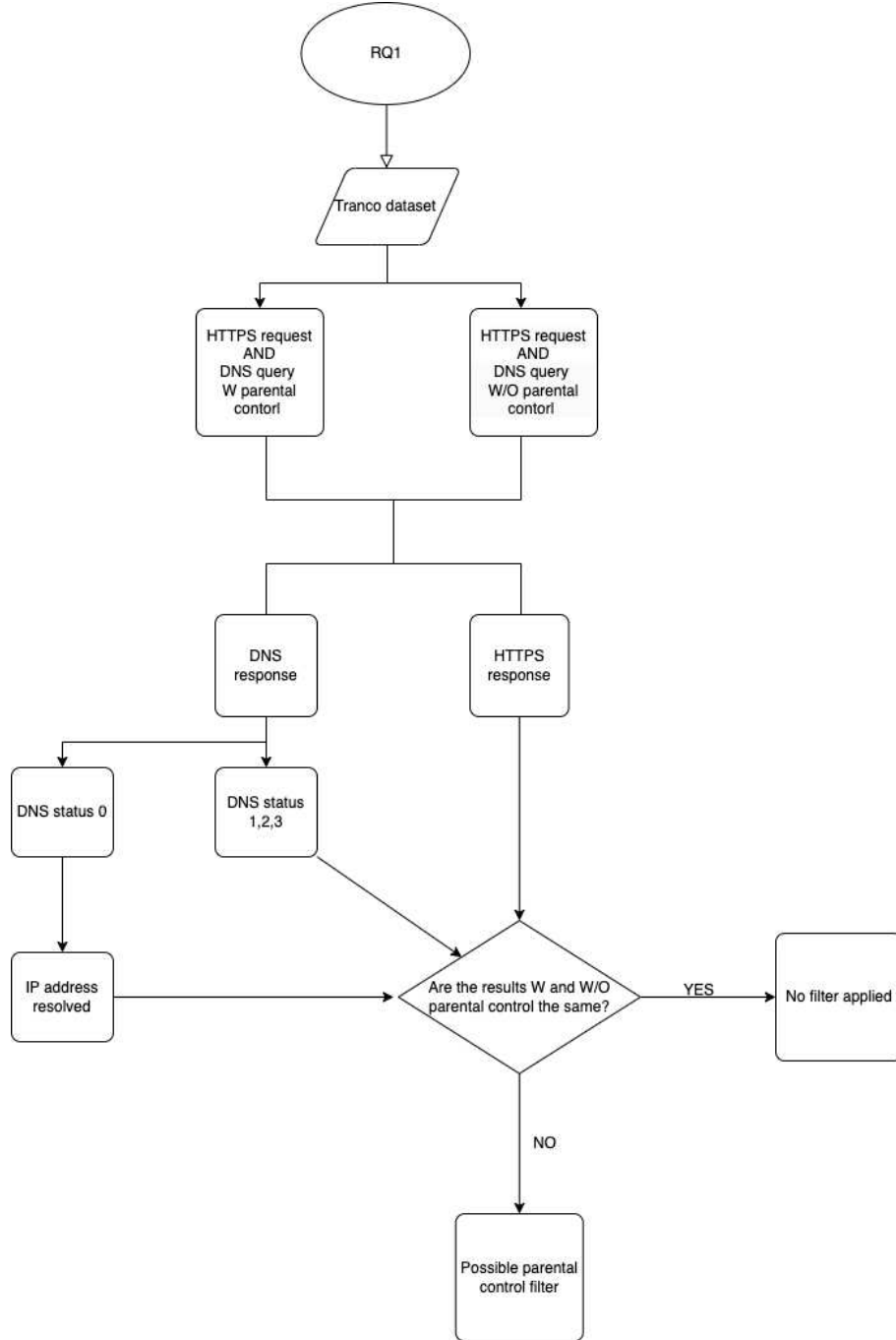


FIGURE 3.1: Flowchart describing the methodology used to address the research question 1

3.3.2 Research question 2

Our primary objective was to investigate the efficiency of parental control mechanisms. Given the filtering issues highlighted in the literature, we seek to answer the following question: *Is there any evidence of underblocking, i.e., domains that should be blocked but are accessible instead for the parental control solutions provided by the targeted ISPs?* In the following section, we outline our approach to addressing this question, as described in 3.2.

General Approach

We considered a domain accessible if the DNS query returned a status code of 0 and a valid IP address. An IP address is deemed valid when it does not redirect to a blocking page and corresponds to an actual content page. Additionally, we verified accessibility through the HTTPS status code, considering a domain accessible if it returns a 200 OK response. A domain was classified as accessible only when all these conditions were met. We conducted the tests on both our datasets obtained from Tranco and DNS for Family, covering a total of slightly over 70,000 domains classified as harmful. To ensure the validity of our data and monitor potential trends in the blocking mechanisms, we repeated the measurements for the domains in both datasets five times at intervals of one to two weeks, spanning from December 2024 to January 2025. Furthermore, to allow for a fair comparison, the tests for all three ISPs were conducted on the same day. Our analysis focused on identifying accessible domains, tracking variations over time, and examining how underblocking differs among providers.

Research question 2.a

In the event that underblocking was observed, we were also interested in analyzing how it varies across different domain categories. As explained in Section 3.1, we categorized the domains in our datasets using Cisco Umbrella’s content classification. Subsequently, for our analysis, these Cisco categories were mapped to the main AGCOM categories, allowing for a broader and more structured evaluation 3.1.

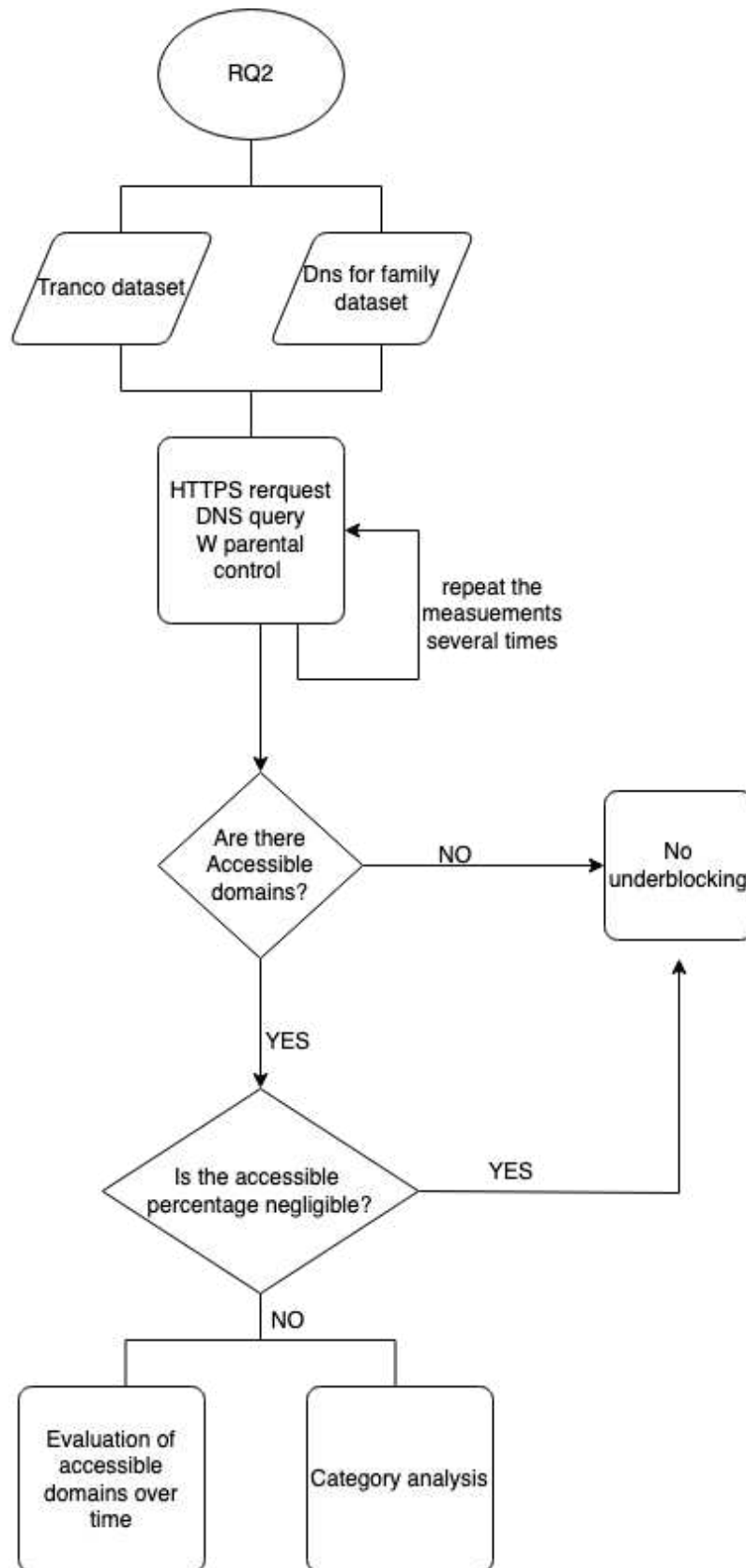


FIGURE 3.2: Flowchart describing the methodology used to address the research question 2

3.3.3 Research question 3

Another crucial aspect in evaluating the effectiveness of parental controls is assessing their robustness against circumvention attempts. This leads us to our final research question: *Can the blocking mechanisms of the targeted ISP’s parental control solutions be bypassed?*. To address this, we tested various bypassing techniques, including VPNs, direct HTTP requests, and Google DNS. In this section, we outline the methodology used to evaluate these techniques 3.3. For the following tests, we selected a random sample of 10,000 domains from our Tranco dataset.

VPN

A Virtual Private Network (VPN) is a service that enhances privacy and security by redirecting internet traffic through a private server. When connected to a VPN, the traffic is encrypted, and a secure tunnel is established between the device and the VPN server. The VPN also provides IP masking by assigning a new IP address, which can be in a different region or country, allowing users to bypass geographical restrictions. Since the encrypted channel routes DNS queries through the VPN’s DNS server, the traffic should, in theory, be untrackable by ISPs, making it a potential method to bypass ISP-level filtering mechanisms. Our research focused on evaluating whether the parental control filters implemented by ISPs can be bypassed using VPNs. To conduct this analysis, we used both a paid commercial VPN, ExpressVPN [13], and the educational VPN provided by the University of Twente. We compared the results of our HTTPS and DNS analyses with and without VPNs for each ISP, ensuring that parental control remained enabled throughout the tests.

HTTP requests

For all our measurements, we sent GET requests using HTTPS. Hypertext Transfer Protocol Secure (HTTPS) is a protocol that encrypts communication and data transfer between a web client and a web server. It is the secure version of the HTTP protocol, as it employs Transport Layer Security (TLS) or Secure Sockets Layer (SSL) encryption. Before the client and server exchange data, they perform a handshake to verify each other’s identity through certificate exchange. Since filtering mechanisms can operate at the TLS/SSL level by invalidating server or client verification and blocking the connection, our aim was to investigate what is the parental control behavior when TLS over HTTP is not enforced. To do so, we compared the results of the measurements by sending HTTP and HTTPS requests. Additionally, we were interested in analyzing the behavior of domains, specifically, how many still support HTTP and how many automatically redirect to HTTPS.

Google DNS

We conducted our tests using the default DNS resolver provided by the ISP, meaning that all DNS queries were directed to the ISP’s DNS servers. If the blocking mechanism operates at the DNS level, we were particularly interested in examining whether changing the DNS resolver affects the outcome. To investigate this, we compared the results of tests conducted using the provider’s DNS with those performed using Google’s public DNS at the address 8.8.8.8.

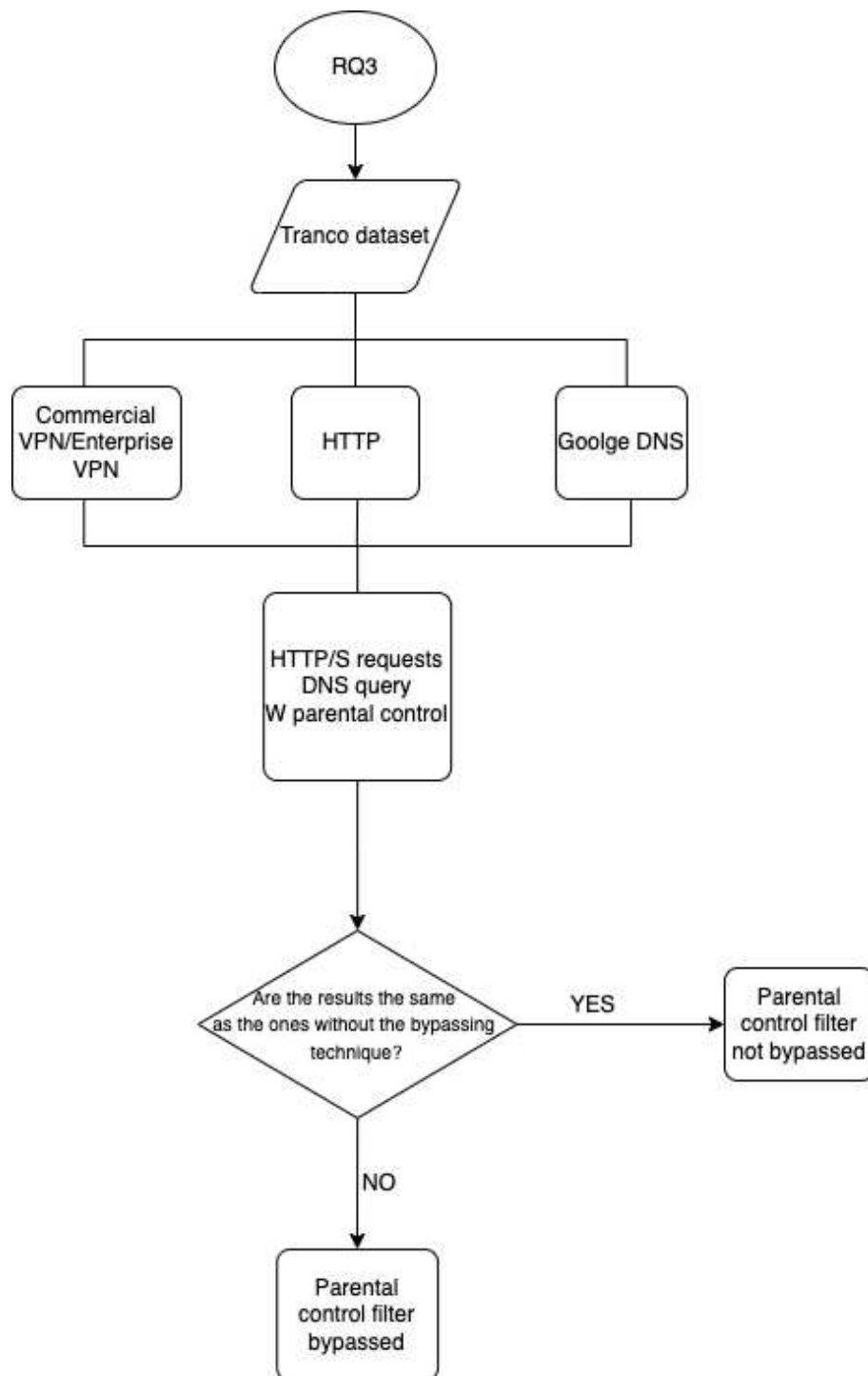


FIGURE 3.3: Flowchart describing the methodology used to address the research question 3

Chapter 4

Results

In the previous chapter, we explained how we planned to answer the research questions. In this chapter, we dive into the implementation of the methodology and present the results obtained. The chapter is divided into subsections based on the research questions addressed. For **RQ1**, we provide an explanation of the technical aspects of the blocking mechanism for each provider, specifically identifying the network level at which it occurs. Based on these results, we can distinguish when a domain is blocked by the parental control filter and when it remains accessible. Additionally, as required by the AGCOM resolution, we include a review of the user experience and customization features of the various solutions. Specifically, the Article 10 of the resolution states that the contract holder should have the possibility to add or customize the contents subject to filtering. Therefore, we verified whether this article is complied. The results from the first research question enable us to further explore **RQ2**, where we identify the presence of underblocking for each provider, assess whether it is a negligible issue, and examine how underblocking varies across domain categories. Finally, we present the results for **RQ3**, addressing the effectiveness of each bypassing technique and evaluating to what extent the block is maintained for each ISP.

4.1 Network-Level Blocking Techniques

In this section, we explain the technical aspects of the filtering mechanisms employed by the parental control solutions of each studied provider, based on the measurement results and analysis we conducted. Additionally, we provide an overview of the user experience and customization options offered by the ISPs for their solutions. The following section presents the results for Tim, WindTre, and Vodafone.

4.1.1 Technical filtering mechanism

In the following sections, for each ISP we provide a detailed explanation of the filtering mechanism implemented based on the DNS and HTTPS responses. Flow diagrams illustrating the blocking process for each ISP are presented in figures [4.1](#) and [4.2](#).

TIM

DNS response analysis The first aspect we examined in our analysis was the DNS response status for measurements taken with parental control enabled. We found that 98% of the responses returned a status of 0, indicating that the queries were successfully processed. Next, we analyzed the IP addresses returned and observed that 34% of the domains in the blocklist tested reported the IP address 85.38.30.24, as it can be clearly seen

in figure 4.3. This IP address belongs to the range 85.38.0.0/16 and is owned by Telecom Italia. It matches the IP of the blocking page to which we are redirected when attempting to access a blocked website using the TIM parental control. Furthermore, the average TTL for DNS responses that returned the targeted IP address was 22 seconds, which is significantly shorter than the TTL for other TIM DNS responses, that is on average 353s, and those from other ISPs which ranges from 5000s to 8000s. Therefore, whenever we attempt to access a website on the blocklist, the ISP's DNS does not resolve the IP address of the requested content but instead redirects us to the blocking page at IP address 85.38.30.24.

HTTPS response analysis The next step in our investigation was to analyze the HTTPS responses. The most common HTTPS response we encountered, reported 53% of the time, was the SSL error: *SSLCertVerificationError("hostname doesn't match either of '*.safeweb-service.tim.it', '*.bb.safeweb-service.tim.it', '*.dr.safeweb-service.tim.it'")*. This error occurs during the TLS handshake over TCP. Specifically, after the TCP connection is established between the client and the server, the client initiates the TLS handshake to establish a secure connection. However, when trying to access a blocked domain, the DNS resolution provided an IP address different from the actual content server's IP. As a result, during the TLS handshake, the client sends a Client Hello message that includes the SNI (Server Name Indication) extension, which contains the hostname of the requested domain. At this point, the server at IP address 85.38.30.24 (rather than the actual domain) responds with a certificate that does not match the domain the client is attempting to visit. This causes an SSL certificate mismatch error, leading to the failure of the connection. As a further validation, all domains that returned the blocking IP address during DNS resolution reported this SSL error during the HTTPS handshake.

Validation findings By comparing the results obtained with parental control enabled to those with parental control disabled, we observed that no SSL errors occurred without parental control. Moreover, in the test without parental control, no IP addresses in the range 85.38.0.0/16 were present. Based on this analysis, we conclude that the TIM ISP parental control works via a DNS blocking mechanism. This mechanism redirects the user to a blocking page using a fake IP address for the domains on the blocklist. The redirection causes the HTTPS connection to fail at the TLS handshake level, as shown in Figure 4.1.

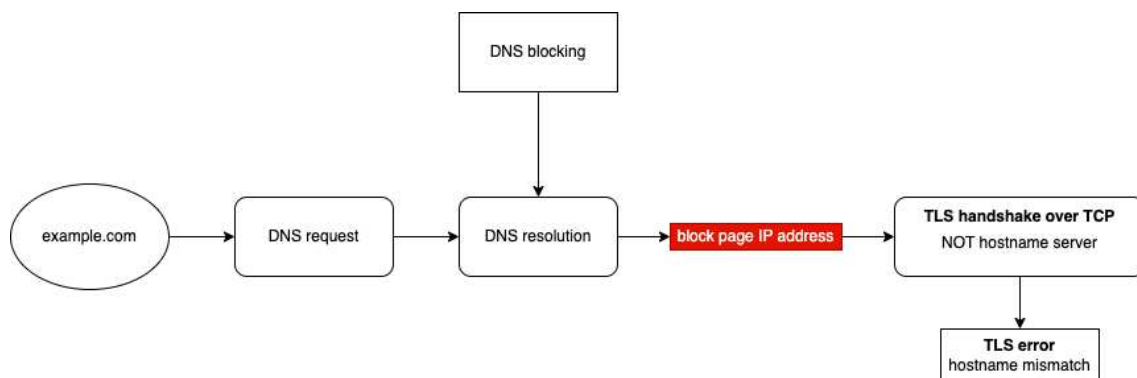


FIGURE 4.1: Flowcharts describing the DNS blocking process implemented by TIM

Vodafone

DNS response analysis For the DNS responses in the Vodafone ISP tests, we observed that, similar to the previous provider, 98.6% of the domains returned a status code of 0. This led us to rule out the possibility of a DNS block that simply fails to process the queries. The next step was to examine the IP addresses resolved for the domains that returned DNS status 0. Specifically, we focused on identifying any IP addresses that were commonly resolved for multiple domains, as such addresses could indicate a redirection to a blocking page. We also checked for the presence of localhost IP addresses or non-routable addresses that could be indicative of filtering. Our analysis revealed no evidence of such IP addresses. The most frequently resolved IPs were 188.114.96.7 and 188.114.97.7, each accounting for 5.8% of the domains. The next most common IP addresses were 104.21.80.1, 104.21.64.1, 104.21.112.1, 104.21.32.1, 104.21.48.1, 104.21.16.1, and 104.21.96.1, each making up 3.1%. All other IP addresses were resolved fewer than 10 times, making it unlikely they were used for blocking. These IPs are owned by Cloudflare, which hosts multiple websites on them, and they do not appear to be associated with blocking services.

HTTPS response analysis Subsequently, we analyzed the HTTPS responses. The most common response returned was an SSL error, followed by status 200 OK, as shown in Figure 4.6.

The SSL error is *SSLCertVerificationError(1, SSL:CERTIFICATE_VERIFY_FAILED certificate verify failed: self signed certificate in certificate chain)*. This error indicates a failure in certificate verification due to a self-signed certificate in the certificate chain exchange between the client and server during the TLS handshake.

Through deeper packet inspection using Wireshark, we observed that all domains reporting this error had certificates signed by Allot [4], an international cybersecurity solutions provider that collaborates with Vodafone. This suggests the presence of TLS traffic redirection and interception by a man-in-the-middle (MiTM) entity, which injects a manipulated certificate. Since the certificate sent from the server to our client is not signed by a trusted certificate authority (CA), the TLS handshake is interrupted, and the connection is closed with an "Unknown CA" alert. A detailed explanation of the blocking process is shown in Figure 4.2.

We found similar results to those obtained by Vesveris et al. [30] that investigated Vodafone ISP for censorship. This study highlighted instances of TLS interception used by Vodafone in Spain, where similar SSL errors were observed during measurements.

Validation findings To further validate our findings, we compared the resolved IP addresses between the measurements with and without parental control. We conducted a weighted analysis, where the weight was based on the number of domains resolved to each IP address. The results showed that 97% of the IP addresses were consistent between both tests. Given that most of the domains returned DNS status 0 and no blocking IP addresses were identified, we concluded that Vodafone's parental control mechanism does not operate at the DNS level. Further, we conducted an analysis to exclude the possibility that other HTTPS responses, such as 403 errors, connection errors, and timeouts, were caused by the parental control filter. By comparing the "with" and "without" results, we observed similar behavior for these HTTPS responses. Notably, when we increased the timeout from 3 seconds to 15 seconds, 30% of the timeout errors were resolved, suggesting no correlation with the parental control filter. Finally, While the SSL error is the most

frequent response when parental control is enabled, it does not appear in the test without parental control, reinforcing our theory.

WindTre

DNS response analysis Based on our results, WindTre ISP has implemented a blocking system with behavior similar to Vodafone's. As discussed earlier, we began our analysis by investigating the DNS response. We observed that 98% of the domains returned a status code of 0. We then proceeded to investigate the IP addresses resolved for the domains that returned a DNS status of 0. Similar to Vodafone, our analysis revealed no evidence of specific IP addresses associated with the blocks. The next most common IP addresses were 104.21.80.1, 104.21.64.1, 104.21.112.1, 104.21.32.1, 104.21.48.1, 104.21.16.1, and 104.21.96.1, each accounting for 3.2%, followed by 188.114.96.12 and 188.114.97.12, each with a 0.5% occurrence. As shown in Figure 4.3, the first group of IP addresses is common across all the providers, although with varying distribution, while the last two appear only for WindTre. However, the low percentage of occurrence and the fact that these IP addresses also appear in the test without parental control suggest they are unlikely to be redirecting IP addresses.

HTTPS response analysis We then analyzed the HTTPS responses. For WindTre, the most common response returned was an SSL error, followed by a status 200 OK, as shown in Figure 4.6. However, although the error indicates a certificate validation failure, it is slightly different from the one observed in the previous case. The SSL error is *SSLCertVerificationError(1, SSL:CERTIFICATE_VERIFY_FAILED certificate verify failed: unable to get local issuer certificate)*. This error suggests a failure in certificate verification due to an incomplete certificate chain sent by the web server to the client. In this case, the certificate is signed by WindTre's security enforcement. As with Vodafone, the HTTPS connection is closed due to an unknown CA alert sent to the server. Therefore, we can conclude that, like Vodafone, WindTre's parental control intercepts the traffic at the TLS level, and the block occurs by injecting a manipulated certificate.

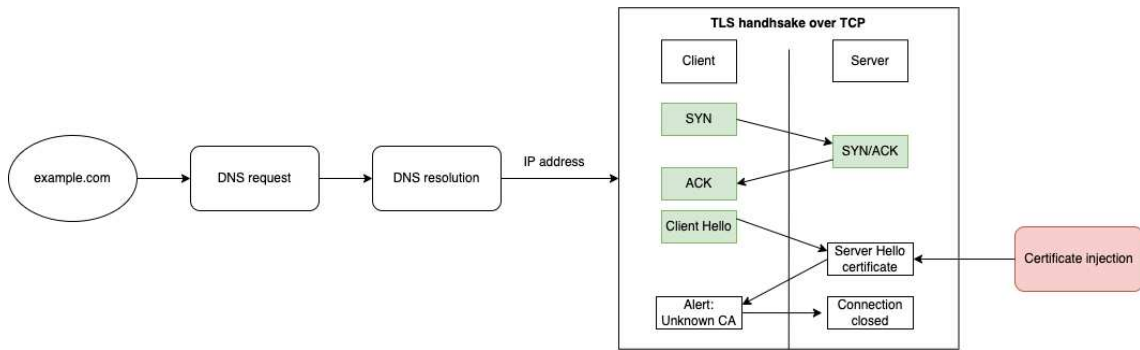


FIGURE 4.2: Flowcharts describing the TLS interception and certificate injection implemented by Vodafone and WindTre

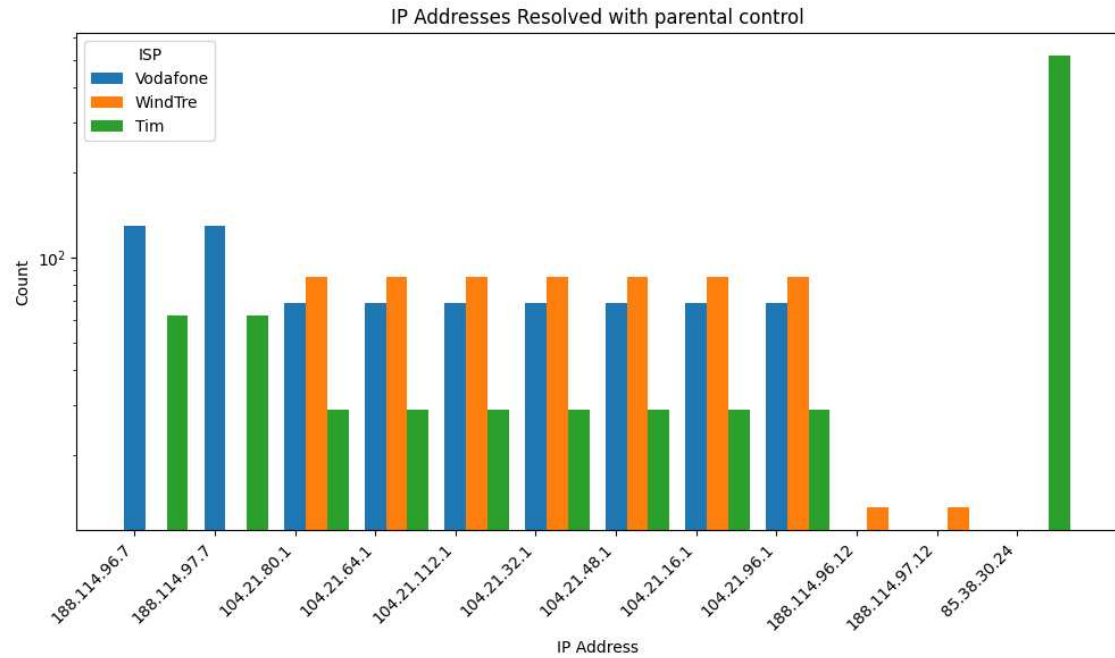


FIGURE 4.3: Comparison of IP addresses returned in our measurements for all the ISP analyzed

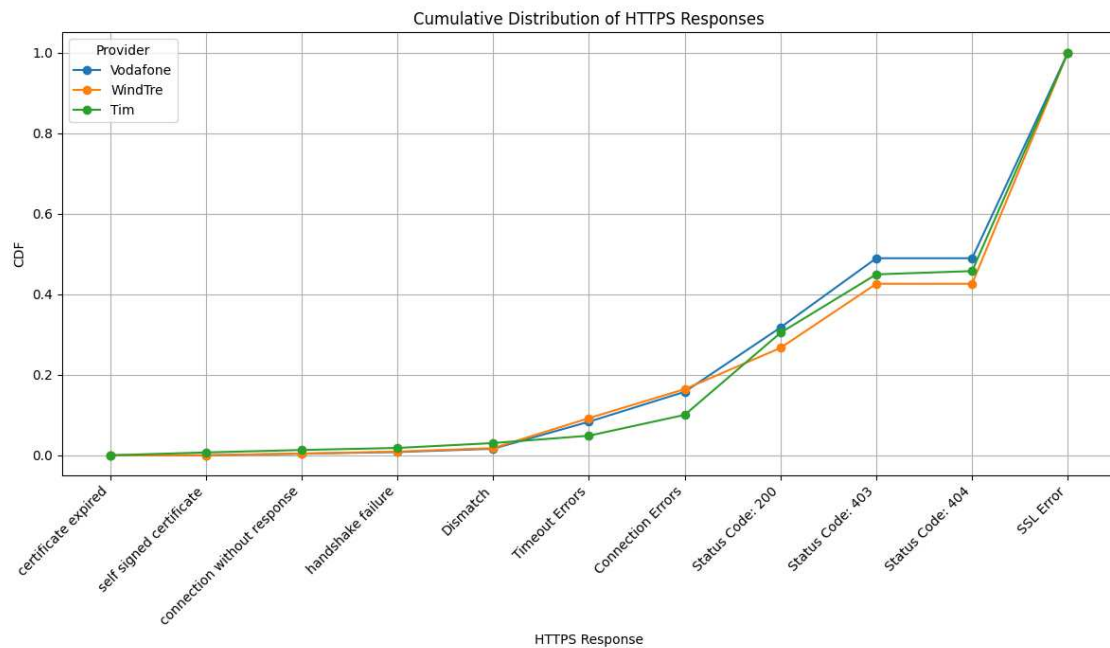


FIGURE 4.4: Cumulative distribution functions (CDF) of the HTTPS responses returned by each ISP in our measurements

4.1.2 User experience and customization

One of the AGCOM resolution requirements for ISPs is to provide a system that allows users to customize the filter and monitor minors' online activity. However, not all the ISPs investigated have implemented such a system, and the level of service varies significantly across providers. To evaluate these features, we manually accessed blocked websites using the Chrome browser to examine the blocking pages returned by each ISP. Additionally, we reviewed the ISPs' official websites and mobile applications to gather relevant information. Regarding the pages displayed for blocked websites, we observed that the content of the blocking page depends on both the browser used and the specific characteristics of the website being accessed. In general, two types of blocking pages were encountered, as illustrated in Figure 4.5.

- Direct redirection to the ISP blocking page where the explanation of why the websites is blocked is given
- Warning page of the browser due to an unsure connection with the server

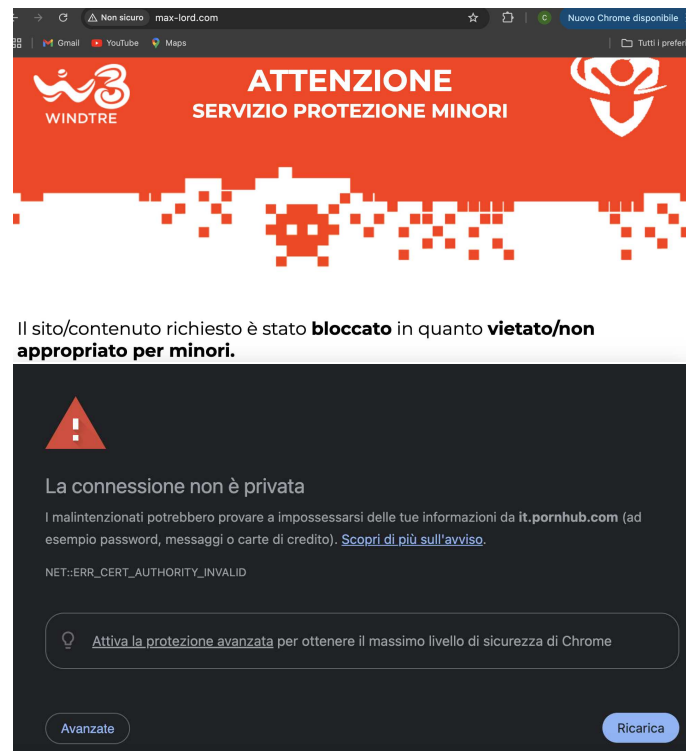


FIGURE 4.5: Examples of blocking pages returned by the ISPs

In the case of Vodafone, when manually accessing a blocked website, users are redirected to a portal that allows customization of filtering categories. Specifically, users can remove or add categories to the blocklist, selecting from a total of 16 predefined categories. Additionally, the portal provides the option to customize a minor's profile by specifying their age range and monitoring their online activity. However, no such customization options

are available through the Vodafone mobile app or website portal. For the other providers, to the best of our knowledge, there is no option to customize filtering preferences, either through a portal or via the app. Although the customer profile menu includes a parental control section, it only provides information about the filtering mechanism and instructions on how to enable or disable it, without offering any customization features.

4.2 Effectiveness of Parental Control Filtering

With the second research question, we aimed to investigate the presence of underblocking—cases where harmful domains that should be blocked remain accessible—for each provider. Based on our previous analysis, we determined when a domain is blocked and when it is accessible for each ISP. A domain is considered accessible if the DNS response returns a status code of 0 with a valid IP address and the HTTPS response returns a status code of 200 OK. Among the ISPs studied, we found that Vodafone and WindTre do not use blocking IP addresses, while TIM implements DNS blocking by redirecting requests to an IP address within the range 85.38.0.0/16. Consequently, in our results analysis for Vodafone and WindTre, we treated all returned IP addresses as valid, as they correctly resolve to the actual content of the requested domains. For TIM, we considered all IP addresses valid except those within the 85.38.0.0/16 range, which indicate redirection to a blocking page.

4.2.1 Underblocking

Our results indicate that underblocking is not only present in the studied ISPs but also occurs at a significant rate. We conducted five different measurements for the domains in the two analyzed datasets, and the results are shown in Figure 4.6. As observed, Vodafone and WindTre exhibit a similar trend, though Vodafone allows access to a higher number of domains. TIM, on the other hand, follows a different pattern due to its parental control system temporarily stopped operating during our testing period. Although we were able to recover it for the final measurement, its operation was short-lived, functioning only intermittently at the beginning of our tests. This explains the spike in accessible domains in the last Tranco measurement for TIM, where the filtering system worked correctly only the first 10000 domains in our tests.

Additionally, from the third measurement onward, we observed an increase in the number of accessible domains for Vodafone and WindTre. This change was due to an adjustment in our request timeout setting, which we extended from 3 seconds to 15 seconds. We made this decision after noticing a significant number of HTTPS responses returning timeout errors and aimed to determine whether these errors were related to the parental control system. As a result, 30% of the previously timed-out domains successfully resolved in subsequent tests, with some proving to be fully accessible.

In Figure 4.7, we present the combined results for the last measurement taken on February 4, 2025, for Vodafone and WindTre, and on January 5, 2024, for TIM (due to its parental control malfunction). The results reveal a non-negligible proportion of accessible domains across all studied ISPs, with percentages ranging from 20.5% to 34.2%. These values are notably close to the percentage of blocked domains, reinforcing the concern that current blocking solutions are not effectively protecting minors from harmful content. Lastly, the category "Others" represents domains that were correctly processed by the ISP's DNS but encountered errors or exceptions at the HTTPS response stage.

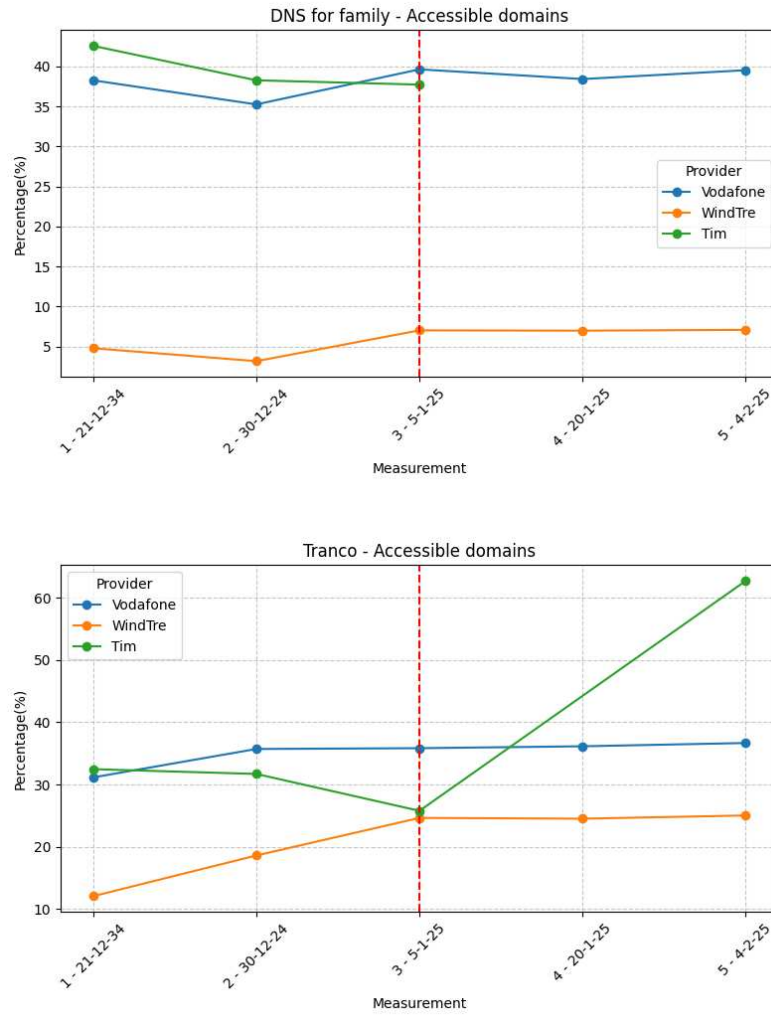


FIGURE 4.6: Trend of accessible domains for the two datasets investigated

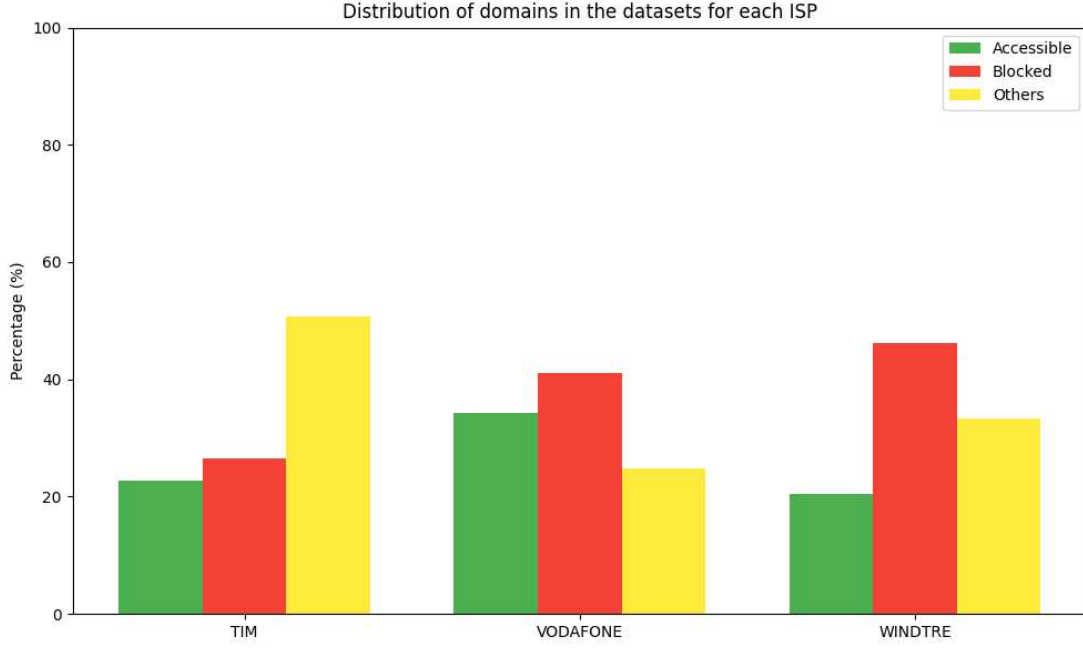


FIGURE 4.7: Distribution of accessible, blocked and unclassified domains for each ISP

4.2.2 Category analysis

After confirming the presence of underblocking, we extended our investigation to assess how filtering efficiency varies across the website categories identified by the AGCOM resolution as harmful and subject to blocking. We refer to Table 3.1, where all Cisco categories used for our dataset classification are grouped into broader categories. The results of this analysis are summarized in Figures 4.8, presenting findings for all ISPs based on the dataset. While our DNS for Family dataset primarily consists of domains classified under adult content followed by gambling and anonymizer categories, the Tranco domains are more broadly distributed across different categories. However, domains categorized as adult content remain the most prevalent in both datasets. Filtering performance varies significantly among ISPs and among datasets. WindTre demonstrates consistent filtering behavior for the adult content categories, showing the lowest percentage of accessible domains across both datasets. However, in the case of gambling domains, while it performs better than the other ISPs in the DNS for Family dataset, with 12% of domains remaining accessible, the percentage of accessible gambling domains increases significantly to 36.7% in the Tranco dataset. Regarding TIM, it shows weaker filtering behavior in the DNS for Family dataset, especially for pornography and nudity domains. On the other hand, Vodafone displays the highest accessibility rates for almost every category in the Tranco dataset, with 41.1% of gambling domains and 61.9% of dating content remaining accessible. Overall, the results reveal inconsistencies both among ISPs and between datasets, suggesting that each ISP has implemented its own filtering policy and blocklist, leading to differences in coverage and effectiveness.

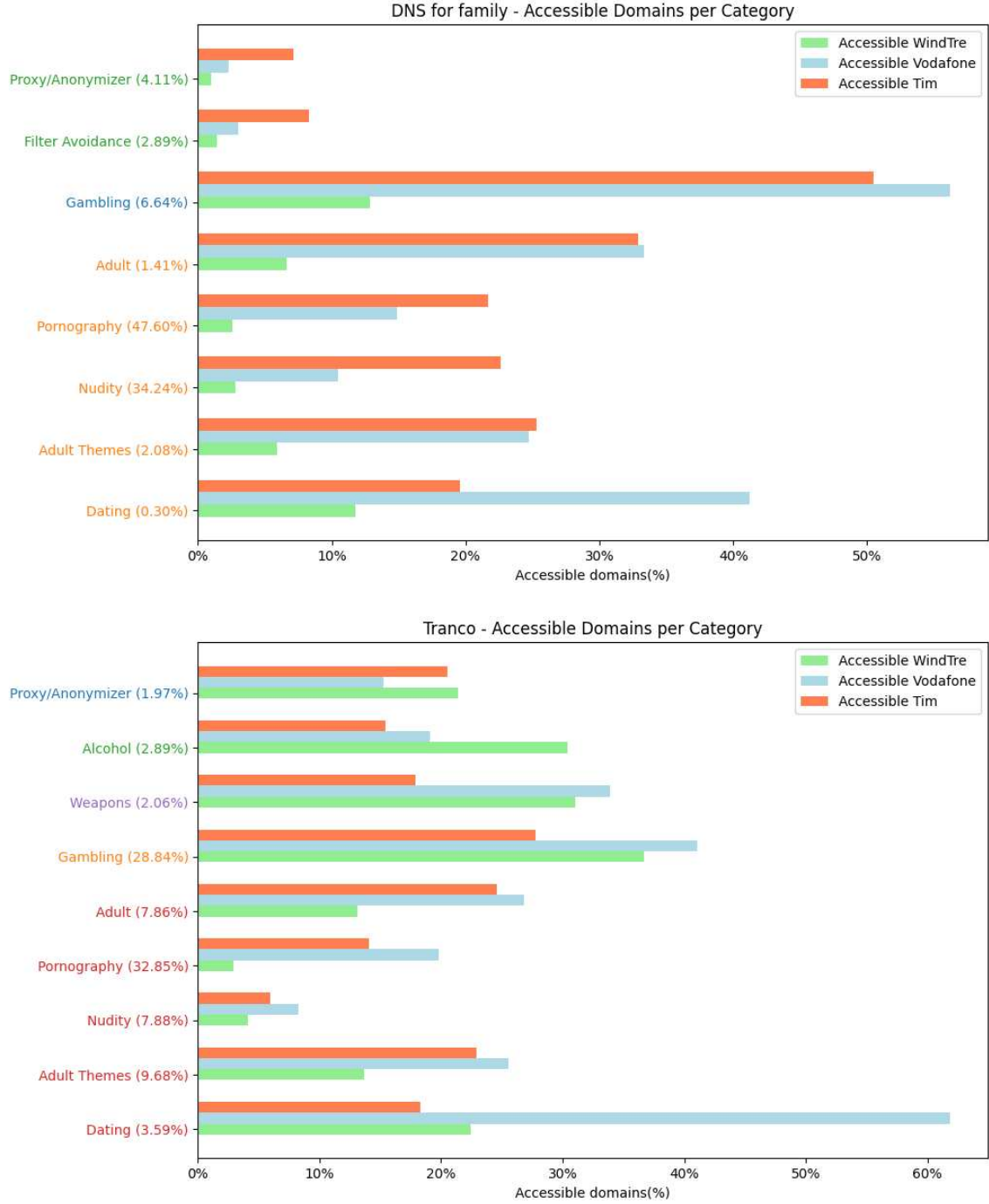


FIGURE 4.8: The figures show, for each dataset, the percentage of accessible domains for each ISP, grouped by content category, along with the distribution of each category in the overall dataset

4.3 Bypassing Techniques

Our final study focused on evaluating the robustness of ISP parental control mechanisms against bypassing attempts. As detailed in the methodology section, we assessed this by testing three different techniques: commercial and enterprise VPNs, HTTPS requests,

and Google DNS. The figures below 4.9 present the results of the measurement obtained with and without these techniques, while keeping parental control enabled at all times. Specifically, in the figures the number of accessible and blocked domains are compared, and the "Others" category was omitted for clarity. The following sections present a detailed analysis of the effectiveness of each bypassing method.

4.3.1 VPNs results

The two figures 4.9 illustrate the results obtained using a paid commercial VPN and the University of Twente VPN. As shown, the results are similar in both cases. For TIM ISP, no domains were blocked when the VPN was enabled. This contrasts sharply with the results obtained without a VPN, where 4,123 domains were blocked and 3,843 were accessible out of a total of 10,000 tested domains from our blocklist. For Vodafone and WindTre, the tests with VPNs enabled showed a small number of domains returning an SSL error, which we initially attributed to the parental control filtering mechanisms. However, further analysis revealed that these errors were common across all ISPs, indicating that they were caused by factors unrelated to the parental control system. From these tests, we conclude that by using the two targeted VPNs, we effectively bypass parental control mechanisms by creating an encrypted tunnel between the client and the VPN servers, preventing ISPs from intercepting or filtering the traffic.

4.3.2 HTTP requests results

The results obtained using the HTTP request technique compared to the HTTPS request technique revealed distinct differences in ISP behavior based on their implemented blocking mechanisms. The chart highlights these variations.

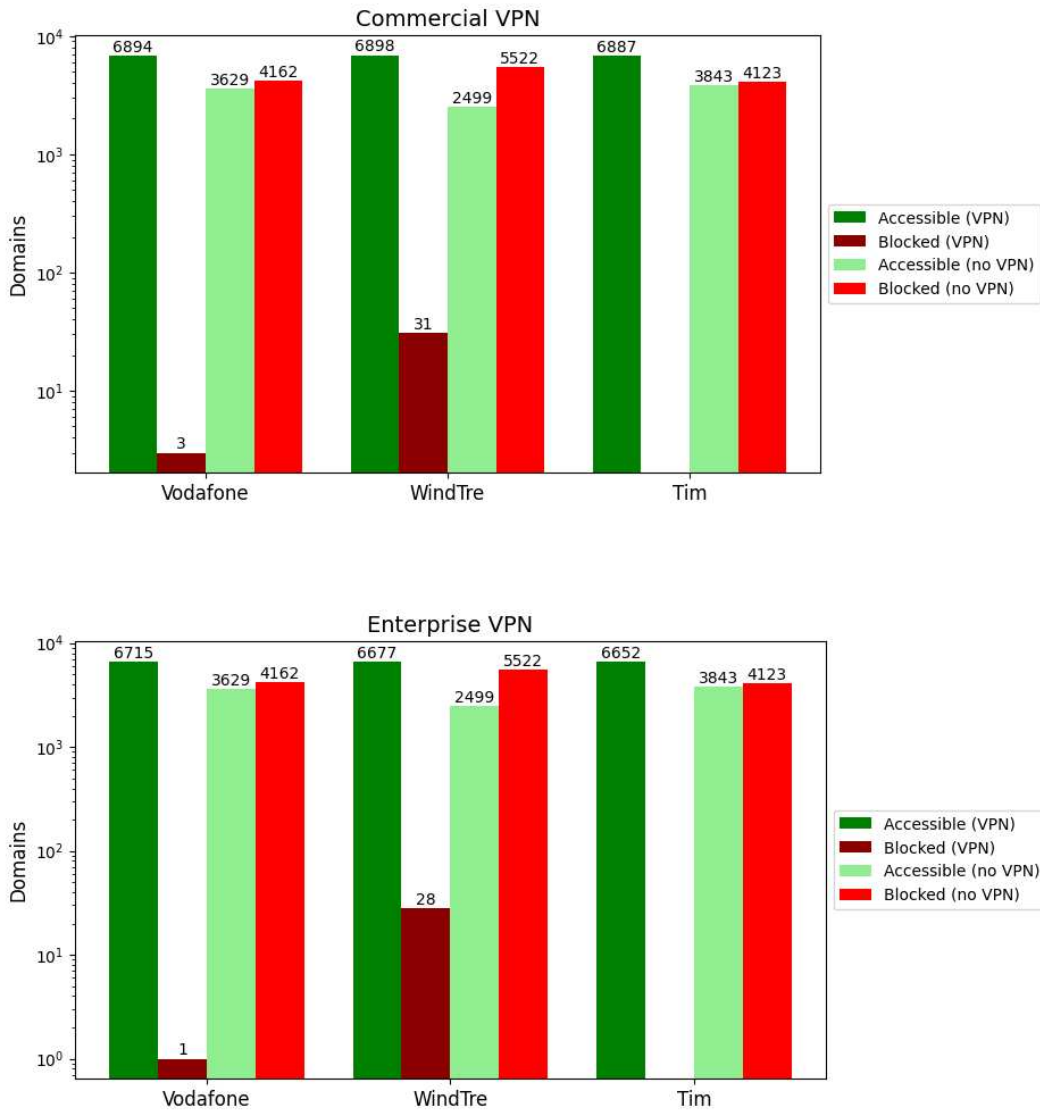
For TIM, which employs a DNS-based parental control filter, the results align with expectations 4.9. In fact, since our analysis focuses on TLS over HTTP requests, TIM's filtering mechanism is unaffected, as the blocking occurs at the DNS level, preventing HTTP requests from being processed in the first place. As a result, the measurements for both HTTP and HTTPS requests show minimal differences that may be attributed to the network conditions, confirming that the DNS-based parental control is independent of the request type.

However, for Vodafone and WindTre, the number of blocked domains in the HTTP test was significantly lower than in the HTTPS test. This discrepancy can be explained by the way HTTP-to-HTTPS redirection works: approximately 30% of the HTTP requests were redirected to HTTPS via status codes 301 or 302, after which the parental control filter operated at the TLS level, blocking some domains.

The remaining 70% of domains, however, continued to be processed over HTTP, effectively bypassing the parental control filters of Vodafone and WindTre, which only operate at the TLS level of HTTPS requests. This raises a critical security concern—not only does the filtering mechanism fail to block harmful content in these cases, but it also indicates the continued use of outdated and insecure infrastructures. HTTP traffic, which lacks encryption, exposes sensitive user data, posing significant privacy and security risks. Despite HTTPS becoming the default web standard in 2020, the high percentage of domains still supporting HTTP highlights the persistence of these vulnerabilities.

4.3.3 Google DNS results

Finally, we evaluated the parental control mechanisms by testing two types of DNS resolvers: the ISP's default DNS and Google's public DNS (8.8.8.8). The results reflect the nature of the filtering mechanisms implemented by each ISP. For Vodafone and WindTre, which enforce parental control at the TLS level rather than the DNS level, the use of Google DNS did not produce significant changes in the number of blocked or accessible domains. The slight variations observed were likely due to network fluctuations rather than differences in the filtering behavior. In contrast, TIM exhibited a clear difference between the two tests, confirming that its DNS-based filtering mechanism can be bypassed simply by changing the DNS resolver. This highlights a fundamental limitation of TIM's parental control system, as users can circumvent it by manually configuring an alternative DNS service.



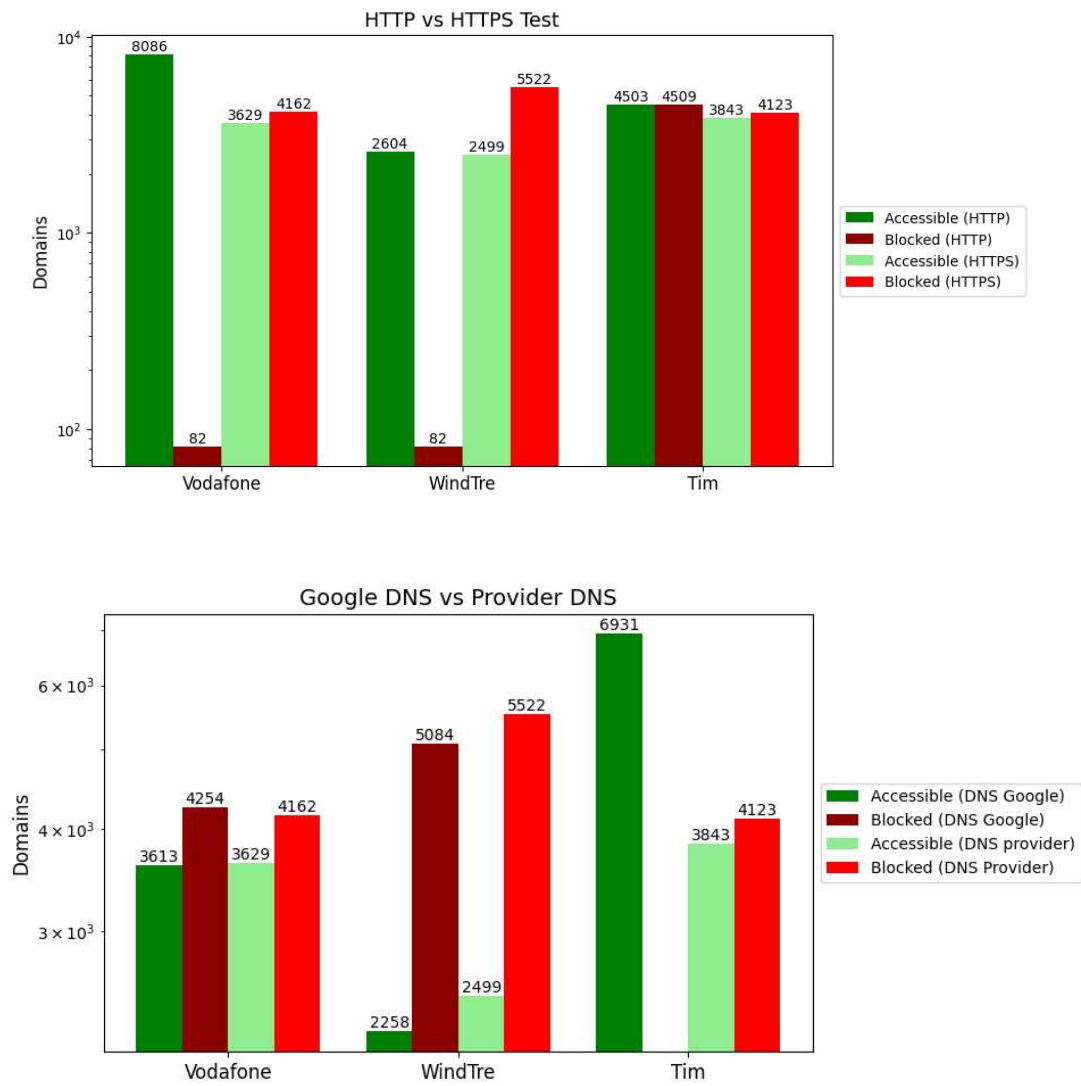


FIGURE 4.9: Results of the measurements for each bypassing techniques

Chapter 5

Discussion

With this research, we contributed to the academic literature by analyzing the technical aspects, effectiveness, and robustness of parental control solutions implemented by the main Italian ISPs—TIM, Vodafone, and WindTre. We assessed the extent to which these solutions comply with AGCOM regulations and how well they protect minors from harmful online content.

All three ISPs have complied with the AGCOM resolution by offering free parental control solutions, but the technical implementations differ significantly. TIM relies on a DNS-based filtering mechanism, which blocks access to restricted domains by returning a blocking IP address. In contrast, Vodafone and WindTre implement filtering at the TLS level, intercepting TLS communications and injecting a manipulated certificate that terminates the connection.

Our analysis highlights differences in effectiveness and robustness across providers. The proportion of unfiltered domains in our blocklists is not negligible, exceeding 20% overall. When focusing on high-risk categories—such as pornography and gambling websites, which are both widely accessible and potentially harmful to minors—the results are particularly concerning. The percentage of unfiltered pornography websites ranges from 21% to 30%, while gambling websites are left accessible in up to 50% of cases.

Moreover, filtering effectiveness varies significantly across providers. WindTre performs well in blocking adult content with percentage significantly lower compared to the other ISPs, but fails to adequately restrict access to gambling, alcohol, and drug-related websites. Vodafone and TIM exhibit similar behavior, with underblocking rates exceeding 30% for most categories and reaching 50% for gambling sites. These disparities suggest that, despite AGCOM’s resolution, not all minors receive the same level of online protection.

Finally, we evaluated the robustness of parental control mechanisms against three potential bypassing techniques. Our results show that for both VPNs we tested, the filtering systems were bypassed for all the providers analyzed. This suggests that VPN usage could significantly undermine the effectiveness of these filtering solutions. Moreover, we observed that the ability to circumvent parental controls depends on the filtering method implemented. Specifically, when filtering occurs at the DNS level, changing the DNS resolver to one different from the ISP’s allows users to bypass the restrictions entirely. Conversely, when filtering occurs at the TLS level, simply switching the URL from `https://` to `http://` causes the filtering system to fail. Additionally, this experiment revealed that, although HTTPS is the default standard for web traffic, only 30% of HTTP requests are redirected to HTTPS. This raises serious concerns not only regarding the ease of bypassing parental controls but also about online security and privacy, as plaintext communication remains exposed.

The results of this study raise important concerns for parents about whether relying solely on ISP-provided parental controls is sufficient to protect children from online risks. Although the AGCOM resolution is titled "Protecting Minors in Cyberspace", our findings highlight a lack of consistency and transparency in how these controls are enforced across different ISPs. Some ISPs fail to effectively block harmful content, while others rely on filtering mechanisms that can be easily circumvented by minors. As a result, children are not equally protected online, creating disparities based on the parental control solution offered by the ISP on which their parents rely.

5.1 Limitations and future work

During the course of our work, we encountered several limitations that impacted the scope and depth of our analysis.

Firstly, since our measurements were conducted in the Netherlands for Italian ISPs, we faced roaming limitations that restricted our data usage to 10 gigabytes per month. As a result, we had limitations on the number of measurements we could conduct, as well as the ability to repeat tests multiple times to ensure consistency and reliability. This constraint affected the comprehensiveness of our dataset and could have introduced some variability in the results.

Additionally, we encountered a technical issue with TIM's parental control, which stopped functioning for almost a month during the testing period. After investigating the potential causes, we were able to restore the service, but the disruption during that period impacted our ability to conduct coherent and fair comparisons between the providers.

Furthermore, our analysis focused exclusively on the mobile solutions provided by the three main ISPs (TIM, Vodafone, and WindTre). A broader study, which includes fixed broadband solutions as well as other ISPs, would provide a more comprehensive overview of the state of parental control in Italy and offer insights into whether these results are consistent across different types of internet access.

Another limitation arises from the way we categorized and filtered the datasets. We relied on the Cisco Umbrella Investigate categorization system and selected domains based on the semantic similarity between category names and descriptions provided by Cisco and those defined in the AGCOM resolution. However, the use of another categorization system could have led to different results, as well as the filtering criteria that are not entirely objective. Consequently, some domains may have been included or excluded based on our interpretation, which could affect the accuracy of the datasets and, by extension, the validity of the results.

Finally, through our tests, we were able to identify the presence of a Man-in-the-Middle (MITM) middlebox in the parental control systems of Vodafone and WindTre that performs certificate injection. However, due to the nature of our tests, we could not definitively determine whether this middlebox functions as a proxy, an invisible proxy, or another type of entity. A future investigation could focus more deeply on the functionalities of parental control systems to clarify the role and impact of such entities in the filtering process.

Chapter 6

Conclusions

In this research, we presented an evaluation of the parental control solutions provided by the main Italian ISPs—Tim, Vodafone, and WindTre—to comply with the AGCOM resolution of November 2023. These solutions offer basic filtering mechanisms free of charge to protect minors from harmful content online. Through a DNS and HTTPS response analysis, we investigated their functional mechanisms to understand at which network level the filtering occurs.

We then evaluated their effectiveness by analyzing the presence of underblocking, i.e., harmful domains that should be blocked but remain accessible instead. We tested two blocklists created from the Tranco and DNS for Family datasets, and our results showed percentages of accessible domains ranging from 20% to 34%, values that we do not consider negligible, as they show significant exposure to harmful content.

We also conducted an analysis based on the domain categories included in our blocklist to assess how underblocking differs across categories. Even in this case, the results were concerning, as for the most risky categories—such as adult content and gambling—the ISPs fail to provide adequate protection.

Finally, we assessed the robustness of the targeted solutions against several bypassing techniques. Our findings revealed that, depending on the technology implemented, the parental control systems exhibit vulnerabilities that can be exploited to bypass them.

In the next section, we address the research questions stated at the beginning of the paper.

6.1 Research answers

We recall the proposed research questions:

- RQ1: *Which network techniques do the three targeted ISPs, i.e., Tim, Vodafone and WindTre use for their parental control blocking mechanism?*

Based on our results, TIM has implemented a DNS filtering mechanism that blocks content matching a domain blocklist. If the queried domain is in the blocklist, the ISP's DNS does not resolve the actual IP address of the domain but redirects the user to a TIM block page.

On the other hand, Vodafone and WindTre have implemented parental control solutions that operate by intercepting TLS traffic during the process of sending HTTPS requests. The blocking occurs during the TLS handshake between the client requesting the content and the web server. Specifically, when the server responds to the Client Hello with its own Server Hello, it sends a manipulated certificate.

In the case of Vodafone, this triggers an alert on the client side because the certificate is signed by an untrusted Certificate Authority (CA). In the case of WindTre, the alert arises due to an incomplete certificate chain. In both instances, these alerts cause the connection to be closed.

- RQ2: *Is there any evidence of underblocking, i.e., domains that should be blocked but are accessible instead for the parental control solutions provided by the targeted ISPs?*
We tested two different blocklists over a period from December 2024 to January 2025, and our results show that the percentage of requests successfully processed ranges from 20.5% for WindTre, 22.7% for Tim, and 34.2% for Vodafone. These results are significant because the percentage of accessible domains is very similar to that of blocked domains, indicating a weak filtering effectiveness. In the case of Tim, the percentage of blocked domains is even lower, at 26.5%.
- RQ2.a: *How does underblocking affects the categories identified by the AGCOM?*
All the domains in our blocklist were classified using the Cisco Umbrella categorization tool to conduct our category analysis. The most common categories in our datasets are adult content, followed by gambling and anonymizer categories. The presence of underblocking varies significantly among ISPs, with the percentage of unblocked domains ranging from 2.9% for WindTre in the pornography category to 21.7%, and up to 56.2% for Vodafone in the gambling category, highlighting inconsistency in filtering effectiveness.
- RQ3: *Can the blocking mechanisms of the targeted ISP's parental control solutions be bypassed?*
We tested four different bypassing techniques: a commercial paid VPN, the enterprise VPN of the University of Twente, sending HTTP URL requests instead of using HTTPS, and using Google DNS instead of the provider's DNS. The outcome of our analysis showed very similar results for the two VPNs, as in both cases, the block was bypassed, regardless of the blocking method implemented. This is due to the fact that when using these VPNs, the traffic is redirected to their servers and encrypted, making it invisible even to the ISPs. When using Google DNS, the performance of Vodafone and WindTre's parental control remained unchanged, as their blocking does not occur at the DNS level. However, in the case of TIM, the results showed that by using Google DNS, the block was bypassed. Finally, sending HTTP requests yielded interesting results. Vodafone and WindTre's parental controls could be bypassed because only 30% of the traffic is redirected to HTTPS, while the remaining 70% is processed via HTTP, where TLS interception cannot occur. This not only poses issues for the robustness of the parental control but also introduces security concerns due to the traffic being unencrypted.

Bibliography

- [1] Agcom. *Comunicato stampa del 07 Ottobre 2024*. Accessed: 2024-10-18. Oct. 2024. URL: <https://www.agcom.it/comunicazione/comunicati-stampa/comunicato-stampa-del-07-ottobre-2024>.
- [2] Agcom. *Delibera 9/23/CONS | Agcom*. 2023. URL: <https://www.agcom.it/provvedimenti/delibera-9-23-cons>.
- [3] S. Ali et al. “Parental Controls: Safer Internet Solutions or New Pitfalls?” In: *IEEE Security & Privacy* 19.6 (2021), pp. 36–46. DOI: [10.1109/MSEC.2021.3076150](https://doi.org/10.1109/MSEC.2021.3076150).
- [4] Allot. *ALLOT - Network Intelligence Security Solutions and Protection*. 2024. URL: <https://www.allot.com/>.
- [5] H. H. M. Altarturi and N. Badrul Anuar. “A preliminary study of cyber parental control and its methods”. In: *2020 IEEE Conference on Application, Information and Network Security (AINS)*. 2020, pp. 53–57. DOI: [10.1109/AINS50155.2020.9315134](https://doi.org/10.1109/AINS50155.2020.9315134).
- [6] H. H. M. Altarturi, M. Saadoon, and N. Badrul Anuar. “Cyber parental control: A bibliometric study”. In: *Children and Youth Services Review* 116 (2020), p. 105134. DOI: [10.1016/j.childyouth.2020.105134](https://doi.org/10.1016/j.childyouth.2020.105134).
- [7] E. Blancaflor et al. “A Vulnerability Assessment on the Parental Control Mobile Applications’ Security: Status based on the OWASP Security Requirements”. In: *11th Annual International Conference on Industrial Engineering and Operations Management*. 2021. DOI: [10.46254/AN11.20211104](https://doi.org/10.46254/AN11.20211104).
- [8] BT. *How to keep your family safe online with BT Parental Controls and the different blocking categories*. URL: <https://www.bt.com/help/security/how-to-keep-your-family-safe-online-with-bt-parental-controls-an>.
- [9] Y. Cheng et al. “In-Depth Evaluation of the Impact of National-Level DNS Filtering on DNS Resolvers over Space and Time”. In: *Electronics* 11.8 (Jan. 2022), p. 1276. DOI: [10.3390/electronics11081276](https://doi.org/10.3390/electronics11081276). URL: <https://doi.org/10.3390/electronics11081276>.
- [10] Jakub Dalek et al. “A method for identifying and confirming the use of URL filtering products for censorship”. In: 2013. DOI: [10.1145/2504730.2504763](https://doi.org/10.1145/2504730.2504763).
- [11] Cisco DevNet. *Check status and categorization of domains*. 2024. URL: <https://developer.cisco.com/docs/cloud-security/check-status-and-categorization-of-domains/>.
- [12] T. J. Dishion and R. J. McMahon. *Parental monitoring and the prevention of problem behavior: A conceptual and empirical reformulation*. Vol. 177. 1998, p. 229.
- [13] ExpressVPN. *ExpressVPN*. 2024. URL: <https://www.expressvpn.com/>.
- [14] DNS for Family. *DNS for Family*. 2024. URL: <https://dnsforfamily.com/>.

- [15] Á. Feal et al. “Angel or Devil? A Privacy Study of Mobile Parental Control Apps”. In: *Proceedings on Privacy Enhancing Technologies* 2020.2 (2020), pp. 314–335. DOI: [10.2478/popets-2020-0029](https://doi.org/10.2478/popets-2020-0029).
- [16] W. Fuertes et al. “On the Development of Advanced Parental Control Tools”. In: *IEEE Xplore*. 2015. URL: <https://ieeexplore.ieee.org/document/7812938>.
- [17] Vahiny Gnanasekaran and Katrien De Moor. “Usability, Security, and Privacy Recommendations for Mobile Parental Control”. In: *Proceedings of the 2023 European Interdisciplinary Cybersecurity Conference (EICC '23)*. 2023, pp. 138–143. DOI: [10.1145/3590777.3590800](https://doi.org/10.1145/3590777.3590800).
- [18] Virginia Media. *Parental Controls and Web Safety for Children | Virgin Media*. 2024. URL: <https://www.virginmedia.com/broadband/parental-control>.
- [19] A. Moctezuma and S. Robila. “A Framework for Evaluating Parental Controls for Streaming Services”. In: *2023 IEEE Integrated STEM Education Conference (ISEC)*. 2023, pp. 199–203. DOI: [10.1109/ISEC57711.2023.10402305](https://doi.org/10.1109/ISEC57711.2023.10402305).
- [20] V. O’Neill B. Dopona. *BIK Policy Map - BIK Portal - BIK Community*. 2014. URL: <https://www.betterinternetforkids.eu/en/policy/bikmap>.
- [21] ofcom. *Protecting children*. URL: <https://www.ofcom.org.uk/online-safety/protecting-children/>.
- [22] D. J. P et al. “Granular Network Traffic Classification on Mobile Application Service for Better Parental Control”. In: *Mar. 2024*. 2024. DOI: [10.1109/iatmsi60426.2024.10503043](https://doi.org/10.1109/iatmsi60426.2024.10503043).
- [23] M. Rowe and R. King. “An investigation into the performance of UK internet providers’ web filters”. In: *Lancaster EPrints* (2015). URL: <https://eprints.lancs.ac.uk/id/eprint/76435>.
- [24] Sky. *Parental controls on Sky Mobile | Sky Help | Sky.com*. 2024. URL: <https://www.sky.com/help/articles/parental-controls-sky-mobile>.
- [25] TalkTalk. *TalkTalk Help & Support*. 2018. URL: <https://community.talktalk.co.uk/t5/Articles/About-mobile-parental-controls/ta-p/2204354>.
- [26] Tim. *TIM Navigazione Sicura Bambini su rete mobile TIM*. https://www.tim.it/content/dam/flytoco-areapubblica-aemfe/tim_it/pdf/fisso-e-mobile/mobile/servizi/tim-navigazione-sicura-bambino/guida-tim-navigazione-sicura-bambini.pdf. Accessed: Oct. 10, 2024. 2024.
- [27] Tranco. *Tranco List*. 2024. URL: <https://tranco-list.eu/>.
- [28] UNICEF. *Protecting Children Online*. 2022. URL: www.unicef.org/protection/violence-against-children-online.
- [29] V. Varadharajan. “Internet filtering-issues and challenges”. In: *IEEE Security & Privacy* 8.4 (2010), pp. 62–65.
- [30] Ververis, V., Ermakova, T., Isaakidis, M., Basso, S., Fabian, B., Milan, S. *Understanding internet censorship in Europe: the case of Spain*. Accessed: 2024-10-18. 2021. URL: <https://doi.org/10.1145/3447535.3462638>.
- [31] R. Viorescu. “2018 Reform Of EU Data Protection Rules”. In: *European Journal of Law and Public Administration* (2017), pp. 27–39. DOI: [10.18662/eljpa/11](https://doi.org/10.18662/eljpa/11).
- [32] Vodafone. *Servizio di controllo parentale*. URL: <https://www.vodafone.it/privati/area-supporto/applicazioni/servizio-di-controllo-parentale.html>.

- [33] WINDTRE. *Servizio Protezione Minori - Parental Control* / WINDTRE. 2024. URL: <https://www.windtre.it/protezioneminori>.
- [34] Wireshark. *Wireshark*. 2024. URL: <https://www.wireshark.org/>.