To Block Or Not To Block? Evaluating Parental Controls Across Routers, DNS Services, and Software

Anonymous Authors

***Abstract*—In today's society, ensuring the safety of children online has become a priority for parents, educators, and policymakers. Parental control systems are widely used to restrict access to age-inappropriate online content, offering a tool for protecting children from potential online risks. However, their blocking behavior across different technologies remains hidden and poorly analyzed. In this paper, we present a comparative analysis of seven parental control solutions: three provided by routers with built-in filtering functionalities (TP-Link, Netgear, and ASUS), two DNS-based services (OpenDNS and DNS.eu), and a software tool (Norton Family). Each system was tested by systematically visiting a large set of domains and recording which were blocked. We then classified all domains into content categories and used this classification to evaluate the consistency and selectivity of each parental control system. Our findings show that parental controls frequently fall short of providing adequate protection: Some systems fail to block up to 96% of inappropriate domains, while others block appropriate content or apply inconsistent and arbitrary filtering rules. For example, ASUS allows access to over 95% of gambling websites, and Netgear exhibits internal inconsistencies, with one-third of the domains blocked under its "Adult" settings remaining accessible under its stricter "Child" setting. These results provide insight into the real-world operation of popular parental control technologies and show a methodology for evaluating category-based filtering at scale.**

***Index Terms*—Network measurements, Use and digitalization, Parental control.**

I. Introduction

The Internet has become an integral part of children's lives, offering a variety of educational resources, entertainment, and opportunities for social interaction. However, this increased digital presence comes with significant risks. Children are exposed to a wide range of online risks, including age-inappropriate content, online predators, privacy invasions, and cyberbullying. These dangers are documented in numerous studies and reports, such as those by UNICEF [1] and Amnesty International [2]. In response, a wide range of parental control tools have become available, promising to safeguard children by filtering inappropriate content, monitoring online activities, setting time restrictions and blocking access to age-inappropriate websites. While parental control systems are an attractive solutions for parents, their effectiveness is difficult to verify due to a lack of transparency in how they operate. Many parental control solutions present themselves as tools capable of addressing a wide range of risks, but they often function as black-box systems, operating while hiding their internal functioning. As a consequence, users remain unaware of which content is being filtered, how filtering decisions are made, and whether these tools truly fulfill their promises. This lack of transparency raises concerns about the actual protection provided, including the risk of under-blocking (where harmful sites remain accessible) and potential over-blocking (where legitimate sites are restricted), affecting user's trust.

This paper aims to empirically analyze how parental control systems operate, focusing on their content blocking mechanisms and their blocking effectiveness. Our study spans a range of deployment types, including consumer routers, DNS-based services, and software solutions, to represent the diversity of tools currently available to families.

Anna @all: we need to revise those contributions once the results are in to ensure consistency

The contributions of this work are as follows:

- We conduct an empirical analysis of parental control mechanisms implemented in three consumer routers, two DNS-based filtering services, and one software solution, using network traffic data collected during domain access attempts.
- We evaluate the behavior of each system on a pre-classified list of popular domains, identifying what content is blocked and inconsistencies or anomalies in the blocking logic.
- We document the techniques these systems appear to use for filtering, such as DNS manipulation, keyword matching, and application-level interception, and reflect on their transparency and reliability.

This paper is organized as follows: Section II provides background information and reviews relevant literature on parental control technologies. Section III details our experimental setup and data collection process, while in Section IV we describe the domain list used to evaluate blocking behavior. In Section V, we present our findings on how different systems implement filtering and what types of content they block. Section VI discusses the limitations of our study. Finally, Section VII presents our conclusions.

II. Background & Related work

A. Parental Controls

Parental control tools are designed to assist parents in monitoring and managing their children's online activities, with the goal of protecting children from potential online risks. Risks include exposure to harmful content, cyberbullying, and

excessive screen time, as such risks can impact children's cognitive, emotional, and social development [3].

Prior work has categorized online risks using the "4Cs" framework (Content risk, Contact risk, Conduct risk, and Contract risk). This framework is used to distinguish between exposure to harmful material, interaction with potentially dangerous individuals, engagement in risky behavior, and privacy issues or commercial exploitation [3], [4]. Parental control tools can help mitigate these risks, targeting content categories such as pornographic or violent content, hate speech, and drug or gambling-related websites. These are frequently highlighted in risk taxonomies studies as high-priority threats to children online [5], [6].

Parental control tools typically include features such as content filtering, time management, and tracking the location and activities of specific devices. They can be deployed on various platforms (e.g. computers, mobile devices, video games consoles, and televisions), and in this case they are called *device-specific* controls. If the controls target a specific application on a device (e.g. a browser), we talk about *application-specific* controls. Finally, we have *network-based* controls, which aim at blocking undesirable activity (e.g. requests to an unappropriated website) at the network level. This class of parental control is implemented, for example, on a router or by specialized software. This paper focuses on network-based parental controls and their content-filtering functionalities.

B. Network-Based Parental Controls

Network-based parental controls are a subset of parental controls that can be implemented using a variety of techniques. These solutions typically rely on methods such as IP blocking, URL filtering, Deep Packet Inspection (DPI) and DNS filtering [7]. Each method serves a different purpose in restricting access to online content. IP-based blocking blocks traffic to specific IP addresses. However, this method can be overly broad, as many unrelated websites may share the same IP, leading to possibly significant overblocking. URL-based filtering inspects the full web address (URL) being accessed and blocks or allows the request based on a filter list. It is more precise than IP blocking but requires deeper traffic inspection, which can be hindered by encryption. Deep Packet Inspection (DPI) inspects not only the headers but also the content of packets to identify content to be blocked. While precise, DPI is resource-intensive and privacy-invasive. A different approach, DNS filtering, act on the Domain Name System, a fundamental service that translates human-readable domain names into numerical IP addresses, enabling users to access websites. DNS filtering works by intercepting DNS queries and modifying the responses (DNS injection), blocking access to restricted websites by preventing the resolution of their DNS queries. This technique is relatively easy to implement and can be done at various levels, including public, ISP, and open DNS resolvers, making it a versatile tool for content control [8].

C. Related work

Despite the increasing availability of parental control tools, the adoption rate remains relatively low [9]. For instance, the internation survey Global Kids Online shows that parental controls are used by less than 3% of parents in several countries, and EU Kids Online data showed usage ranging from 11% to around 33% across European nations [10]. The reasons for this low adoption rate are multiple, including the balance between security, privacy and usability. Parental control tools are at times perceived as difficult to use, too invasive or ineffective [9]. In addition, these tools often operate as black-box systems, offering limited insight into their operations, including which types of content are blocked and how such content is identified [11]. The lack of transparency raises concerns about the actual protection provided and the overall impact on user trust. While recent tools increasingly integrate AI, edge computing, and cloud-based monitoring to improve filtering accuracy and usability [12], [13], these advancements do not fully resolve the conflict between functionality, transparency, and privacy. In this paper, we investigate the blocking behavior of parental controls, in particular focusing on over-blocking or under-blocking of content.

The broader field of parental control has received attention in terms of user experience and effects on family dynamics, often in the form of user-focused surveys [9], [10], [14] However, to the best of our knowledge, only a limited number of studies have examined how parental controls work from a technical point of view. Ali et al. [15] investigates the security and privacy risks of various parental control implementations, including router-based solutions. Feal et al. [16] reveals serious privacy and security flaws in mobile parental control tools and Ali et al. [11] claims to be the first comprehensive study analyzing different parental control solutions across multiple platforms. These findings highlight how parental control internal mechanisms remain poorly explored and evaluated. This work aims to addresses this gap by a network-level analysis of their real-world behavior.

At the technical level, several of the approaches supporting network-based parental control have been investigated in the context of other application areas. Works like [17]–[20] investigated the use and characteristics of IP-based blocklists. Feal et al. [17] investigate 2093 free and open source IP-based blocklists to gain a better understanding about, among others, their coverage, liveliness and scope. Ramanathan et al. [18] focus on the interplay between IP-blocking and IP address reuse, arguing that IP-based blocking could lead to un-intended blocking of legitimate users. The work of Li at al. [19] uses network measurements to infer the use of IP blocklists in the wild. URL filtering and DPI are often studied in the context of Internet censorship, as for example in [21]–[24]. Bock et al. [22] report that censorship middleboxes can react to TLS server name indication (SNI) fields. At the DNS level, a recent study by Liu et al [25] investigate the security implications of relying on protective DNS services, namely DNS resolvers that implement content filtering. Our paper also includes DNS-

based filtering, but focuses instead on the filtering of content harmful to young Internet users. DNS-filtering has also been studied in the context of newer DNS implementations (e.g. DoH [26]) and to assess the risk of content underblocking [8], or overblocking [27], [28].

III. Methodology

Parental controls are typically marketed with generic claims about their functionalities, promising to protect children while providing no information about their functioning. To investigate their actual implementations and effectiveness, we examined the consumer market for parental control solutions across hardware, software and the DNS. To try and ensure broad coverage and relevance we aimed to identify devices and services that are likely to be adopted by actual families. Our selection was guided by three criteria:

1) Popularity, which we determined trough user reviews and ratings (considering both the amount of reviews and the score) on retail platforms such as Amazon.
2) Market relevance, recognized from mentions in consumer-focused rankings (e.g. blog-posts about the best parental controls of the year on technology forums) and professional reviews.
3) Diversification, aiming to include major providers to reflect the available commercial offerings, rather than focusing on a single vendors solutions.

From this process, we selected three router solutions, one software solution and two DNS-based filtering mechanisms, which are summarized in Table I.

Solution	Type
TP-Link (Archer AC1900)	Router
Netgear (Nighthawk AX5400)	Router
ASUS (RT-AX57)	Router
Norton Family	Software
OpenDNS FamilyShield	DNS
DNS.eu	DNS

TABLE I
Parental control solutions analyzed in the study

To evaluate how each parental control system handles different types of content, we designed our tests around a popularity list of domains (see Sec. IV) and the set of identified parental control solutions. To identify which domains should be blocked by a parental control, we rely on content categories, such as porn and gambling, using domain classification data (see Sec. IV). In particular, we tested how each solution blocks or allows each of the domains in the popularity list. Since we expected the selected solutions to have different blocking methodologies (e.g. DNS or HTTPS), we used different approaches targeting the specific solutions. We therefore first define a behaviour baseline for each solution (See V-A).

Each of the identified solutions necessitates of a specific data collection methodology, which we explain in the following subsections.

A. Routers

To analyze how each router implements its parental control functionalities, we built a controlled environment involving three physical routers and three dedicated client PCs. Each PC is connected via LAN to one router. Each router is connected to the Internet via its WAN port, allowing it to resolve DNS queries and apply filtering policies as it would in a real-world scenario. To capture and analyze traffic, we tapped both the LAN and WAN connections: The LAN capture records the traffic sent by the client PC to the router and the router's responses. The WAN capture monitors traffic between the router and the Internet, including DNS lookups and connections to web servers.

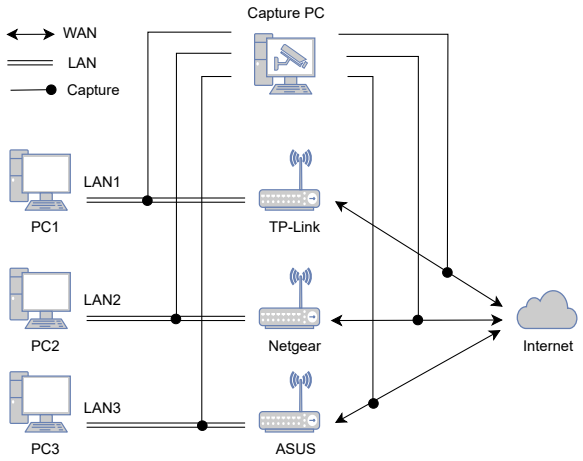

Fig. 1. Capture setup to investigate the behaviour of the routers.

This setup, shown in Figure 1, gives us an overview of the traffic entering and exiting each router. We captured PCAPs of each connection and used Wireshark to analyze them and extract information about the behaviour of the routers. By comparing what is sent by the PC and what is allowed to leave (or blocked) by the router, we can observe exactly how each device enforces its parental control policies, especially at the DNS level. Each router was tested with parental controls enabled and disabled, allowing us to isolate and understand the specific blocking behavior.

1) The TP-Link case: The TP-Link router uses a set of user-defined keywords as base for its filtering mechanisms. This implies that we need to initialize such a set before conducting our experiments. In this case, we tried to put ourselves in the shoes of a parent confronted with such a task, and we therefore asked ChatGPT to create a list of keywords (31 words, as this is the maximum allowed by TP-Link), based on a set of website categories of inappropriate or harmful content (see Sec. IV).

B. Software

For software-level parental controls, we analyzed Norton Family. We installed the software solution on a Windows machine. We iterate over the input list of domains using a curl

script and we monitored the outgoing traffic using Wireshark to collect possible evidence of blocking behaviour.

C. DNS

OpenDNS FamilyShield and DNS.eu Kids are publicly accessible DNS resolvers that advertise family-safe filtering. To analyze their behavior, we created a measurement script that uses the dig to send DNS queries to each service. We limited the number of daily queries sent to each service to prevent rate-limiting or automatic blocking.

IV. Dataset

In the following sections we provide details on our list of input domains and the classification data.

A. List of Input Domains

We use the Tranco Top 1 Million (downloaded on 19-02-2025, root domains) for our list of diverse input domains. Tranco[1] is a research-oriented ranking of the most-visited websites and includes domains spanning a variety of content categories such as social media, news and search engines. The list aggregates multiple other top lists, including Cloudflare Radar, the Cisco Umbrella Popularity List, the Majestic Million, and the Crome User Experience Report (CrUX). As Tranco aggregates these lists it addresses some of their shortcomings, which include instability, inter-list disagreement, and susceptibility to rank manipulation. We use the standard version of the list with apex domains.

B. Input Data for Domain Classification

We use the Cisco Umbrella Investigate API for domain name classification. The *domain status and categorization* data[2] offers a detailed set of categories. To make these data more manageable and relevant for evaluating parental controls, we developed a heuristic to consolidate the categories into a smaller and simplified set. This is guided by three key principles:

1) We analyze the co-occurrence of categories across all input domains to identify cases where multiple categories are frequently assigned to the same domain (an example will follow in section IV-C). Such categories are candidates to be collated into a simpler category;
2) We then collate similar and overlapping categories that provide additional granularity but no meaningful distinction in terms of web content filtering. For example, we merge *Radio*, *Music*, and *Entertainment* into the simpler *Entertainment*;
3) We consider categories through the specific lens of parental control system evaluation, and infer if a category should typically be blocked for children. For example, we expect the simpler *Adult Content* category, which collates Cisco categories such as *Pornography* and *Dating*, to not be meant for children.

[1]https://tranco-list.eu/

[2]https://umbrella.cisco.com/products/umbrella-investigate

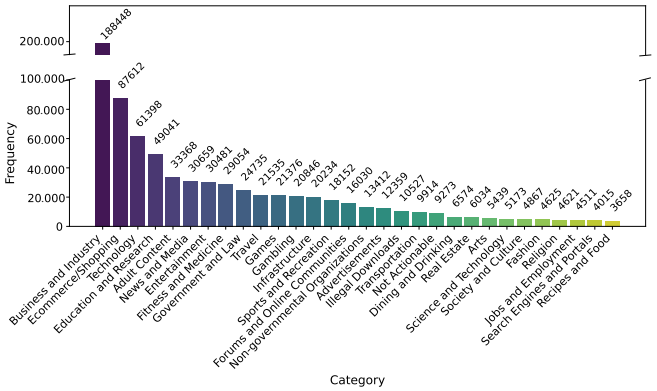

Fig. 2. Tranco Lisy Top 30 Content Categories by Frequency.

This results in a simplified, content-oriented classification that facilitates the measurement of blocking behavior across different parental control solutions. The consolidated categories serve as the basis for our exploration of which types of content are blocked by each parental control system.

C. Content Distribution Across Input Domains

We mapped the Tranco list to categories to understand the distribution of content. We were able to map 779,591 domains using the Cisco data. For the other 220,409 domains the Cisco data does not offer a category. Figure2 shows the most prevalent categories among the Tranco domains.

Further analysis revealed that uncategorized domains appear more frequently toward the long tail of the list. Only 8K domains missing a category are among the Top 100K of the Tranco list. We can only speculate as to the reasons for this difference, but the position of domains in the list suggests that categorization services may be more effective for high-traffic domains.

We next analyzed the extent to which domains were assigned multiple categories. The vast majority of domains (95%) mapped to a single class, with fewer than 4% assigned two classes, and less than 0.5% assigned three or more. We conducted an analysis of category co-occurrences, which revealed that several classes appeared together so frequently that they effectively functioned as a single category, validating our intuition to collate and simplify categories. For example, the *Shopping* category co-occurred with *Ecommerce/Shopping* in 86,7K cases, while the prior appeared without the latter fewer than 2K times.

By collating Cisco categories into simpler ones, we reduced the number of domains still assigned multiple categories by several orders of magnitude, from hundreds of thousands to a few thousand at most. Of the consolidated categories, we consider *Adult Content*, *Gambling*, *Hate/Discrimination* and *Terrorism* to likely be candidates for blocking by parental control systems. In the remaining of this paper, we will use those categories as reference for comparing the blocking capabilities of different solutions. In Figure 3 we show the cumulative distribution of domains under these four categories

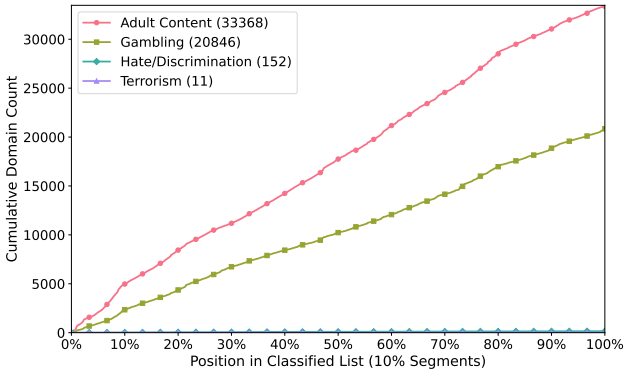

Fig. 3. Distribuion of Domains in Sensitive Categories.

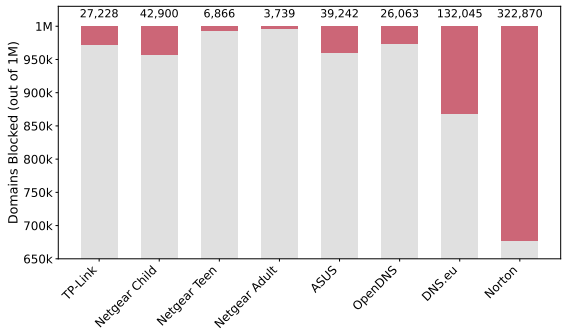

Fig. 4. Total number of domains blocked by each parental control system.

in order of rank. As shown, the number of domains belonging to *Hate/Discrimination* and *Terrorism* is extremely low, while *Adult Content* and *Gambling* are much more prevalent. In general, we observer that domains in these categories are quite evenly spread throughout the entire Tranco list.

V. Results

In this section, we present the results of our analyses. First, we show the baseline domain filtering analysis of all the parental control systems, providing insights into how each system blocks domains. Next, we show their overall blocking behavior and examine their blocking mechanisms by category, showing how each system handles different types of content.

A. Baseline Domain Filtering

As outlined in Section III, we analyzed the PCAPs of each router's LAN and WAN connections to assess their parental control mechanisms. We established a baseline for blocking behavior by testing HTTPS connections both with and without parental control. Our findings show that the routers do not apply HTTPS-based filtering. Instead, they rely on DNS-based filters, each using a different approach.

TP-Link, for instance, implements a basic keyword-based filter mechanism, blocking DNS queries that contain user-defined keywords. Our packet capture analysis confirms this behavior: such DNS queries are visible on the LAN connection between the computer and the router, but are absent on the WAN connection between the router and the Internet. This suggests that the router intercepts and drops these DNS requests before they are forwarded to the external network. In contrast, Netgear uses a more advanced approach. It intercepts DNS queries and validates them against an internal API at *https://urldb.meetcircle-netgear.co*. If a domain is flagged as inappropriate or harmful, the router blocks access by responding with an IP address from the *10.123.0.0/16* range. Additionally, this router allows users to fine-tune blocking with four protection levels: Child, Teen, Adult, and None. ASUS, on the other hand, redirects all DNS queries to Cloudflare for Families (*IP address: 1.1.1.3*) when parental control is enabled. If a domain is deemed inappropriate, Cloudflare for Families returns the IP address *0.0.0.0* instead of the correct address, blocking access by preventing the resolution of the domain. Based on our analysis, the decision on whether a domain is allowed or not is made entirely by the Cloudflare service.

Regarding the Norton software, our initial approach was similar to the one used for the routers, capturing PCAPs using Wireshark during browsing sessions. However, Norton was not suited for this type of measurement. We observed inconsistent signs of activity, such as failed TLS handshakes, or unexplained delays. These artifacts were not regular or reliable enough to support a large scale measurement. For this reason, we developed an alternative detection strategy. After exploring several approaches, including monitoring the behavior of the browser extensions installed by Norton, we settled on a curl-based script that determine if a domain is blocked by analyzing its HTML content. Specifically, we look for known indicators of a Norton block page, such as the presence of the term "NortonLifeLock" and references to the logo file used in the block page.

Finally, for the DNS resolvers, OpenDNS FamilyShield and DNS.eu Kids, we send a DNS query and record the response for domains from the Tranco list. A domain is classified as blocked if the response is empty (for DNS.eu Kids) or if it contains a known block IP address (*146.112.61.106* for OpenDNS).

Key takeaway: Routers and DNS resolvers implement parental control systems through DNS-based filters, involving user-defined keywords, external services, or redirecting to a blocking page. In contrast, software-based systems require analyzing the HTML content of the webpage. This attests the diversity of blocking methodologies among parental controls.

B. Overall Blocking Behavior

After analyzing the blocking mechanisms of each system, we evaluated the total number of domains blocked to assess their effectiveness in filtering inappropriate content. As shown in Figure 4, Norton blocks 322,870 domains, significantly more than any other system, followed by DNS.eu with 132,045 blocked domains. Netgear's blocking behavior varies significantly depending on the setting. For example, in the Adult setting, it blocks fewer than 4k domains, while in the Child setting, it blocks over 10 times as many (42,900). The teen

setting, on the other hand, blocks 6,866 domains. The variation is expected, as these configurations are designed for different use cases. However, we also expected less variation between the Child and Teen settings than between the Teen and Adult settings, as both Child and Teen options are supposed to filter content for younger users, while the Adult option is typically less restrictive. In comparison, ASUS blocks a closer number of domains (39,242) to the Netgear's child setting. TP-Link, which allows users to specify keywords to block, blocks approximately 27,228 domains, more than OpenDNS that blocks approximately 26,000 domains. As a result, OpenDNS blocks the fewest domains among the systems evaluated, excluding Netgear's Teen and Adult filters.

To further understand the nature of this blocking, we also examined the ranking of the blocked domains in the Tranco list, specifically assessing whether the systems prioritize blocking higher-ranked and potentially more popular domains. This is important because blocking popular domains may have a greater impact on real-life user experience and exposure to content compared to blocking low-traffic sites.

Figure 5 shows the cumulative distribution function of blocked domains across the Tranco Top 1 million list. The logarithmic scale is used to highlight differences in blocking behavior across domains with different popularity levels. The x-axis represents the domain rank, from most to the least popular, while the y-axis shows the cumulative number of blocked domains up to each rank, with the axis ending at the total of one million domains.

Netgear Child shows the highest concentration of blocked domains in the early ranks, with 50% of its blocked domains appearing within the top 31% of the Tranco list (rank 310,472), and 90% reaching rank 826,542. This suggests that, despite its limited overall coverage, Netgear Child prioritizes blocking more popular domains. Similarly, Netgear Adult follows a comparable trend, with 50% of its blocked domains reaching rank 329,750. In contrast, Netgear Teen shows a more gradual distribution, with 50% of its blocked domains appearing only by rank 472,049. This suggest that its filtering is more spread out across the ranking, giving less emphasis on popular domains compared to the other Netgar configurations. In contrast, TP-Link shows a relatively flat progression, reaching 50% of its blocked domains only by rank 629,440, and 90% by rank 952,000. This pattern is not due to an inherent filtering policy but reflects the limitations of the keyword-based blocking mechanism. Since domains are blocked only if their names contain a blocklist term, this approach disproportionately affects lower-ranked or obscure domains, where such terms are more likely to appear. Norton, despite its much larger blocking volume, shows a more gradual rise: half of its blocked domains appear by rank 650,143, and 90% by 930,832. DNS.eu, OpenDNS and ASUS reach the 50% thresholds at ranks 430,748, 459,953 and 480,093 respectively, around middle of the Tranco list, indicating a more balanced focus between popular and unpopular domains. On average, across all eight systems, 50% of blocked domains appear by rank 470,331, and 90% by 886,945, indicating that most tools concentrate their blocking in the lower half of the Tranco list.

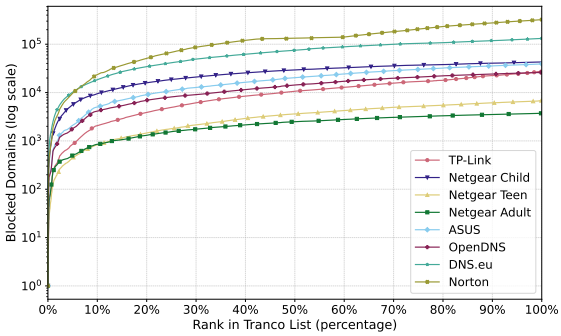

Fig. 5. Cumulative Distribution of Sensitive Domains Across the Tranco Top 1 million.

Key takeaway: Among the different solutions, we observe a large variability in the number of blocked domains with variations in the range of 300k to 7k, which raise questions about the effectiveness of these solutions.

C. Domain Blocking Across Categories

To better understand the types of content most frequently blocked by each system, we evaluated the proportion of blocked domains across various content categories. We calculated the ratio of blocked domains in each category to the total number of domains in that category, as shown in Table II. Several notable patterns emerge, revealing strong differences in filtering priorities. We observe that systems like OpenDNS, DNS.eu, and Norton block a high percentage of domains across the sensitive category *Adult Content*, with blocking rates of 73.5%, 70.5% and 74.5% respectively. Additionally, Norton stands out as the system that blocks over 73% of domains in the *Gambling* category and achieves the highest blocking rates in *Hate/Discrimination* and *Terrorism*, further confirming its overall stricter filtering policy. However, on the other hand, Norton applies its broad filtering even to non-sensitive domains, blocking over 15% of *Education*, *Technology*, and *Business and Industry* sites, raising concerns about overblocking.

In contrast, router-based systems tend to block fewer domains and show more variability in their filtering across categories. Notably, despite being the most restrictive among Netgear's settings, the Child configuration performs poorly in inappropriate categories, blocking only 5.0% of *Adult Content* and 3.7% of *Gambling* domains. Additionally, *Hate/Discrimination* and *Terrorism* categories are blocked at rates of 19.1% and 18.2%, respectively, which is relatively low. It is important to note that the *Terrorism* category is very small (only 11 domains), so even slight differences in absolute blocking result in large percentage differences. Instead, it blocks a larger percentage of non-sensitive domains, such as 28.2% of *News and Media*, 28.5% of *Government and Law*, and 13.4% of *Fitness and Medicine*, notable percentages for categories that typically would not require such strict filtering.

Category (Total)	TP-Link	Netgear Child	Netgear Teen	Netgear Adult	ASUS	OpenDNS	DNS.eu	Norton
Business and Industry (188448)	0.3%	1.7%	0.1%	0.1%	0.3%	0.1%	7.6%	17.8%
Ecommerce/Shopping (87612)	0.4%	4.7%	0.1%	0.1%	0.5%	0.1%	2.3%	17.2%
Technology (61398)	0.3%	2.5%	0.2%	0.2%	0.6%	0.2%	8.2%	23.5%
Education and Research (49041)	0.2%	2.8%	0.1%	0.1%	0.4%	0.0%	6.8%	12.7%
Adult Content (33368)	**27.7%**	**5.0%**	**5.6%**	**6.3%**	**63.8%**	**73.5%**	**70.5%**	**74.5%**
News and Media (30659)	0.5%	28.2%	0.2%	0.1%	0.2%	0.1%	2.3%	12.7%
Entertainment (30481)	0.5%	6.1%	0.4%	0.2%	5.0%	0.6%	4.9%	24.8%
Fitness and Medicine (29054)	0.4%	13.4%	0.1%	0.0%	0.6%	0.0%	4.4%	15.6%
Government and Law (24735)	0.2%	28.5%	0.1%	0.1%	0.1%	0.0%	11.2%	6.2%
Travel (21535)	0.7%	1.2%	0.2%	0.0%	0.1%	0.1%	3.8%	9.4%
			...					
Gambling (20846)	**16.9%**	**3.7%**	**4.4%**	**0.4%**	**0.9%**	**0.2%**	**13.9%**	**75.3%**
Hate/Discrimination (152)	**0.7%**	**19.1%**	**9.2%**	**0.7%**	**10.5%**	**7.2%**	**13.8%**	**46.1%**
Terrorism (11)	**0.0%**	**18.2%**	**27.3%**	**0.0%**	**54.5%**	**0.0%**	**18.2%**	**36.4%**

TABLE II
PERCENTAGE OF DOMAINS BLOCKED BY EACH PARENTAL CONTROL SYSTEM, FOR THE MOST FREQUENT CATEGORIES IN THE TRANCO LIST AND THE FOUR SENSITIVE CATEGORIES.

This suggests a broad filtering strategy that fails to prioritize the inappropriate content typically expected from a parental control system.

ASUS, while more effective in targeting *Adult Content* (63.8%), shows limited blocking in other sensitive categories, suggesting a narrower and more focused approach. Finally, TP-Link, as explained in Sec. III, blocks domains containing any of 31 user-defined keywords and performs well in categories such as *Adult Content* (27.7%) and *Gambling* (16.9%). This suggests that domains related to Adult Content and Gambling are more likely to contain keywords associated with those categories, such as "bet" or "casino", whereas domains in the Hate/Discrimination and Terrorism categories have entirely different names that do not include these keywords.

Overall, it appears that only DNS-based systems and Norton provide robust protection for categories commonly associated with harmful or inappropriate content, while router-based systems, particularly Netgear, struggle to effectively block these inappropriate domains.

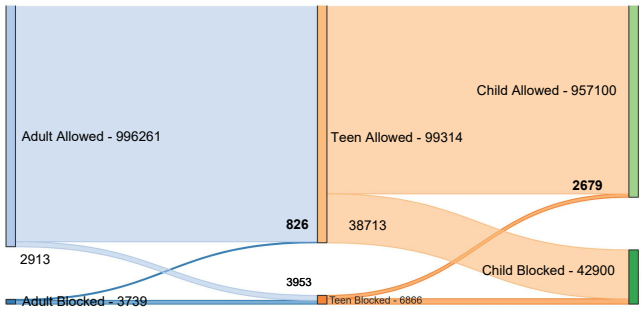

Fig. 6. Domain Blocking Decisions Across Netgear's Child, Teen, and Adult configurations.

Figure 6 shows the flow of domain blocking across the three Netgear's parental control settings. Each column represents one of these settings, while the edges illustrate how domains are handled across different parental control levels. At the Adult level, 3,739 domains are blocked, and 996,261 are allowed. Of the 3,739 domains blocked at the Adult level, most (2,913) remain blocked in the Teen setting, but 826 domains are unexpectedly allowed in Teen despite being blocked in Adult, which contradicts the expected progression from stricter to more permissive filtering. Additionally, 6,866 domains are blocked in the Teen setting. This includes 2,913 domains that were already blocked in the Adult setting, as well as 3,953 domains that were allowed in the Adult setting but are now blocked in the Teen setting. At the Child level, most of the domains that were blocked in the Teen setting are still blocked, but 2,679 domains that were previously blocked in Teen are now allowed, indicating that the Child setting fails to block content that the supposedly more permissive Teen level restricted. Additionally, 38,713 domains transition from "Teen Allowed" to "Child Blocked" , and 957,100 domains remain allowed throughout, which are expected transitions given the increasing strictness of the settings. The blocking logic appears inconsistent, suggesting that each configuration is applying its own filtering criteria, rather than following a coherent, tiered policy. This leads to contradictions among the different settings.

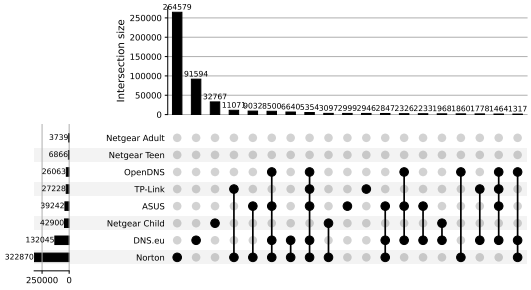

Fig. 7. Overlap in Blocked Domain Across Parental Control Solutions
.

Figure 7 presents an UpSet plot showing the intersection of blocked domains across all parental control systems analyzed in this study. We show only intersections containing more than 300 domains for better visibility. Netgear Adult is the only exclusive set not represented in this graph, due to its small size (21). Each bar represents a group of domains blocked by the combination of systems indicated by the dots below. If a column contains a single dot, it corresponds to the set of domains blocked exclusively by the associated parental control system. The largest set (264,579) consists of domains blocked only by Norton, highlighting it as the most restrictive tool among those measured. DNS.eu follows with 91,594 blocked domains exclusively by its filtering system. The distinctiveness of these exclusive sets highlights how these two tools enforce restrictions more independently and strictly than the others. Most of the largest intersections involve Norton, which is expected, as it has the largest number of blocked domains. Notably, Netgear Child, Teen and Adult, rarely overlap, further reflecting the inconsistent blocking behaviors observed in the previous analysis. Although too small to be clearly visible in the plot, only 205 domains are blocked by all eight parental control solutions, illustrating the lack of alignment across tools and the absence of a unified approach to blocking content.

Key takeaway: The majority of solutions perform better with adult content domains than other sensitive categories. However, over-blocking in non-sensitive categories is often present. Additionally, more restrictive settings do not automatically mean safer blocking and can lead to inconsistency.

VI. Discussion

Our analysis reveals significant variation across parental control systems in both scope and strategy. Norton and DNS-based tools exhibit the broadest and most assertive filtering, blocking hundreds of thousands of domains, including many that are uniquely targeted by them. In contrast, router-based solutions demonstrate more limited coverage and, in some cases, internal inconsistencies that undermine their tiered structure. TP-Link's keyword-based blocking shows success in specific categories but lacks broader effectiveness. The lack of consensus across tools, shown in both the UpSet plot and category-level analysis, highlights the fragmented nature of web content filtering. These findings raise important questions about transparency, policy alignment, and the user's ability to make informed decisions when selecting parental control solutions.

Our analysis is subject to a few limitations. First, our evaluation relies on a third-party domain classification system, in this case Cisco Umbrella Investigate, to assign categories to websites. While Cisco Umbrella Investigate provides a broad and well-maintained categorization, a number of domains (22%) remained unclassified or were labeled as "unknown". This inherently creates a level of uncertainty as we are not able to judge if those domain should be blocked or not. However, it should be noted that the vast majority of unknown domains appeared in the final third of the ranked domain list, meaning they are less popular and less frequently accessed. Consequently, their impact on the overall results, particularly concerning popular domains, is considered to be limited. Secondly, we have further grouped the website categories in those that should be blocked and those that should not. This division was guided by prior literature and common sense, but it clearly contains a degree of subjectivity. Different stakeholders, such as parents, educators, policy makers, and parental control manufacturers may disagree on what content should be considered acceptable for different age groups.

Ale: we actually did more than one measurement, no? This was the case at least for the Cisco classification. Saying "one shot" is really dangerous.

Thirdly, we performed a one-shot measurement to assess the blocking behavior of the selected parental control solutions. This approach gives us an overview of the overall behavior of the parental control systems we selected, but it does not allows us to draw conclusions about possible changes over time of the blocking behavior. A longitudinal study capturing updates in filtering behavior and changes in classification could provide insight into whether the effectiveness of these tools changes.

Finally, we have worked with the version of the Tranco list that does not include subdomains. This means that parental control solutions could potentially react differently to a root domain and a subdomain. We leave this analysis as future study.

VII. Conclusions

In this work, we conducted a measurement and analysis of parental control systems across three types of deployments: routers, DNS providers, and software. Using a classified list of domains, we assessed how each system handled inappropriate and sensitive content, revealing differences in both coverage and implementation strategies.

Our findings show that the effectiveness of parental controls remains uneven across tools and vendors, with no single solution offering comprehensive protection, although DNS-based filtering seems to have a consistently higher efficiency. Often, high effectiveness in blocking inappropriate and sensitive content is accompanied by overblocking for all other categories, mostly among popular domains. And while one could argue that, for the sake of the children, blocking more is better, the effect this has on the general user experience and therefore the adoption of parental controls, remains unclear. In addition, we realized that an higher filtering granularity does not necessarily mean a better blocking effectiveness. In the case of Netgear's router-based controls, some domains blocked at lower levels of control are allowed at levels that should have been more restrictive. By comparison, a simple keyword-based solution like the one implemented by TP-Link easily outperform more complex approaches for categories like Adult Content and Gambling. These inconsistencies, combined with the lack of transparency in how filtering decisions are made, highlight the limitations of current parental control solutions, which work as a black-box without sufficient insight on their blocking strategies.

In regard to possible future work, there several feasible directions. As indicated in Sec. III, our study focused on a set of market leading parental control solutions. However, we realize this ecosystem is in continuous evolution. In particular, our study could be expanded to cover a broader range of parental control systems, including ISP-level filtering solutions, mobile devices and newer AI-based tools. Finally, improving classification coverage and accuracy in web classification remains an open challenge. Using multiple classification systems or developing a custom classifier trained for child safety could reduce the number of unknown websites and improve the analysis of blocking behavior.

REFERENCES

[1] M. Stoilova, S. Livingstone, and R. Khazba, "Investigating Risks and Opportunities for Children in a Digital World: A Rapid Review of the Evidence on Children's Internet Use and Outcomes," UNICEF, Innocenti Discussion Papers, May 2021, series: Innocenti Discussion Papers. [Online]. Available: https://www.un-ilibrary.org/content/papers/10.18356/25211110-2020-03

[2] A. International, "Driven into Darkness: How TikTok's 'For You' Feed Encourages Self-Harm and Suicidal Ideation," Nov. 2023. [Online]. Available: https://www.amnesty.org/en/documents/pol40/7350/2023/en/

[3] S. Livingstone and M. Stoilova, "The 4Cs: Classifying Online Risk to Children," *CO:RE Short Report Series on Key Topics*, 2021. [Online]. Available: https://www.ssoar.info/ssoar/handle/document/71817

[4] M. Stoilova, M. Rahali, and S. Livingstone, "Classifying and responding to online risks to children," Tech. Rep., 2023. [Online]. Available: https://www.lse.ac.uk/business/consulting/assets/documents/Classifying-and-responding-to-online-risk-to-children-Good-practice-guide.pdf

[5] N. Alqahtani, "A state of the art review of Internet risks on children," in *2017 2nd International Conference on Anti-Cyber Crimes (ICACC)*. IEEE, Mar. 2017.

[6] R. M. Machado Fernandes, L. F. Rust da Costa Carmo, and C. L. Rebello da Motta, "A taxonomy proposal of cyber threats involving children and adolescents," in *2022 XVII Latin American Conference on Learning Technologies (LACLO)*, Oct. 2022.

[7] Internet Society, "An Overview of Internet Content Blocking." [Online]. Available: https://www.internetsociety.org/resources/doc/2017/internet-content-blocking/

[8] Y. Cheng, Y. Liu, C. Li, Z. Zhang, N. Li, and Y. Du, "In-Depth Evaluation of the Impact of National-Level DNS Filtering on DNS Resolvers over Space and Time," *Electronics*, vol. 11, no. 8, Apr. 2022.

[9] Z. Iftikhar, Q. R. u. Haq, O. Younus, T. Sardar, H. Arif, M. Javed, and S. Shahid, "Designing Parental Monitoring and Control Technology: A Systematic Review," in *Human-Computer Interaction – INTERACT 2021*, 2021.

[10] M. Stoilova, M. Bulger, and S. Livingstone, "Do parental control tools fulfil family expectations for child protection? A rapid evidence review of the contexts and outcomes of use," *Journal of Children and Media*, vol. 18, no. 1, Jan. 2024.

[11] S. Ali, M. Elgharabawy, Q. Duchaussoy, M. Mannan, and A. Youssef, "Betrayed by the Guardian: Security and Privacy Risks of Parental Control Solutions," in *Annual Computer Security Applications Conference*. ACM, Dec. 2020.

[12] N. I. A. Razak, S. Kamarudin, M. I. M. Shuhud, M. L. M. Zakaria, S. M. Mohd, and A. N. A. Wahab, "PiWall as a home traffic controller: enabling parental control and monitoring," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 6, Dec. 2024.

[13] S. Ramezanian, T. Meskanen, and V. Niemi, "Parental Control with Edge Computing and 5G Networks," in *2021 29th Conference of Open Innovations Association (FRUCT)*, May 2021.

[14] G. Wang, J. Zhao, M. Van Kleek, and N. Shadbolt, "Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety," in *Proceedings of the ACM on Human-Computer Interaction*, Oct. 2021.

[15] S. Ali, M. Elgharabawy, Q. Duchaussoy, M. Mannan, and A. Youssef, "Parental Controls: Safer Internet Solutions or New Pitfalls?" *IEEE Security & Privacy*, vol. 19, no. 6, Nov. 2021, conference Name: IEEE Security & Privacy.

[16] A. Feal, P. Calciati, N. Vallina-Rodriguez, C. Troncoso, and A. Gorla, "Angel or Devil? A Privacy Study of Mobile Parental Control Apps," in *Proceedings on Privacy Enhancing Technologies (PETS 2020)*, Jul. 2020.

[17] A. Feal, P. Vallina, J. Gamba, S. Pastrana, A. Nappa, O. Hohlfeld, N. Vallina-Rodriguez, and J. Tapiador, "Blocklist Babel: On the Transparency and Dynamics of Open Source Blocklisting," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, 2021.

[18] S. Ramanathan, A. Hossain, J. Mirkovic, M. Yu, and S. Afroz, "Quantifying the impact of blocklisting in the age of address reuse," in *Proceedings of the ACM Internet Measurement Conference (IMC 2020)*, 2020.

[19] V. G. Li, G. Akiwate, K. Levchenko, G. M. Voelker, and S. Savage, "Clairvoyance: Inferring Blocklist Use on the Internet," in *Passive and Active Measurement (PAM 2021)*, O. Hohlfeld, A. Lutu, and D. Levin, Eds., 2021.

[20] L. Metcalf and J. M. Spring, "Blacklist Ecosystem Analysis: Spanning Jan 2012 to Jun 2014," in *Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security (WISCS '15)*, 2015.

[21] J. L. Hall, M. D. Aaron, A. Andersdotter, B. Jones, N. Feamster, and M. Knodel, "A Survey of Worldwide Censorship Techniques," RFC 9505, Nov. 2023. [Online]. Available: https://www.rfc-editor.org/info/rfc9505

[22] K. Bock, A. Alaraj, Y. Fax, K. Hurley, E. Wustrow, and D. Levin, "Weaponizing middleboxes for TCP reflected amplification," in *30th USENIX Security Symposium (USENIX Security 21)*, Aug. 2021.

[23] J. Dalek, B. Haselton, H. Noman, A. Senft, M. Crete-Nishihata, P. Gill, and R. J. Deibert, "A method for identifying and confirming the use of URL filtering products for censorship," in *Proceedings of the 2013 Conference on Internet Measurement Conference (IMC 2013)*, 2013.

[24] A. McDonald, M. Bernhard, L. Valenta, B. VanderSloot, W. Scott, N. Sullivan, J. A. Halderman, and R. Ensafi, "403 Forbidden: A Global View of CDN Geoblocking," in *Proceedings of the Internet Measurement Conference 2018 (IMC 2018)*, 2018.

[25] M. Liu, Y. Zhang, X. Li, C. Lu, B. Liu, H. Duan, and X. Zheng, "Understanding the Implementation and Security Implications of Protective DNS Services," in *Network and Distributed Systems Security (NDSS) Symposium 2024*, 2024.

[26] D. Vekshin, K. Hynek, and T. Cejka, "DoH Insight: detecting DNS over HTTPS by machine learning," in *Proceedings of the 15th International Conference on Availability, Reliability and Security (ARES 2020)*. Association for Computing Machinery, Aug. 2020.

[27] N. P. Hoang, A. A. Niaki, J. Dalek, J. Knockel, P. Lin, B. Marczak, M. Crete-Nishihata, P. Gill, and M. Polychronakis, "How great is the great firewall? measuring china's DNS censorship," in *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021.

[28] N. P. Hoang, M. Polychronakis, and P. Gill, "Measuring the Accessibility of Domain Name Encryption and Its Impact on Internet Filtering," in *Passive and Active Measurement (PAM 2022)*, O. Hohlfeld, G. Moura, and C. Pelsser, Eds., 2022.

APPENDIX

A. Ethical Considerations

Our measurement approach involves active measurements, both at the DNS level (routers and DNS services) and the HTTPS level (software). We took care to avoid undue load on the on the infrastructure of the services we were testing. For the routers, we performed fewer than 120 DNS $requests/s$, which is well below the number of requests a resolver is able to handle in operational settings. For the DNS service, taking into account that they involve open resolvers dedicated not only to parental control functionality, we set an even more cautious pace, limiting ourselves to 300K $requests/day$ (3.5 $requests/sec$). Finally, the software solutions appeared not to

directly rely on a back-end infrastructure, so we did not apply any explicit pacing as the measurement is naturally quite slow. To further limit any possible impact, we focused on a one-shot measurement. For website categorization we rely on data provided by Cisco. A number of Top sites contain potentially disturbing content. During the measurements we took care that we never visually load any of those pages.