Analyzing the Influence of Online Advertisements on ISP Network Traffic

Renan Augusto Redel[1], Douglas Zietz[1], Vitor Uchikawa[2], Ricardo J. Pfitscher[2]

[1] Centro Universitário SOCIESC (Unisociesc) - Joinville, SC – Brazil

[2]Mobility Engineering Department
Federal University of Santa Catarina (UFSC) - Joinville, SC – Brazil

renan.redel@outlook.com, douglaszietz@hotmail.com,

vitor.uchikawa@grad.ufsc.br, ricardo.pfitscher@ufsc.br

Abstract. *With the constant internet access increase and with a large offer of content, a priori free of charge, there has been a significant increase in data consumption on the network, and online advertising follows this trend. Online advertising allows advertisers to leverage their capital through its platform and allows the dissemination of products and services within the network. However, the advertises load, which, for the most part, is unwanted for users, impacts internet access plans contracted by customers. This work aims to measure the impact of such advertisements on network traffic. We rely on two network traces to characterize consumption patterns: one from a regional access ISP with 33 customers and 15 days of capture and one from a publicly available dataset with 2000 customers and 3 minutes of capture. The results show that at least 11% of all the traffic volume relates to advertising content on the regional ISP and 38% of the public dataset.*

1. Introduction

With the internet's exponential growth, web advertising has become a highly profitable business for large companies, which make their advertising (AD) platforms available and for the companies that use it to publicize their products widely [Goldfarb and Tucker 2019, Würfel et al. 2021]. However, the traffic originating from advertisements burdens the end consumer, consuming their network bandwidth to download content that users may not desire [Pujol et al. 2015, Mouawi et al. 2015].

Even though internet users can rely on ad-blockers [Pujol et al. 2015] and users are more concerned about computational costs related to ads than the data traffic costs [Gao et al. 2021], the network bandwidth consumed by ads has an intrinsic impact on the contracted internet access plan. More specifically, plans that work on an allowance-based model (a.k.a. plans with data caps), in which users have a monthly limit for download/upload, may have a more explicit consequence, as operators can reduce the provisioned bandwidth or even charge for extra bytes when the downloaded amount of data reaches the contracted limits [Chillemi et al. 2020]. On the other hand, content providers argue that they should leverage advertising to provide the content in a free of charge manner, and the usage of ad-blockers hampers the free access to content.

Efforts in literature have been designated to discuss and analyze web advertisements on the network traffic. [Gao et al. 2021] analyzed smartphone user experience

regarding apps containing ads, in which results demonstrated that users are more concerned about the impacts on the battery than on network traffic. In a similar app context, but at an older research, [Mouawi et al. 2015] evaluated battery and network consumption in different app-ADs categories, results reveal that ad traffic varied from 50 to 1,100 KB per minute. In the seminal work [Pujol et al. 2015] evaluated the impacts of ads and ad-blockers in a European ISP and found that 18% and 1% (requests and bytes) of the total traffic relates to ad traffic. On the other hand, several works investigated the use of ad-blockers and their impact on user experience, revenue of providers, and commercial campaigns [Aseri et al. 2020, Gupta and Panda 2020].

In this work, we measured the traffic related to ads in a regional access ISP. We analyzed 1.4 TB of network traffic collected in a trace from 33 customers. Using a publicly available blocklist of ADs used by an ad-blocker tool[1], we identified the ads' sources through the DNS requests and answers. Then, we accounted for the respective bytes for the following requests from the identified IP address sources. The results show that at least 11% of the ISP traffic relates to ads. Also, for validation purposes, we rely on a public available network trace, the analysis show that advertising data varies from 15% to 39% during 3 minutes capture of network traffic from 2000 customers of an anonymous ISP[Yar 2023].

The remainder of this paper is organized as follows. Section 2 discusses the research works in the literature that discuss advertising and its impact on traffic. Section 3 provides an essential background for understanding this research and discusses the measurement methodology. Section 4 presents the results of ad measurement on the ISP traffic and discusses the limitations of our findings. Finally, Section 5 summarizes this research conclusion.

2. Related work

Research efforts to understand online advertising's impact on traffic are split into two major categories: those investigating the amount of traffic related to ads and those investigating the usage of ad blockers. Interestingly, there is a chronological break in the research direction, with dated work focusing on the first category and more recent works on the second.

In 2015, Mouawi et al. [Mouawi et al. 2015] evaluated the bandwidth and energy consumption of ads in apps and conducted comparisons among several popular ad networks. The authors developed an Android application for the study and connected it to several ad networks. Then, they accessed the resource consumption related to the ads' types: banner, interstitial (pop-ups), and interstitial video. Regarding the network consumption, the results reveal that ad traffic varied from 50 to 1,968 KB per minute. Also in 2015, the seminal work of Pujol et al. [Pujol et al. 2015] evaluated the impacts of ads and ad-blockers in a European ISP. Through passive and active monitoring, the authors captured the ISP traffic and conducted request to the Alexa top 1000 sites. Among other relevant results, the study found that 18% of requests and 1% of bytes in the captured traffic relates to advertising.

The more recent works are going in another direction, looking for the impacts of ad-blocker usage [Aseri et al. 2020, Gupta and Panda 2020]. In their work, Aseri et

[1]https://github.com/nicholasb2101/PiHole

al. discuss the impacts of ad-blocker usage on the revenue of providers and commercial campaigns; the paper also models the revenue of publishers according to the type of users (ad-blockers and non-blockers). Gupta et al., on the other hand, discuss the usage of anti-ad-blockers by publishers. An analysis of the top 500 sites in three classes of sites (top 500 in Germany, top 500 in DACH region, and top 500 news sites) through classification models (i.e., random forest, naive Bayes, and J48) shows that the number of sites that rely on anti-ad-blockers varies from 1.4% to 3.1%.

In a distinct direction, the study conducted by [Gao et al. 2021] examined the user experience of smartphone users with in-app advertising. The paper proposes the RankMiner, an approach to quantify user concerns about specific app issues, including performance costs. The research work collected reviews on three platforms: Android (178,477 reviews), iOS (249,212 reviews), and Windows Phone (33,143 reviews). The findings revealed that most users are more concerned with performance (mainly battery usage) than network traffic-related ones.

In this paper, we differentiate ourselves from recent related work by providing the community with an up-to-date view of the byte traffic associated with advertising. To do this, we collected traffic from an access ISP over 15 days and measured the total bytes of advertising whose source comes from a public list of advertising-supplying domains. Our results show that 11% of the ISP's traffic is associated with advertising, notably higher than the results discussed in the literature.

3. Measuring advertisement traffic

3.1. Background

Web pages represent a fundamental component of the modern digital landscape, serving as platforms for information dissemination, communication, and commercial activities. Although the COVID-19 pandemic imposed a substantial shift to traffic patterns, web protocols, and social media applications are still dominant over the general traffic, accounting for at least 40% of current traffic [Feldmann et al. 2020, wid 2024]. A typical web page comprises various objects, including HTML files, images, videos, and scripts [Kurose and Ross 2021]. Figure 1 illustrates a typical web page from a Brazilian content provider.

As Figure 1 shows, a typical page includes, in addition to the content of interest for customers (i.e., text, images, and other multimedia), advertising content. Many content providers rely on advertising to publish their content free of charge for customers, covering their costs (CAPEX and OPEX) and allowing them to profit with access [Goldfarb and Tucker 2019, Würfel et al. 2021]. Thus, advertising plays a significant role in the ecosystem of web pages, serving as a primary revenue source for many online platforms and content creators. Advertisements are placed on web pages to capture users' attention and promote products, services, or brands. The placement of ads is often guided by factors such as user behavior, content relevance, and monetization objectives.

In the process of downloading the web pages, in a typical request using the HTTP or HTTPS protocol, the browser requests a page associated with the URL entered by the user. The DNS protocol is then used to convert the domain name into an IP address and make the HTTP request to the web server associated to the IP. The server then

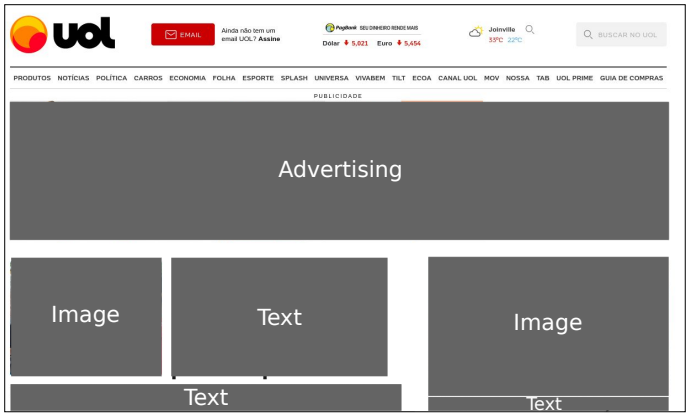

Figure 1. Typical web page and their objects, for copyright reasons we hide the original content.

sends the file of the requested page, composed of links to multiple objects. These objects will then be requested following the same flow, with DNS requests (when the domain name hosting the object differs from the main page) and HTTP(S) for each object [Kurose and Ross 2021].

Considering that most of the advertisement objects are hosted by ad providers (e.g., Google ads), ad-blockers (e.g., Adblock Plus [adb 2024]) rely on filter lists that contain domains that provide ads to block the DNS requests and adjust the pages accordingly [Pujol et al. 2015, Gupta and Panda 2020]. Consequently, the customer's browser does not request the ads objects, saving their bandwidth, and the content provider does not receive its revenue. Another method used by ad-blockers is the hiding of objects. The blockers search in the page's source code for keywords, such as 'banner' or 'ads', to hide the ad-related elements. With the advertisements not being displayed, the blocker should use CSS selectors and rules in the page's source code to reposition the elements.

In this work, we take advantage of filter lists to measure in the network trace the amount of bytes related to advertising. We discuss details of our methodology in Section 3.3.

3.2. Network traces capture

The research conducted in this paper relies on two distinct network traces to analyze the number of bytes related to advertising on ISP traffic. A private company, the Zlink.net, provided the first network trace (named as *private dataset*). The second network trace is publicly available in the Kaggle platform [Yar 2023] (named as *public dataset*). We rely on the second dataset to validate our method and to provide a more up-to-date view of advertising on network traffic.

Zlink.net, an access ISP located in Garuva, Santa Catarina State, Brazil, contributed to this research by providing the private dataset. The company's network topology has clients on one end who make their accesses and requests. These access requests are directed to a concentrator, which consolidates the optical fibers and directs the data to the company's local switch. Finally, the requests pass through a firewall that filters the requests and stores the data that has traveled through the network.

The data collected corresponds to the internet traffic of a network segment that

represents a regional access point to a neighborhood in the city; this point maintained an average of 33 active devices from February 13 to 27, 2021. This network traffic was captured using the packet capture tool available in PAN-OS [pal 2024] and saved in PCAP files, allowing for the further data analysis. For privacy reasons, we can not publish the PCAP files of this trace.

The second network trace contains internet traffic data captured by an anonymous Internet Service Provider (ISP). According to the description available on the publishers' site [Yar 2023], the collection process used sniffer tools over the Mikrotik SDN Controller. The data set includes traffic from over 2000 customers who use Fibre to the Home (FTTH) and Gpon internet connections. The available traces consist of three files with data from one hour of capture: May 6, 2023, 09:29:18 to 09:30:44; May 6, 2023, 09:23:21 to 09:24:47, and May 6, 2023, 09:20:29 to 09:21:19.

3.3. Data processing

The data processing methodology consists of two steps: first, we rely on the Wireshark tool to define the filters and establish what fields we should use to account for the number of bytes related to advertising properly; second, we developed a Python script to consolidate the analyzed data.

3.3.1. Understanding the captured data

Based on the data captured by the ISP, the packets were filtered. The first filter used was for the DNS protocol. This step makes it possible to identify the domain of origin and destination of the requests, which, compared to a list of advertising services, makes it possible to count traffic originating from these sources. A list of domains previously identified as "advertisement" was used to identify advertisements. This list comes from one of the most popular ad blockers and has more than 1 million domains known to be advertisements [pih 2024].

Figure 2 shows a fragment of the Wireshark screen with some responses obtained from the DNS server for the local computer, in which it is possible to identify the resolved domain and the IP addresses that respond to that particular domain. Based on the DNS answers, it is possible to determine the domain being resolved and all the IP addresses the domain in question responds to. Figure 3 illustrates a domain that originated a response with more than one IP address. The figure shows that the domain `a1947.dscb.akaamai.net` responds to the IP addresses `186.192.138.248`, `186.192.138.243`, `186.192.138.232`, and `186.192.138.233`. Thus, for accounting for the advertising-related bytes, one could sum the length of all the packets originating from the IPs whose domains are in the ad list.

Figure 4 illustrates the Wireshark screen when a filter for a specific source (IP address source is equal to `50.116.194.21`) is applied. As the figure shows, with such a filter, one can account for all packet lengths originating from the source, independently of the used protocol.

Figure 2. Visualization of a DNS filter in Wireshark

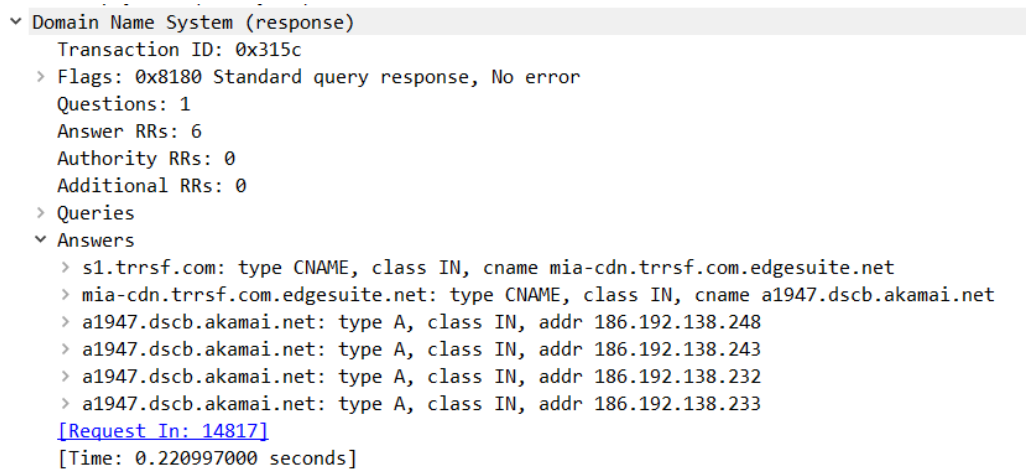

Figure 3. Example of a DNS answer content

3.3.2. Script for data consolidation

Considering that the filtering process explained in Section 3.3.1 allows us to account for the bytes related to ads, we developed a Python script using the DPKT library[2] to automate and consolidate such data. For reproduction purposes, the script is available on GitHub[3]. However, we can not publish the original PCAP files of the private dataset for privacy reasons. Thus, we only published the PCAP files of the public dataset.

The first stage of the script is responsible for reading all the data used during execution. First, the script loads the list of domains with advertisements and saves all the occurrences in a list of strings. We used a list containing 1,335,182 (one million, three hundred and thirty-five thousand, one hundred and eighty-two) domains identified as advertising providers [pih 2024]. Also, we validated our results using an updated list [Black 2024].

The second stage of the script is responsible for filtering the loaded data. As the script reads the PCAP file, it analyzes the DNS responses. These responses rely on the UDP protocol, allowing access to the domains and IP addresses that the DNS server has resolved. For each packet, the script checks if the resolved domain is in the list of domains identified as advertising providers. If this domain is recognized as an advertisement, its IP address and domain address are saved in a Python dictionary. Finally, as shown in

[2]https://github.com/kbandla/dpkt
[3]https://github.com/Vuchikawa/pcapAdFilter

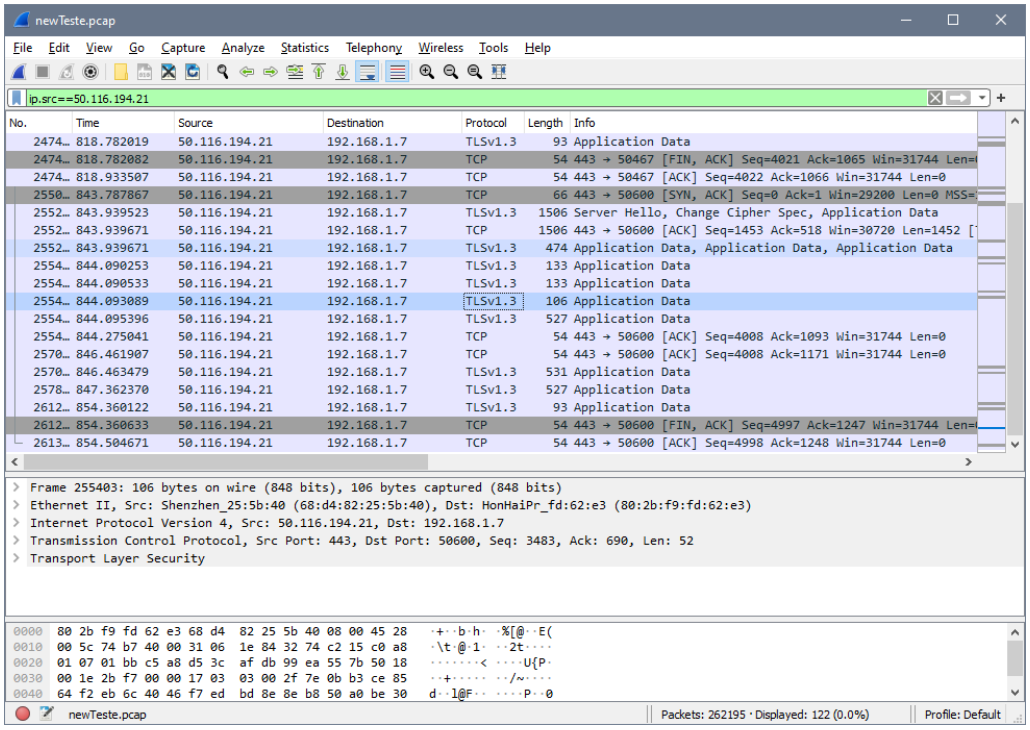

Figure 4. Example of a filter for packets from a specific source

Figure 5, a search is carried out on the entire PCAP file to determine the number of advertisements transmitted.

```
75        for domains in dictionaryAds.values():
76            for ips in domains:
77                if inet_to_str(ip.src) == ips:
78                    print(ips)
79                    print('IP: %s -> %s   (len=%d)\n' % \
80                          (inet_to_str(ip.src), inet_to_str(ip.dst), ip.len))
81                    sizelen = sizelen + ip.len
82                    print("Total Bytes Ads: ", sizelen)
```

Figure 5. Code snippet for data filtering

The code snippet shown in Figure 5 is a script segment developed and used to filter the previously collected data. Line 75 indicates the start of the `for` repetition structure, which will iterate through all the items in the dictionary (`dictionaryAds`), which was filled previously with the ad domain list. In line 76, we have a new repetition structure, which will go through all the possible IP addresses each domain has and search for all the traffic from these IPs. The IP filter applies on line 77, and the values are saved in an accumulator variable (`sizelen`) on line 81.

```
41            if (inet_to_str(ip.src) != 'IPLOCAL'):
42                print('IP: %s -> %s   (len=%d)\n' % \
43                    (inet_to_str(ip.src), inet_to_str(ip.dst), ip.len))
44                sizelen = sizelen + ip.len
45        print("Total Bytes Trafegados: ", sizelen)
```

Figure 6. Code snippet for total bytes downloaded by each customer

The last stage of the algorithm consists of reading the entire PCAP file and identifying the amount of data that has traveled from the internet to the client (download).

Figure 6 shows part of the script used to read the entire traffic. Line 41 has a conditional operation that validates whether the analyzed packet originates from an external IP address and saves the traffic values in an accumulator, as per line 44. It is also necessary to replace the string "IPLOCAL" with the local IP address of the device being analyzed. The local IP addresses were obtained from a simple individual analysis of the files provided by the private company ISP.

During the consolidation process we identified that many name resolution queries run over encrypted DNS requests, such as DNSSEC [Hoffman 2023] and DNS over QUIC [Huitema et al. 2022]. Such an evolution on the DNS protocol occurred for privacy reasons. However, the encryption of DNS requests and answer hinders our ad identification method. Thus, we also decided to account for the amount of data related to encrypted DNS requests.

4. Results

This section discusses the results of the present research. First, we assess the time needed to process the network traces. Afterward, we present the amount of advertising in both the available traces. Finally, we discuss the limitations of this research.

4.1. Processing times

The private dataset contains 15 files in PCAP format of varying sizes. Each file averaged 103.2 GB per day analyzed, which amounted to approximately 1.548 TB of traffic. Due to the large number of files, the analysis script had to be run on two devices to optimize the analysis time. Table 1 specifies the configurations of each device. Both devices have comparable configurations, so the difference in processing time between the devices is irrelevant to this analysis.

Specification	Device 01	Device 02
Processor	Intel Core i5-8250U 1.80GHz	AMD Ryzen 3 1200 3.10 GHz
RAM memory	8.00 GB	8.00 GB
OS	Linux Kernel Version 5.8	Windows 10 Pro

Table 1. Hardware specification of processing devices for the private dataset

The average script execution time was approximately 23 hours for each day analyzed, which represents approximately 7 GB per hour of script execution. The total script execution time was approximately 349 hours. Table 2 details the script execution time in hours, the file size for each day analyzed in GBs, and the execution time.

The public dataset, published on Kaggle [Yar 2023], contains three PCAP files with one hour of capture in each one. The aggregated size of the three files is 2GB. The processing time of the three files was around 1 hour on a Linux virtual machine (VirtualBox) on top of a 12th Gen Intel(R) Core(TM) i7-12700H at 2.30 GHz. The VM had one vCPU and 8GB allocated RAM.

4.2. Analysis on advertising data

Figure 7 depicts the traffic consumption during the fifteen days of capture in the private dataset.

Measurement day	Execution time (hours)	File size (GB)
1	25	101.00
2	29	144.55
3	21	109.50
4	24	103.19
5	30	118.02
6	23	108.33
7	22	99.46
8	26	111.23
9	29	125.97
10	21	96.18
11	19	84.55
12	16	75.45
13	19	87.53
14	21	83.25
15	23	100.70
Standard deviation	4.01	17.7
Average	23	103.26
Total	349	1548

Table 2. Data processing times of the private dataset

For each day, the chart depicts three bars: the amount of ad traffic, the total traffic including ads and excluding QUIC protocol, and the captured amount of bytes without filters. The graph also highlights the captures of weekends and holidays (carnival). Three significant conclusions arise from the analysis of these results: (i) although the traffic pattern shows an increased consumption on non-commercial days, the amount of bytes related to advertising is relatively constant; (ii) the QUIC protocol, generally used in video applications, accounts for a substantial part of the total traffic (37.9% in average, or 67.5 GB per day), and; (iii) the advertising traffic accounts for at least 14 Gigabytes per day, which represents, in average, 11% of total traffic and 18% of unencrypted traffic.

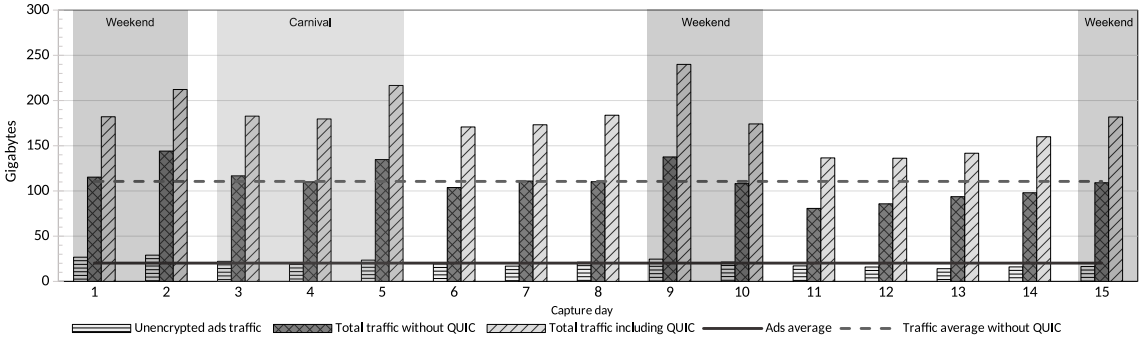

Figure 7. Summary of captured traffic from the private dataset

Figure 7 shows that the total data consumption was above the overall average on only five days of analysis (days 1, 2, 3, 5, and 9), with days 1, 2, and 9 being on weekends. Ads consumption exceeded the overall average on six days of capture (days 1, 2, 3, 5, 9, and 10). A slight increase in the overall data traffic consumption was noted

on Sundays (an increase of 27.3%) and holidays (an increase of 21.8%) compared to the overall average. In advertising consumption, there was an increase in Sundays (an increase of 42.5%) and holidays (an increase of 16.4%). It was impossible to determine if these numbers represent a consumption pattern throughout the month due to the analysis period provided by the ISP not covering more extended periods than the one presented in this research.

To have a more precise view of the impact of ads on unidentified traffic source (excluding QUIC and encrypted DNS requests), we plot the percentage of advertisements during the 15 days of capture. Figure 8 exhibits the percentage evolution and customer traffic consumption. An analysis of ad traffic by customer (the gray continuous line) does not show a clear pattern, with the number of bytes related to advertising varying from 0.4 GB to 0.9 GB. Regarding the percentage analysis, the chart shows that the proportion of advertising on the traffic varies from 15% (days 7, 13, and 15) to 23% (day 1), with an average of 18%. Such results demonstrate that advertisement accounts for a substantial portion of the ISP's daily traffic. Substantially higher than that found in the literature (Section 2, which indicates a growth in ad traffic.

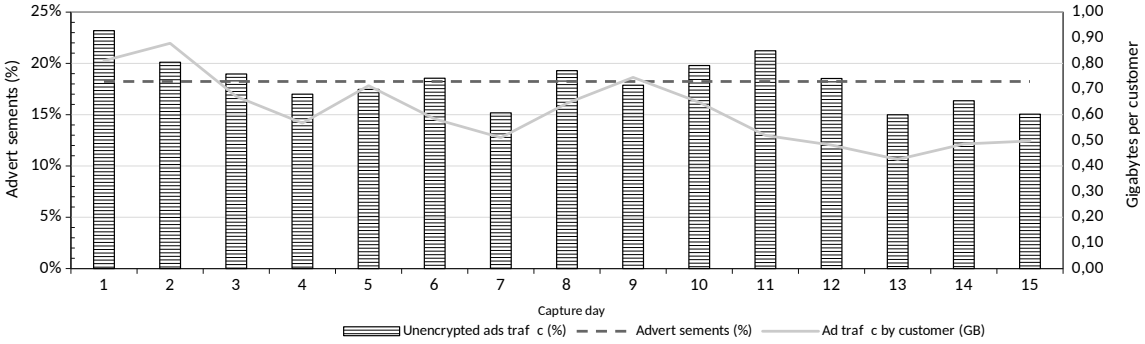

Figure 8. Percentage of advertisements by day in the private dataset

For validation purposes, we analyzed the amount of bytes related to ads in the public dataset. Table 3 depicts the results of advertising traffic analysis on the public dataset.

Interval	Total traffic [MB]	Ad traffic [MB]	Ad percentage
09:29:18 - 09:30:44	775.4	302.2	38.97 %
09:23:21 - 09:24:47	784.8	312.3	39.80 %
09:20:29 - 09:21:19	82.25	12.34	15.00 %
Total	1642.45	626.84	38.16%

Table 3. Traffic analysis of the network trace from the public dataset

The results shown in the Table 3 are remarkable. Although the capture duration is shorter than the private data (3 minutes versus 15 days), the traffic analysis shows that the amount of advertising-related data is substantially higher than that found previously. The first and second minutes contain 39% advertising, and the last minute contains 15%. Such an impressive result supports the need to rely on more extended network traces for a fair comparison and, consequently, have representative results.

4.3. Limitations of the study

Besides our results demonstrating that advertising accounts for a significant amount of ISP traffic, we understand that the impact of ads may have been underestimated, and we point out the following limitations to this research findings:

- *Lack of measurements on encrypted DNS requests* - Figure 7 shows that a significant part of traffic relates to QUIC protocol. Considering that QUIC is an encrypted protocol, it hinders the identification of DNS over QUIC resolutions. Thus, we performed the measurements and analysis disregarding this protocol. As a result, we disregard a significant portion of video-related consumption in the analysis, given that most video provider domains (e.g., YouTube) use the QUIC protocol [Mazhar and Shafiq 2018].
- *Lack of packet content analysis* - as we do not analyze the packets' payload, we could be missing ads hosted by the content provider. Thus, instead of advertisements, we marked ads originating from the same address as regular traffic. An option to circumvent such a limitation is to rely on the information from packet payloads, searching for keywords like 'ad' or 'banner.'
- *Too short network trace* - Although the public dataset contains a network capture of 2000 customers of an anonymous ISP, the capture files contain only 3 minutes of traffic. Such a small duration may result in a biased analysis, since the traffic is susceptible to bursts and specific customer access.

5. Conclusion

In today's Internet architecture, web advertising is essential for content providers offering free-of-charge content. However, the traffic originating from advertisements implicitly impacts end-users, consuming the contracted bandwidth. To understand the size of this impact, we analyzed the advertising-related traffic of an access ISP in this paper. The first step was obtaining traffic files from two ISPs. We developed an automated script to analyze the captured traffic using Python and the DPKT library. Finally, by combining the script and the files provided by the ISPs, we measured the amount of advertising consumed by customers of the two ISPs.

The analysis shows that, on average, 18.2% of the accessed internet content relates to advertising on the private dataset and 31% on the public. During the period analyzed, the average daily advertising consumption was approximately 610 MB per device/household. Consumption also increased on specific days of the week, such as Sundays (42.5% increase in total traffic) and carnival holidays (16.4% increase in total traffic). However, the proportional value of advertising traffic remained within the deviations from the average.

In future works, we intend to analyze the total consumption of advertisements by including the QUIC protocol or other encrypted DNS protocols. We envisage using an AI-based classification tool to account for encrypted ad traffic. In our ongoing research, we are actively monitoring websites and labeling the traffic related to online advertising. For future steps, we will compare classification algorithms to understand which one can obtain better precision for classifying the traffic based on packet features.

References

(2024). Adblock Plus — The world's 1 free ad blocker — adblockplus.org. https://adblockplus.org/. [Accessed 04-04-2024].

(2024). MAWI Working Group Traffic Archive — mawi.wide.ad.jp. https://mawi.wide.ad.jp/mawi/. [Accessed 04-04-2024].

(2024). PAN-OS — docs.paloaltonetworks.com. https://docs.paloaltonetworks.com/pan-os. [Accessed 04-04-2024].

(2024). Pi-hole &x2013; Network-wide Ad Blocking — pi-hole.net. https://pi-hole.net/. [Accessed 04-04-2024].

Aseri, M., Dawande, M., Janakiraman, G., and S. Mookerjee, V. (2020). Ad-blockers: A blessing or a curse? *Information Systems Research*, 31(2):627–646.

Black, S. (2024). Unified hosts file with base extensions. https://github.com/StevenBlack/hosts. [Accessed 12-04-2024].

Chillemi, O., Galavotti, S., and Gui, B. (2020). The impact of data caps on mobile broadband internet access: A welfare analysis. *Information Economics and Policy*, 50:100843.

Feldmann, A., Gasser, O., Lichtblau, F., Pujol, E., Poese, I., Dietzel, C., Wagner, D., Wichtlhuber, M., Tapiador, J., Vallina-Rodriguez, N., et al. (2020). The lockdown effect: Implications of the covid-19 pandemic on internet traffic. In *Proceedings of the ACM internet measurement conference*, pages 1–18.

Gao, C., Zeng, J., Sarro, F., Lo, D., King, I., and Lyu, M. R. (2021). Do users care about ad's performance costs? exploring the effects of the performance costs of in-app ads on user experience. *Information and Software Technology*, 132:106471.

Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1):3–43.

Gupta, R. and Panda, R. (2020). Block the blocker: Studying the effects of anti ad-blocking. *arXiv preprint arXiv:2001.09434*.

Hoffman, P. E. (2023). DNS Security Extensions (DNSSEC). RFC 9364.

Huitema, C., Dickinson, S., and Mankin, A. (2022). DNS over Dedicated QUIC Connections. RFC 9250.

Kurose, J. F. and Ross, K. (2021). *Computer networking: A top-down approach 8th edition*. Pearson.

Mazhar, M. H. and Shafiq, Z. (2018). Real-time video quality of experience monitoring for https and quic. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pages 1331–1339. IEEE.

Mouawi, R., Elhajj, I. H., Chehab, A., and Kayssi, A. (2015). Comparison of in-app ads traffic in different ad networks. In *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 581–587.

Pujol, E., Hohlfeld, O., and Feldmann, A. (2015). Annoyed users: Ads and ad-block usage in the wild. In *Proceedings of the 2015 Internet Measurement Conference*, IMC '15, page 93–106, New York, NY, USA. Association for Computing Machinery.

Würfel, M., Han, Q., and Kaiser, M. (2021). Online advertising revenue forecasting: An interpretable deep learning approach. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 1980–1989.

Yar, M. A. (2023). Internet traffic data set. `https://www.kaggle.com/dsv/5658579`.