Measuring DNS-over-HTTPS Downgrades: Prevalence, Techniques, and Bypass Strategies

JINSEO LEE, KAIST, Republic of Korea
DAVID MOHAISEN, University of Central Florida, USA
MIN SUK KANG, KAIST, Republic of Korea

DNS-over-HTTPS (DoH) is a privacy-enhancing protocol that encrypts plaintext query data in DNS resolution. However, DoH often faces accessibility challenges due to phenomena known as DoH downgrades, where DoH queries are reverted to plaintext DNS queries. Unlike downgrades in other security protocols, which are undoubtedly malicious, the act of downgrading DoH queries can be both desirable and undesirable depending on the context; e.g., enterprise networks are officially advised to avoid or downgrade DoH for security reasons. Recent research has drawn attention to the deeper examination of the phenomena of DoH downgrades, focusing on the prevalence, techniques, and potential bypass strategies. However, existing studies on DoH downgrades have several limitations, notably that they severely overestimate the severity of DoH downgrades across the globe as they lack any distinction between desirable and undesirable downgrades of DoH. In this work, we conduct a large-scale measurement study to provide a more accurate depiction of the DoH downgrade landscape. By minimizing the influence of desirable downgrades of DoH in our measurement probes, we show a skewed long-tail distribution of DoH downgrades across the globe. Our stateful probing techniques also reveal hidden DoH filtering mechanisms that were previously undetected. Furthermore, we design near perfect bypass strategies against existing DoH downgrades. Our study expands our limited understanding of DoH downgrades, offering a more accurate, fine-grained, and comprehensive view of the phenomena.

CCS Concepts: • **Networks** → **Network measurement**; • **Security and privacy** → **Security protocols**; **Privacy-preserving protocols**.

Additional Key Words and Phrases: DNS-over-HTTPS, DNS Privacy, Downgrade, Measurement, Bypass

ACM Reference Format:
Jinseo Lee, David Mohaisen, and Min Suk Kang. 2024. Measuring DNS-over-HTTPS Downgrades: Prevalence, Techniques, and Bypass Strategies. *Proc. ACM Netw.* 2, CoNEXT4, Article 28 (December 2024), 22 pages. https://doi.org/10.1145/3696385

1 Introduction

The Domain Name System (DNS) has used plaintext messages for queries and responses of domain names since its inception. Its plaintext nature, however, has been criticized for leaking sensitive information, such as the domains individual users wish to visit, to third parties [22, 35, 43]. Such criticisms have led to the development and standardization of encrypted DNS protocols, such as DNSCurve [6], DNS-over-TLS (DoT) [26], and DNS-over-HTTPS (DoH) [24], with DoH being the most widely adopted protocol to date. By encapsulating DNS queries within HTTPS, DoH aims to protect user data from eavesdropping and tampering. Despite its privacy benefits, however, DoH

Authors' Contact Information: Jinseo Lee, jinseo.vik.lee@kaist.ac.kr, KAIST, Daejeon, Republic of Korea; David Mohaisen, mohaisen@ucf.edu, University of Central Florida, Orlando, USA; Min Suk Kang, minsukk@kaist.ac.kr, KAIST, Daejeon, Republic of Korea.

ACM 2834-5509/2024/12-ART28
https://doi.org/10.1145/3696385

often faces accessibility challenges due to phenomena known as *DoH downgrades*, where DoH queries are reverted to plaintext DNS queries.

Unlike downgrades in other security protocols, such as TLS downgrades [2, 7] or WPA3 downgrades [44], which are undoubtedly malicious, the legitimacy of DoH downgrades exists in a nuanced gray area. This complexity is illustrated by advice given to enterprise networks, which are often recommended to avoid or downgrade DoH due to potential negative impacts on their perimeter security measures; see the official advice from NSA [15] and CISA [14]. In these contexts, DoH downgrades can be even desirable for maintaining specific security postures.

However, for the average Internet user, downgrades that revert DoH queries to plaintext DNS are predominantly undesirable. These users, unknowingly reverted to much less secure plaintext DNS, lose the privacy protections that DoH offers. This inherently nuanced view has drawn attention in recent research [4, 23, 27, 30], prompting a deeper examination of the prevalence, techniques, and potential bypass strategies for DoH downgrades.

Recent measurement studies on DoH downgrades, however, fall short of offering a precise understanding of the nuanced phenomena due to several limitations. First, studies often overestimate the severity of DoH downgrades across the globe as they do not differentiate between desirable and undesirable downgrades of DoH. For instance, a recent large-scale study [4] reported that approximately 30% of DoH queries are consistently downgraded across many regions worldwide. However, this figure is likely overestimated, as the measurement may have included an unknown number of desirable DoH downgrades, such as those occurring in enterprise networks. Second, the exact mechanisms behind these downgrades have been analyzed in a straightforward, stateless manner, leading to limited and coarse-grained insights into the technical intricacies of DoH downgrades occurring in the wild. Finally, no existing study has explored or presented practical strategies to bypass real-world DoH downgrades. Proposed solutions, such as running privately owned DoH resolvers [23], are impractical for the general user base due to the non-trivial resources and expertise required.

To address these gaps, we conduct a large-scale measurement study to provide a more accurate depiction of the DoH downgrade landscape. By minimizing the influence of desirable downgrades of DoH in our measurement probes, we offer a clearer understanding of the state of undesirable downgrades targeting average Internet users. We show a severely skewed long-tail distribution of DoH downgrades across the globe, with Internet users in a few countries experiencing severe downgrades and many others facing low-intensity downgrades. Our stateful probing techniques also reveal hidden DoH filtering mechanisms that were previously undetected, offering finer insights into the multi-layered techniques employed by real-world downgraders. Moreover, we discover and report the behavior of resilient DoH resolution, observed across three widely used real-world public DoH resolvers, that helps us design highly effective (nearly 100%) bypass strategies against existing DoH downgrades.

As with many other Internet-scale measurement studies, our measurements also have limitations, as we detail in Section 6: despite our best efforts to minimize the influence of enterprise networks, our results may still include some desirable downgrades of DoH; the country- or AS-level aggregation of measured data may contain some noise as these groupings are known to be noisy [42]; and our bypass strategies may not be effective against future DoH downgrades that employ more sophisticated techniques [10, 16, 45]. Yet, our study greatly expands our limited understanding of DoH downgrades, offering a more accurate, fine-grained, and comprehensive view of the phenomena.

We hope that our new findings and insights on the nuanced DoH downgrade landscape will begin the conversation on whether and how to address the privacy concerns of average Internet users who are affected by DoH downgrades. We also believe that our study will help the community

to develop more effective DoH protocols (e.g., a zero-knowledge-proof-based DoH resolution [21]) for better enforcement of query privacy on the Internet.

2 Related Work and Our Motivation

In this section, we first discuss the advent of encrypted DNS protocols. Then, we review existing studies on DoH downgrades and their limitations, which serve as motivation for our work.

2.1 DNS Encryption Protocols

In response to the growing prevalence of DNS-based surveillance [20, 28], there has been an increasing emphasis on encrypting DNS traffic to enhance DNS query privacy. Early encrypted DNS protocols, such as DNS-over-TLS (DoT) [26, 48], aimed to provide TLS protection between clients and DNS resolvers. Later, in 2018, DNS-over-HTTPS (DoH) was proposed to exchange DNS queries and responses over HTTPS [24].

Since DoT and DoH run over TLS, they provide similar security and privacy guarantees. That is, network operators cannot eavesdrop on or tamper with the domain names in DNS queries and responses unless TLS encryption suites are broken or the DoT/DoH resolvers are compromised.

2.2 Downgrading to Plaintext DNS

Confronted with the growing adoption of encrypted DNS protocols, a new technique has emerged and is being deployed in practice: disabling encrypted DNS protocols through downgrades. Network operators can disable encrypted DNS protocols by blocking them and force clients to retry their queries over plaintext DNS.

Huang et al. [27] have reported DoH downgrades, showing that most client browsers are configured to fall back to plaintext DNS by default when they fail to receive timely responses over encrypted DNS. This fallback happens *silently* without any user notification, and thus, users may be left with a false sense of security that their DNS traffic is encrypted and protected. For successful DoH downgrades, the ability to identify encrypted DNS traffic is necessary. Once the encrypted DNS traffic is identified, network operators can block them and force clients to retry over plaintext DNS. Several studies [10, 16, 45] show that identifying DoH packets is possible for downgrades with in-network machine learning (ML) capabilities.

Note that in this paper we focus on DoH downgrades, not DoT downgrades, because (1) DoH is much more widely deployed than DoT, and (2) DoH is designed to blend in with other HTTPS traffic while DoT is not, making DoH downgrades more challenging.

2.3 Measurement Studies and Their Limitations

Several recent studies have examined the state of DoH downgrades, including research by Jin et al. [30], Basso [4], and Hoang et al. [23]. However, these studies have limitations that hinder a comprehensive, precise understanding of DoH downgrades. For instance, Jin et al. [30] conducted their study with a relatively small number of vantage points (around one thousand) and relied on commercial VPN providers. Hence, the generalizability of their findings is limited.

Since Lu et al.'s early work [32] on the overall accessibility of DoH services, more recent studies by Basso [4] and Hoang et al. [23] have further explored the accessibility of DoH services from VPN-based vantage points distributed worldwide. The Open Observatory of Network Interference (OONI) project has also provided public data on the accessibility of DoH services from various countries since December 2020 [5]. These studies have shed light on the *inaccessibility* of DoH services from various countries; yet, they offer limited, if not misleading, insights into the accurate landscape of DoH downgrades. We summarize the limitations of these studies in the rest of this section.

Limitation 1: Downgrades in enterprise networks. Existing studies fail to adequately filter out DoH downgrades occurring in networks that are specifically advised to disable DoH, such as enterprise and organization networks. The NSA [15] and CISA [14] have strongly recommended enterprises and government agencies to disable DoH because DoH may render existing security tools ineffective (e.g., DNS-based malware and phishing detection systems). Therefore, measuring DoH downgrades using vantage points that may represent these networks can lead to inaccurate assessments of the extent of DoH downgrades in the wild, as such downgrades can be desirable for the enhanced security of these entities.

Due to this limitation, previous studies resulted in the overestimation of DoH downgrades in the global landscape. To highlight this problem, we compare our results with those of previous studies. Figure 1(a) illustrates the quantitative differences between our results and those of Basso [4] for the same period. It shows how many DoH queries are downgraded in countries (for more complete definitions, see Section 4.1). We overlay Basso's results on our findings to compare the two. According to Basso's findings, approximately 20–50% of DoH queries are downgraded in most countries, suggesting that DoH downgrades are prevalent worldwide and the intensity of downgrades does not vary significantly across countries. In contrast, our results indicate that DoH downgrades show a wide range of intensities across countries, resulting in a long-tail distribution; see the details of our results in Section 4.1. Hoang et al.'s findings [23] also show that about half of the ASes (9 out of 20) where they detected DoH downgrades are enterprise, cloud, or education networks (e.g., Oracle, DigitalOcean, Alibaba, a university network, etc.).

Limitation 2: Missing details on DoH downgrades. No existing studies consider the detailed factors that affect the existing DoH downgrades, such as the browsers that make queries, HTTP methods, header styles, and their combinations (see Section 4.2 for details). While Hoang et al. use kdig [31] and Basso implements a portion of the OONI project, neither accurately replicates the behaviors of the actual browsers. In Appendix A.2, we provide a detailed comparison of this matter.

Additionally, existing studies infer the filtering techniques used in the real world with simple, stateless probes. Simply sending some probes to differentiate DoH downgrade techniques without storing the state of each probe result can identify only the first filtering technique employed, even if multiple techniques are deployed in sequence.

Limitation 3: Lack of generalizable bypass techniques. No existing studies propose generalizable bypass techniques against the DoH downgrades they observed. Hoang et al. [23] mention that a private DoH resolver can be used to bypass DoH downgrades, but this is not a feasible solution for mass adoption because it requires non-trivial knowledge and resources for each individual.

2.4 Our Approaches

We address the above limitations by employing the following approaches:

- We use residential proxies (see Section 3.1) and also conduct manual sanitization (see Section 3.2) to exclude networks that are advised to disable DoH, such as enterprise and organization networks.
- We examine different combinations of key factors that influence the effectiveness of DoH downgrades (see Section 4.2) and perform stateful probes to identify the deployed filtering techniques (see Section 4.4).
- We discover previously unknown phenomena within major DoH resolvers, then propose generalizable bypass techniques based on these discoveries (see Section 5.1). Additionally, we experimentally demonstrate that these techniques can nearly perfectly bypass existing DoH downgrades in the wild (see Section 5.2).

Table 1. Characteristics of recent measurement studies.

Study	RESIP	Protocol	# Country	Conference
Lu et al. [32]	Proxyrack, Zhima	DNS, DoT, DoH	166	IMC '19
Chhabra et al. [11]	BrightData	DoH	224	IMC '21
Qiu et al. [39]	BrightData	DoH	10+	ISCC '23
Bhowmick et al. [8]	BrightData, Proxyrack	DNS	220	PAM '23
Our study	Proxyrack	DoH	220	-

3 Measurement Methods

We explain our detailed measurement methods, focusing on the vantage points, our manual sanitization, the target browsers under test, and the overall measurement tasks and procedures. As discussed in section 2, all these methods are particularly designed to address the limitations of prior studies. We end this section by discussing the ethical considerations in our measurement study.

3.1 Vantage Points

One crucial requirement for our accurate measurements is to have large numbers of vantage points across many countries to observe the current landscape of DoH downgrades. We choose to use Proxyrack's residential proxy (RESIP) network, where users voluntarily opt-in to share their residential computing resources with financial compensation as vantage points [37]. These RESIP networks are actively used in many recent measurement studies due to their unique characteristic of providing residential computing resources [8, 11, 32, 39]. See Table 1 for a brief comparison of our study with these. With this commercial RESIP network that provides SOCKS proxies deployed in residential networks, we can send DoH queries while mimicking the behavior of various browsers. Overall, we use more than 400,000 vantage points in this RESIP network from 220 countries and analyze the results of 110 countries where we have at least fifty vantage points.

To ensure the reliability of our measurement results, we take extra care in selecting and managing the vantage points: ensuring vantage point consistency across repeated measurements and verifying the population representativeness of our vantage points. We obtain the Autonomous System Number (ASN) and Internet Service Provider (ISP) information corresponding to each vantage point using the MaxMind GeoLite2 database [33].

Consistency across repeated measurements. When we evaluate whether DoH downgrades have changed over time in a country, we conduct multiple, repeated measurements. In all the repeated measurements, we keep the number of vantage points *at the ISP level* consistent across the measurements so that the results are comparable. In other words, the vantage points we use for repeated measurements are always from the same set of ISPs with the same number of vantage points per ISP. If the numbers of vantage points per ISP across repeated measurements differ, we randomly exclude some from the set to strictly enforce the ISP-level consistency.

Population representativeness. We also ensure that our vantage points are well distributed across different ISPs in each country so that the conclusions we draw from our measurement data reflect the DoH downgrades experienced by the typical population in each country. We utilize the APNIC AS population dataset [3] to estimate the number of users our vantage points can effectively represent in each country. This process involves counting the users within ASes where we have at least one vantage point, and then dividing this number by the total number of users in each country. In the countries used in our main analysis, our dataset represents an average of 81.23% of the total population, with a standard deviation of 21.61%.

3.2 Manual Sanitization

Yet another critical requirement for our measurement study is to ensure that the vantage points we use represent the average Internet users in each country. To approximate the vague notion of the average Internet user, we make our best effort to ensure that our vantage points are located within residential networks, not within enterprise, cloud, or education networks.

First, unlike all existing measurement studies on DoH downgrades that utilize VPN nodes, we use residential proxies known to be recruited from residential networks [37]. While these proxies might include some vantage points that are not representative of residential networks, they remain one of the best options available for approximating the average Internet users in each country; see many other recent measurement studies that use residential proxies [8, 11, 32, 39].

We conduct a heavy manual inspection of our dataset to further sanitize it and ensure that our vantage points are located within residential networks. We manually review our dataset to rule out any vantage points confirmed to be located within enterprise, cloud, or education networks. To be specific, for the 110 countries where we have a sufficient number (i.e., at least fifty) of vantage points, we manually verify the top-10 ASes in each country (based on the number of vantage points) and check whether they are managed by ISPs that provide residential Internet services in that country. That is, if the network is an enterprise, cloud, or education network (or if the ISP is not based in the corresponding country), we exclude that AS from our dataset. To associate each ASN with the organization of the AS, we use the IPinfo ASN database [29] and PeeringDB [36]. We then access the website of each organization listed in the datasets. If two datasets point to different websites, we prioritize the latest accessible one. Then, we check whether the organization offers residential Internet service by examining their price or service lists. If the website is not in English, we use Microsoft Edge's translation service. As a result, we have removed 13,164 vantage points (3.2% of the total 406,395 vantage points) from our dataset.

3.3 Other Details: Browsers, Resolvers, and Domains

We outline the details of generating DoH queries for our measurements: client browsers, DoH resolvers, and queried domains.

3.3.1 Client browsers. All the prior measurement studies on DoH downgrades are *browser-agnostic*; thus, the effect of client browsers has been overlooked. Specifically, DoH queries in previous studies were generated by a custom script or DNS lookup utilities compliant with the DoH protocol, but the generated DoH queries are often different from those generated by actual browser implementations.

Instead, we generate browser-specific messages for the DoH protocol, motivated by the observation that different browsers tend to generate slightly different message patterns. Any slight difference in the message patterns between browsers is a critical factor in our DoH downgrade study because downgraders may create filters based on some browser-specific features in messages, and then such DoH downgrade filters may not work for other browsers. We observe different browsers experience drastically different DoH downgrade behaviors; see Section 4.2.

Since our RESIP vantage points do not provide host control, we first monitor the browser-specific messages for sending and receiving DoH queries and mimic the same message patterns using our own script that generates DoH queries. Our detailed analysis of the message patterns of the major browsers is presented in Appendix A. We open-source the script we used to generate DoH queries to facilitate the reproducibility of this work.[1]

3.3.2 DoH resolvers. We use three major DoH resolvers that are commonly listed in the major browsers we test: Cloudflare (cloudflare-dns.com), Google (dns.google), and Quad9 (dns.quad9.net).

[1] https://github.com/NetSP-KAIST/DoH-Downgrades

These three major resolvers have been the subjects of previous studies repeatedly [4, 23, 30, 32] and their information (e.g., hostnames) is also found in the public domain; e.g., see the list of public resolvers [17]. We do not use any custom DoH resolvers in our measurements because our measurement targets, i.e., average Internet users, are unlikely to use such non-major DoH resolvers that are not listed in the browsers.

3.3.3 Queried domains. Network operators will be unable to learn the queried domain names from our measurements, as all the DoH queries we send in this study are encrypted. Nonetheless, we utilize uncensored domain names in each country we test to be extra careful. We particularly avoid using domain names known to be censored in the OONI dataset [19].

3.4 Measurement Tasks and Procedures

The main metric we use is the *DoH downgrade rate*, which is defined as the ratio of the number of downgraded DoH queries to the total number of DoH queries sent. We design the following three measurement tasks:

- **Week-long global measurements.** To understand the DoH downgrade landscape in the world, we conduct our measurement study from a large number of vantage points (around 400,000) in 220 countries. One cycle of this global-scale measurement task takes about a week to complete, and we repeat the same measurement cycle four times in total during our measurement period, which is from August to September 2023. Through the repeated measurements, we observe how the impact of DoH downgrades changes or remains the same over time in each country while we keep the ISP-level consistency (see Section 3.1) over the four consecutive global measurements.
- **Fine-grained time-series measurements.** In addition to the weekly global measurement task, we also conduct a fine-grained time-series measurement task to understand how DoH downgrades change in the order of hours. We conduct this additional measurement in cases where we observe highly fluctuating rates of DoH downgrades over time in the weekly global measurement task. In this study, we take measurement samples made from China as examples and repeat the same measurement cycle with 76 vantage points in China every 2 hours for eleven days, from October 7 to October 17, 2023.
- **Bypass measurements.** To understand how DoH downgrades can be bypassed, we conduct bypass measurements at the vantage points in some countries where we observe a certain level of DoH downgrades (i.e., downgrade rate $> 5\%$) in the weekly global measurement tasks. We carry out the bypass schemes (see Section 5) and measure the changed DoH downgrade rates, from September to October 2023. This experiment confirms that our bypass schemes work nearly perfectly in real-world networks.

Network-condition checks. We ensure that any sign of DoH downgrades in our measurement results (e.g., a DoH query is unsuccessful) is *not* caused by other factors such as network congestion or temporary DoH resolver failures. Ruling out such factors is crucial since we may otherwise draw a wrong conclusion about DoH downgrades.

For any measurement we conduct, we issue five DoH queries from a single vantage point to a single DoH resolver, resulting in five query downgrade results for each pair. We perform a network condition check before and after the five DoH queries to ensure that the vantage point is active and has reliable Internet connectivity. Moreover, we randomly perform additional condition checks during the five DoH queries to ensure stable Internet connectivity. We also test whether a vantage point can directly access a few popular websites for network condition checks. We store the DoH downgrade results *only when all* the network condition checks succeed. A DoH query is said to be downgraded if the vantage point fails to receive any valid response from the resolver.

We further maintain a minimum four-hour interval between the completion of one experiment and the start of the next. These precautions are taken to mitigate any potential bias or disruption arising from excessive query volumes or frequent experimentation.

3.5 Ethical Considerations

Our measurements utilize the commercial residential proxy platform Proxyrack, which provides an opt-in mechanism for volunteers when registering their computing resources for monetary compensation [37]. Proxyrack also informs nodes that they can opt out at any time at their discretion. We purchased the services and always adhered to Proxyrack's Terms of Service. Additionally, our measurements only use benign public DoH resolvers, with a minimum four-hour interval between experiments. Therefore, they do not cause any practical harm to the vantage points, to Proxyrack, or to the public resolvers. We also received a waiver from the Institutional Review Board (IRB) at the authors' institution for our measurement study.

4 State of DoH Downgrades

To present the multi-faceted state of the current DoH downgrades around the world, we present the measured DoH downgrade rates from various perspectives, including countries (Section 4.1), request combinations (Section 4.2), time of measurements (Section 4.3), and the underlying filtering techniques (Section 4.4). We supplement our measurements with geolocation analysis (Appendix B.1). We publish all the measurement results on a website to complement the selected results presented in this section.[2]

Caveats on country-level analysis. In the data analysis in this section, we summarize the statistics of the DoH downgrade rates for each *country* based on the vantage points' IP locations. Yet, we do *not* claim that each country employs centralized, uniform DoH downgrades, nor that these downgrades are conducted by the states. In fact, ASes within the same country exhibit varied downgrade behaviors; see Appendix B.1. In addition, our country-level analysis has inherent limitations because mapping IP addresses to country codes is not perfectly accurate. We acknowledge that our results are not free from this well-known issue and could, therefore, contain errors.

4.1 Global Landscape of DoH Downgrades

Figure 1(a) depicts the downgrade rates of DoH queries sent by clients residing in 110 different countries. The DoH *downgrade rate* of a country is computed as the ratio of the number of downgraded DoH queries to the total number of DoH queries sent from all the RESIP vantage points in that country. To associate each RESIP vantage point with its country of residence, we map the IP address of the vantage point to the country code using the MaxMind GeoLite2 database [33].

DoH downgrade rates may vary depending on the request combinations — we defer the definition and the detailed discussion to Section 4.2. To simplify the discussion of our country-level analysis, we visualize the average (×), maximum (▲), and minimum (▼) downgrade rates across all request combinations for each country.

The most notable observation from Figure 1(a), which sets our study apart from previous work, is the long-tail distribution of DoH downgrade rates observed across countries. Vantage points in a few countries experience extremely high DoH downgrade rates, sometimes nearing 100%, while those in the majority of countries have relatively low rates. This contrasts sharply with the findings of Basso [4], which indicated that DoH downgrades occur consistently across countries with similar rates in the 30–50% range, as shown by the black circles in Figure 1(a). First, there exist a non-negligible number of countries that exhibit severe downgrades of DoH. To better report

[2]https://sites.google.com/view/conext24-doh-downgrade/main

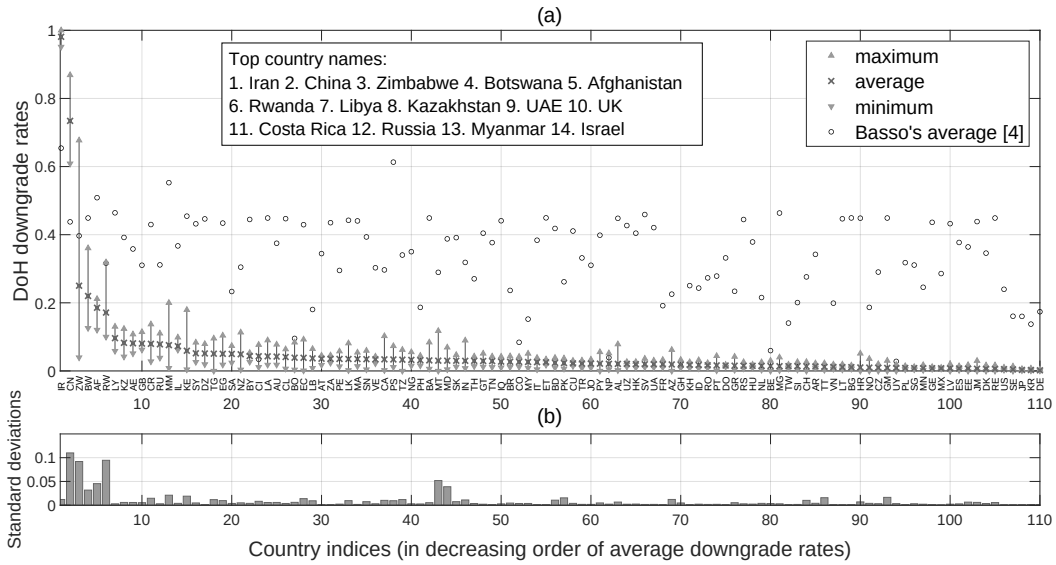

Fig. 1. (a) Four-week global-scale measurement of DoH downgrades. Average, maximum, and minimum downgrade rates across 12 request combinations are shown in 110 countries, ordered by the average downgrade rates. Average downgrade rates from Basso's work [4] are compared. (b) Standard deviations across four weekly measurements.

Table 2. Detailed configurations for twelve request combinations. We anonymize the names of DoH resolvers to prevent potential misuse of the information. Chromium-based browsers include Microsoft Edge, Google Chrome, Opera, Brave, and others.

Request combination	Browser	HTTP method	DoH resolver
C01 / C07	Chromium-based / Firefox	POST	DoH Resolver A
C02 / C08			DoH Resolver B
C03 / C09			DoH Resolver C
C04 / C10		GET	DoH Resolver A
C05 / C11			DoH Resolver B
C06 / C12			DoH Resolver C

those countries, we spell out the names of the 14 countries with the highest DoH downgrade rates (when the average downgrade rates are considered), including Iran, China, Zimbabwe, etc., in the box in Figure 1(a). Second, low-intensity downgrades of DoH (e.g., where about 5% of DoH queries are downgraded) exist in the vantage points from the majority of the countries we test. The low-intensity downgrades of DoH queries on such a widespread scale are yet unreported in the prior work [4, 23], where they overestimated downgrade rates in most countries by including desirable downgrades occurring in enterprise/cloud networks.

We avoid making any definitive conclusions about the countries with low-intensity downgrades (e.g., whether these downgrades are intentional or accidental) as we cannot rule out the possibility of minor noises in our vantage point selection and country grouping. However, the presented result suggests that users in the majority of countries still benefit from DoH query privacy, while users in a few countries experience severe downgrades.

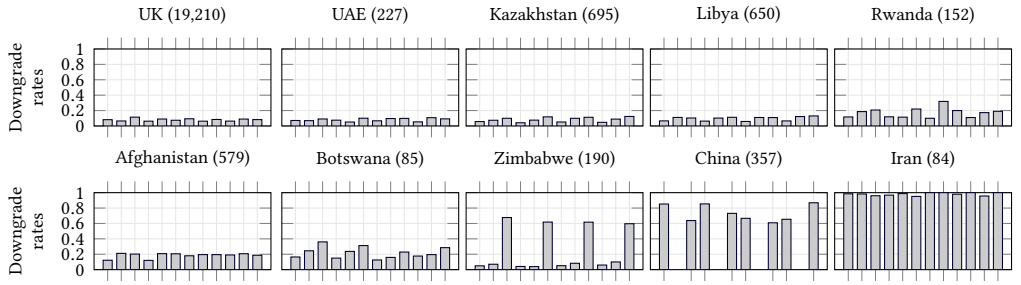

Fig. 2. DoH downgrade rates often vary by request combinations (indices for twelve combinations on the x-axis are ordered alphabetically but omitted for brevity) in ten selected countries. We show only eight combinations for China as we avoid the Google resolver for ethical concerns. Numbers next to the country names indicate the number of vantage points used in the measurement.

4.2 DoH Downgrades *by Request Combinations*

We investigate the DoH downgrade rates for different request combinations to understand the heterogeneity of current DoH downgrades. That is, vantage points at different countries experience widely varying degrees of downgrades for different request combinations.

A *request combination* represents a unique set of three variants in DoH requests: (1) browsers, (2) DoH resolvers, and (3) HTTP methods. For a single DoH query to be sent by a client host, a series of interactions need to exist between the client and a DoH resolver. For example, the DoH resolver's IP address may need to be resolved first, a TCP connection should be made, a TLS handshake should be completed, and then a DoH query can be finally sent to the DoH resolver; see these steps in detail in Figure 4.

Considering the major browsers' code base (i.e., Chromium-based browsers, Firefox), major DoH resolvers (i.e., Cloudflare, Quad9, Google), and two available HTTP methods (i.e., GET, POST), we define a total of twelve unique DoH request combinations. All these request combinations are implemented in our open-source tool. Table 2 summarizes the twelve request combinations and their detailed configurations. We then ask whether the current DoH downgrading networks show any difference in effectiveness when downgrading queries from these different request combinations.

Figure 2 shows that DoH downgrades vary for different combinations of DoH requests. For this in-depth analysis, we selected ten countries: the UK, UAE, Kazakhstan, Libya, Rwanda, Afghanistan, Botswana, Zimbabwe, China, and Iran. These countries exhibit the highest DoH downgrade rates throughout the weekly global measurements. Our request-combinations-based analysis in several countries reveals that the current DoH downgrades are *not homogeneous* within a country. Take Rwanda, Botswana, and China as examples. The results from these countries display heterogeneous patterns, where some request combinations experience significantly higher DoH downgrade rates than others. Zimbabwe is the most notable example, where eight request combinations experience only low-intensity downgrades while the other four combinations experience about 60% downgrade rates. In contrast, results from other countries (e.g., the UK, UAE, Kazakhstan, Libya, Afghanistan, and Iran) show smaller variations in the effectiveness of DoH downgrades for different request combinations.

This result suggests that the current DoH downgrades may not be designed and implemented to realize a *homogeneous* DoH downgrades within most countries. Instead, it appears that, at least for the time being, the DoH downgrades have been conducted in a rather *ad hoc* manner.

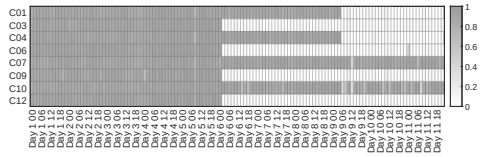

Fig. 3. DoH downgrade rates in China, measured every two hours.

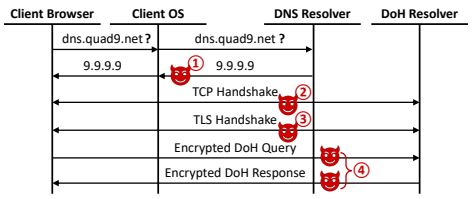

Fig. 4. Four major steps in the DoH resolution process and messages exchanged between a DoH client and resolvers at each step.

4.3 DoH Downgrades *by Times of Measurement*

We perform repeated measurements of DoH downgrades to evaluate whether the current DoH downgrade landscape in the wild is reliable over time. We have two types of time-series measurements that are designed with varying degrees of time intervals between measurements and the target countries: (1) 4-week-long, weekly repeated measurements for all the countries we test and (2) 1-week-long, two-hourly repeated measurements for China.

Weekly repeated measurements. Figure 1(b) shows the standard deviation across four repeated measurements for each country. We observe that the DoH downgrade rates for the same request combination in the same country remain relatively stable over time. To be specific, DoH downgrade rates within the four consecutive weeks demonstrate only a small variance; i.e., the average standard deviation of the DoH downgrade rates is less than 0.009. Even when there exist some seemingly increasing or decreasing patterns in the DoH downgrade rates over time, they are all isolated cases and do not show any consistent patterns, except in China. Our in-depth analysis of the four-week measurements in China suggests that the case of China requires a finer-grained time-series measurement due to unexpected fluctuations in DoH downgrade rates, which we discuss next. Overall, this low variance in DoH downgrade rates over time suggests that the current DoH downgrade landscape in the wild is relatively stable over time, excluding China.

Two-hourly repeated measurements. DoH downgrade rates for some request combinations in China show significant fluctuations. Thus, we perform a finer-grained time-series measurement for China and analyze the phenomenon in more detail. Figure 3 shows the DoH downgrade rates for all eight request combinations (as we did not test the Google resolver in China due to ethical concerns) measured every two hours over eleven days. We make several notable observations from this figure. First, DoH downgrade rates do not vary much across the eight request combinations except for sudden changes. When there is no sudden change, the DoH downgrade rates for all request combinations remain relatively stable over time, showing no sign of any diurnal patterns or weekly patterns. Second, perhaps most notably, DoH downgrade rates for Request Combinations 3, 6, 9, and 12 show a sudden decline (from nearly 100% to almost 0%) around 02:00 local time on the sixth day. Similarly, DoH downgrade rates for Request Combinations 1 and 4 display a sharp drop around 06:00 local time on the ninth day. These sudden decreases persist for the remainder of the measurement period, except for a single instance of downgrading at 00:00 local time on the eleventh day. The affected request combinations involve using `DoH resolver C` with all browsers or `DoH resolver A` with Chromium-based browsers. While it is fundamentally challenging to draw a clear explanation for the abrupt decreases in DoH downgrade rates for these request combinations, the synchronicity of the events suggests the possibility of intentional changes or large-scale, temporary failures in the DoH downgrade policy with `DoH resolver A` and `DoH resolver C` in China.

4.4 DoH Downgrades *by Filtering Techniques*

Last, we investigate the types of filtering techniques used for the current DoH downgrades and how downgraders are currently utilizing them. Our results show that most downgraders employ multiple filtering techniques for DoH downgrades, and the types of filtering techniques used vary greatly across different locations (i.e., countries).

Prior studies [4, 23] have attempted to measure a similar phenomenon; however, the results are limited because of their simple, stateless probing methodology. Thus, when more than one filtering technique is used for DoH downgrades in sequence, the prior studies could detect the first filtering technique only but miss the rest of the filtering techniques that follow.

Let us first explain four possible types of filtering techniques for DoH downgrades, introduce our new probing methodologies for accurate measurement, and then present our findings.

Types of the filtering techniques. The four major steps where the DoH downgrade can be conducted are shown in Figure 4. Notice that a traditional, unencrypted DNS resolution is typically conducted as the very first step of the DoH resolution process. Following the DNS resolution, a TCP connection is established between the DoH client and resolver with the resolved IP address. Then, the DoH client conducts a TLS handshake with the DoH resolver to establish a secure communication channel. The last step is the actual DoH resolution, where the DoH client sends a DoH query to the DoH resolver and the DoH resolver sends a DoH response back to the DoH client.

Based on the above steps, we enumerate the following possible types of filtering techniques that can be used for DoH downgrades (see the circled numbers, ①–④, in Figure 4):

- *DNS filtering* ①: A DoH downgrader can read the plaintext domain names in the plaintext DNS queries/responses to prevent the client from obtaining the IP address of the DoH resolver.
- *IP filtering* ②: The DoH downgrader can block any packets destined for the DoH resolver's IP address.
- *Hostname filtering* ③: The DoH downgrader can see the Server Name Indication (SNI) field in the TLS client hello message and prevent the TLS session establishment, based on the hostname.
- *DoH message filtering* ④: The DoH downgrader can identify and block the encrypted DoH queries/responses.

Notice that the first three filtering techniques require simple match-and-drop operations as long as the match rules are known to the DoH downgrader. On the other hand, the last filtering technique requires deep packet inspection (DPI) of encrypted HTTPS traffic to distinguish DoH queries/responses from other encrypted web traffic. There have been a number of research works [10, 16, 45] that show the feasibility of DPI-based DoH downgrades.

Stateful and proactive measurement. To find exactly which filtering techniques are used for the current DoH downgrades even when multiple filtering techniques are used in sequence, we devise a *stateful and proactive* probing method to test the existence of one or more filtering techniques. We test the existence of any filtering techniques by attempting to make a DoH query to a DoH resolver. When we discover that a vantage point experiences DoH downgrades, we investigate the filtering techniques used. Depending on the results of the existence tests, we further test the existence of the other, possibly layered filtering techniques.

First, we test the types of filtering techniques used for the current DoH downgrades by attempting to bypass each filtering technique proactively. If our earlier DoH query is downgraded but not downgraded after the bypass attempt, we can conclude that the corresponding filtering technique is indeed used. For this test, we implement bypass methods for the first three filtering techniques; we defer the details of our bypass methods to Section 5. We do not implement a bypass method for

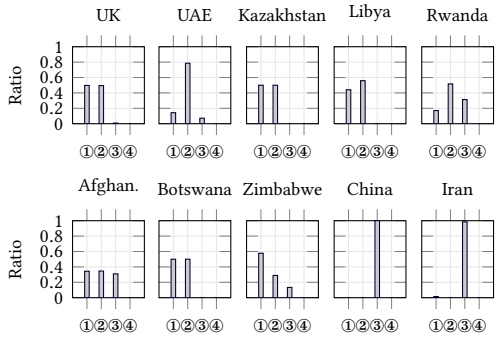

Fig. 5. Filtering profiles of ten selected countries. ① for DNS, ② for IP, ③ for hostname, and ④ for DoH message filtering.

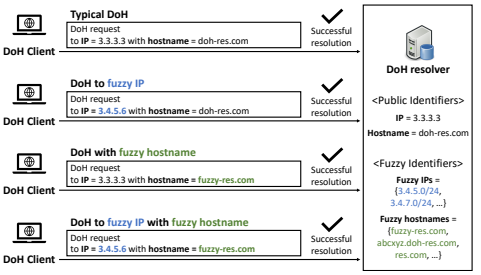

Fig. 6. Illustration of the DoH resilient resolution mechanisms that successfully bypass existing DoH downgrades.

the last filtering technique because we never have to; i.e., we never observe the existence of the fourth filtering technique in any of the measurement probes.

Our proactive probing method is already more sophisticated and accurate than the prior work [4, 23]; yet, we further enhance the accuracy of our measurement by making stateful probing. That is, when we are unable to confirm the existence of a filtering technique (i.e., when we fail to bypass the deployed filtering techniques) from a vantage point, we employ a combination of multiple bypass techniques. This approach allows us to test for the presence of layered filtering techniques, which is in contrast to the prior work [4, 23] that could only identify the initially activated filtering technique. For this, we maintain the *state* of the existence of each filtering technique for each vantage point and continue to test the existence of the other filtering techniques until we find all the filtering techniques used for the current DoH downgrades.

Filtering profiles of countries. Our measurements show that users in most countries we test experience a *combination* of two or more filtering techniques, and the ratio of employed filtering techniques differs depending on the countries where queries are made from. We call such a combination of filtering techniques the *filtering profile* of a country. This is one yet undisclosed important observation about DoH downgrades in the wild.

Figure 5 shows the filtering profiles of the same ten selected countries as in Figure 2. The analysis of the filtering profiles is conducted exclusively with `DoH resolver A`, because bypassing certain filtering techniques is feasible only with `DoH resolver A`. The first three filtering techniques are used in various combinations across the countries we illustrate in the figure. For example, in China and Iran, hostname filtering is extensively (and almost exclusively) employed for DoH downgrades, while a combination of DNS and IP filtering is utilized in the UK, Kazakhstan, Libya, and Botswana. Another interesting observation is that DoH message filtering is not used in any of the countries we test, possibly due to the high cost of per-packet analysis of encrypted TLS packets.

5 Bypassing DoH Downgrades

After understanding that DoH downgrades are conducted in various ways, we now investigate whether one can bypass each of the filtering techniques we examined earlier.

In the following, we review the three filtering techniques and suggest their corresponding bypasses:

- *To bypass DNS filtering:* ***hard-coding the resolver IP addresses.*** When downgraders see and block the very first plaintext DNS queries (such as cloudflare-dns.com), one simple but effective bypass (also suggested by Huang et al. [27]) is to hard-code the IP addresses of DoH resolvers in the client browsers.
- *To bypass IP filtering:* ***sending queries to unblocked IPs.*** When downgraders block any packets going to DoH resolvers based on their IP addresses, one possible bypass is to use some other IP addresses that are not yet blocked. This bypass works only if packets destined for those IP addresses are successfully redirected to the intended DoH resolvers.
- *To bypass hostname filtering:* ***sending queries with unblocked hostnames.*** When downgraders disrupt DoH based on the hostname written in the plaintext SNI field, one possible bypass is to use some other hostnames that are not yet blocked. This bypass works when packets with those hostnames are successfully redirected to the intended DoH resolvers.

While the practical implementation of the first bypass is straightforward, the second and third are non-trivial because they may require non-conventional network-level packet handling, and support from the DoH resolvers might be necessary. One contribution of this work is to show that the second and third bypasses are indeed possible for the existing DoH resolvers to a great extent.

Let us first explain in detail how one can send DoH requests to unblocked IP addresses and with unblocked hostnames. We call this very phenomenon of allowing DoH resolution with inaccurate IP addresses and hostnames *resilient resolution*, as DoH resolution is conducted with the indicators (i.e., IP addresses and hostnames) that are not specifically intended or lesser-known for the DoH resolution service. Similarly, we refer to such IP addresses and hostnames as *fuzzy IP addresses* and *fuzzy hostnames*, respectively. Figure 6 illustrates several examples of fuzzy IPs and fuzzy hostnames.

5.1 Resilient Resolution Mechanisms

A DoH resolver is said to support resilient resolution if it can resolve a DoH request solely based on the (encrypted) DoH query payload *without* being disrupted by the incorrect IP address or the inaccurate SNI field (i.e., hostname) used to make the request. As the DoH resolver with resilient resolution can reliably resolve DoH requests even with the wrong IPs or inaccurate hostnames, one can resolve DoH requests with IPs and hostnames that are not yet blocked. This would enable us to design bypass strategies for IP and hostname filtering.

One important question is *whether such resilient resolution mechanisms exist in the wild.* From our experiments, we found that some DoH resolvers do support resilient resolution on the Internet. Note that we focus on the observed functional properties of resilient resolution mechanisms in the existing DoH resolvers as the inner workings and hidden motives behind the functionality are challenging for an outsider to understand.

The implementations of resilient resolution may have two different features: support for *fuzzy IPs* and support for *fuzzy hostnames.*

5.1.1 Fuzzy IP Addresses. We call the IP addresses that are not intended for DoH resolution but can be used for that purpose as *fuzzy IP addresses.* The existence of fuzzy IP addresses is observed at least in one major DoH resolver, which we anonymize as `DoH resolver A`. Our experiment shows that DoH resolution requests to almost any IP address of `DoH resolver A`'s certain subnets are resolved successfully. For this experiment, we modify the IP addresses to send requests and confirm that 87,385 unique IPv4 addresses (5.73% of the total 1,524,706 IP addresses) of `DoH resolver A` are fuzzy IP addresses. Note that all our tests for fuzzy IP identification are conducted with the advertised hostname and the correct DoH query payload.

Interestingly, these fuzzy IP addresses are not merely some unused IP blocks of `DoH resolver A`. They are often IP addresses of actively used domains with popular web services that are completely unrelated to DoH resolution. Notable cases include a national major airline company and a major online bank in Asia.

The exact internal mechanism of fuzzy IP resolution is proprietary to individual operators of DoH resolvers and is therefore not fully transparent. One plausible explanation is that `DoH resolver A` might be reading the SNI field of the incoming packets and redirecting them to its DoH server based primarily on the hostname, rather than the associated IP addresses.

5.1.2 Fuzzy Hostnames. We call the hostnames that are *not* accurate but still usable for DoH resolution as *fuzzy hostnames*. According to RFC 6066 [1], which defines the SNI extension, the only supported server names are DNS hostnames. However, our tests confirm that certain inaccurate hostnames are accepted by all the major DoH resolvers we test and can therefore be used for DoH resolution. For instance, most DoH resolvers allow the use of slightly altered domain names (e.g., a domain like 3333.resolver instead of dns.resolver) to establish TLS connections with them. Most of the fuzzy hostnames we identify are listed in the certificates of DoH resolvers, while some are not found in them. Based on our experiments, we identify 12 fuzzy hostnames across the three major DoH resolvers we examined, including extensible hostnames: five of these fuzzy hostnames encompass any subdomain of them (e.g., abc.fuzzy.com, xyz.fuzzy.com, etc.). Additionally, we test and confirm that all the major DoH resolvers correctly resolve requests when the SNI extension is not used at all, which we consider as another form of fuzzy hostname.

5.1.3 Related Ideas. While the idea of our resilient resolution of DoH is new, several related ideas deserve some clarification.

Encrypted SNI. Several proposals, e.g., ESNI [40] and ECH [41], suggest encrypting the plaintext SNI field in the TLS handshake. While they may be used to hide the actual hostname, they cannot be used to bypass DoH downgrades because downgraders can easily block the ESNI or ECH traffic if they want to. In fact, the ESNI and ECH traffic are already blocked in some countries, such as China [9, 13], making them unusable for bypassing DoH downgrades at all.

Anti-censorship techniques. While our resilient resolution and general anti-censorship techniques share some similarities, they are different in scope and purpose. Since the use of DoH is not illegal in any state, to the best of our knowledge, the use of resilient resolution is not illegal either. The design of resilient resolution is not to circumvent censorship but to ensure reliable DoH resolutions. In contrast, the primary objective of anti-censorship techniques is to circumvent censorship, which is typically stronger and more challenging to bypass. Thus, anti-censorship techniques such as domain fronting [18] and refraction networking [25, 46, 47] would be an overkill for bypassing DoH downgrades.

5.2 Results

By combining the three effective bypass methods, we have achieved nearly 0% DoH downgrade rates in all the countries we test. For the bypass experiments, we select 19 countries that demonstrate more than 5% downgrade rates during the four consecutive global measurements. Upon implementing our bypass techniques, we achieved 0% downgrade rates in 12 countries, resulting in an average downgrade rate of 0.222% with a standard deviation of 0.36%. Our bypass strategies demonstrate consistently high effectiveness, even when confronted with heavily DoH-downgrading networks.

6 Limitations

In this section, we summarize our study's limitations. We did our best to minimize possible errors, but we acknowledge that the following limitations might introduce some errors despite our efforts.

First, the distinction between desirable and undesirable DoH downgrades is inherently blurry. Despite our best effort to focus solely on undesirable downgrades by utilizing RESIPs and excluding vantage points residing in enterprise, cloud, or education networks (see Section 3.2), we cannot guarantee that our measurements are entirely free of desirable downgrades. In other words, due to the lack of ground truth, our analysis and conclusions are based on carefully designed heuristics. For instance, some of our vantage points might still be located in such networks if the corresponding AS serves both residential and business users, even after manual sanitization. Nonetheless, despite these potential inaccuracies, our study has successfully revealed the previously unknown landscape of DoH downgrades — that is, the long-tail distribution of DoH downgrades.

For analysis, we utilized various mappings (e.g., from ASNs to organizations) included in public datasets, which might contain inaccuracies. As a result, our datasets could also reflect this misinformation, potentially introducing some errors.

Our bypass experiments were conducted against DoH downgrading techniques we observed. Therefore, our bypass strategies might not be as effective as demonstrated (average downgrade rate of 0.222%) if a more sophisticated downgrading techniques are used in the future; e.g., DoH message filtering [10, 16, 45].

Finally, we rely on a single platform [38] for the vantage points in our measurements. Conducting similar measurements on different platforms may further enhance our understanding of global DoH downgrade practices.

7 Conclusion

Precise measurement of the undesirable use of security and privacy protocols on a large scale is critical for the design and improvement of the Internet's security functions, such as the query privacy in domain name resolution. This measurement study has expanded our understanding of the current status of the widely used DoH protocol, the de facto privacy-enhancing protocol for DNS. Our findings indicate the prevalence of DoH downgrades affecting average Internet users, the existence of multi-layered DoH downgrading techniques, and the development of effective bypass schemes. These insights highlight the clear need for a more enforceable and usable DNS encryption protocol. We hope to continue the discussion and efforts towards more privacy-preserving domain name resolution systems.

Acknowledgments

This work is funded by the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (MSIT) under grant RS-2024-00464269.

References

[1] Donald E. Eastlake 3rd. 2011. Transport Layer Security (TLS) Extensions: Extension Definitions. RFC 6066. https://doi.org/10.17487/RFC6066

[2] David Adrian, Karthikeyan Bhargavan, Zakir Durumeric, Pierrick Gaudry, Matthew Green, J. Alex Halderman, Nadia Heninger, Drew Springall, Emmanuel Thomé, Luke Valenta, Benjamin VanderSloot, Eric Wustrow, Santiago Zanella-Béguelin, and Paul Zimmermann. 2015. Imperfect Forward Secrecy: How Diffie-Hellman Fails in Practice. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (Denver, CO, USA) *(CCS '15)*. ACM, New York, NY, USA, 5–17. https://doi.org/10.1145/2810103.2813707

[3] APNIC. 2024. Visible ASNs: Customer Populations (Est.). Retrieved April, 2024 from https://stats.labs.apnic.net/aspop

[4] Simone Basso. 2021. Measuring DoT/DoH blocking using OONI probe: a preliminary study. In *Proceedings of the 2021 NDSS DNS Privacy Workshop* (Virtual Event) *(DNSPRIV '21)*. The Internet Society, Reston, VA, USA, 10 pages. https://www.ndss-symposium.org/ndss-paper/auto-draft-123/

[5] Simone Basso. 2022. *ts-028-dnscheck*. OONI. Retrieved April, 2024 from https://github.com/ooni/spec/blob/master/nettests/ts-028-dnscheck.md

[6] Daniel Julius Bernstein. 2009. *DNSCurve: Usable security for DNS.* DNSCurve. Retrieved August, 2023 from https://dnscurve.org

[7] Benjamin Beurdouche, Karthikeyan Bhargavan, Antoine Delignat-Lavaud, Cédric Fournet, Markulf Kohlweiss, Alfredo Pironti, Pierre-Yves Strub, and Jean Karim Zinzindohoue. 2015. A Messy State of the Union: Taming the Composite State Machines of TLS. In *Proceedings of the 2015 IEEE Symposium on Security and Privacy* (San Jose, CA, USA) *(S&P '15)*. IEEE, New York, NY, USA, 535–552. https://doi.org/10.1109/SP.2015.39

[8] Protick Bhowmick, Md. Ishtiaq Ashiq, Casey Deccio, and Taejoong Chung. 2023. TTL Violation of DNS Resolvers in the Wild. In *Proceedings of the 24th International Conference on Passive and Active Network Measurement* (Virtual Event) *(PAM '23)*. Springer, Cham, Switzerland, 550–563. https://doi.org/10.1007/978-3-031-28486-1_23

[9] Kevin Bock, iyouport, Anonymous, Louis-Henri Merino, David Fifield, Amir Houmansadr, and Dave Levin. 2020. *Exposing and Circumventing China's Censorship of ESNI.* Technical Report. Geneva.

[10] Lionel F Gonzalez Casanova and Po-Chiang Lin. 2021. Generalized Classification of DNS over HTTPS Traffic with Deep Learning. In *Proceedings of the 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference* (Virtual Event) *(APSIPA ASC '21)*. IEEE, New York, NY, USA, 1903–1907. https://ieeexplore.ieee.org/document/9689667

[11] Rishabh Chhabra, Paul Murley, Deepak Kumar, Michael Bailey, and Gang Wang. 2021. Measuring DNS-over-HTTPS performance around the world. In *Proceedings of the 21st ACM Internet Measurement Conference* (Virtual Event) *(IMC '21)*. ACM, New York, NY, USA, 351–365. https://doi.org/10.1145/3487552.3487849

[12] The Chromium Projects. 2023. Chromium. Retrieved June, 2023 from https://www.chromium.org/Home/

[13] Catalin Cimpanu. 2020. *China is now blocking all encrypted HTTPS traffic that uses TLS 1.3 and ESNI.* ZDNet. Retrieved August, 2023 from https://www.zdnet.com/article/china-is-now-blocking-all-encrypted-https-traffic-using-tls-1-3-and-esni/

[14] Catalin Cimpanu. 2020. *DHS CISA to provide DoH and DoT servers for government use.* ZDNet. Retrieved March, 2024 from https://www.zdnet.com/article/dhs-cisa-to-provide-doh-and-dot-servers-for-government-use/

[15] Catalin Cimpanu. 2021. *NSA warns against using DoH inside enterprise networks.* ZDNet. Retrieved March, 2024 from https://www.zdnet.com/article/nsa-warns-against-using-doh-inside-enterprise-networks/

[16] Levente Csikor, Himanshu Singh, Min Suk Kang, and Dinil Mon Divakaran. 2021. Privacy of DNS-over-HTTPS: Requiem for a Dream?. In *Proceedings of the 2021 IEEE European Symposium on Security and Privacy* (Virtual Event, Vienna, Austria) *(EuroS&P '21)*. IEEE, New York, NY, USA, 252–271. https://ieeexplore.ieee.org/document/9581227

[17] Sara Dickinson. 2023. *Public Resolvers.* DNS Privacy Project. Retrieved June, 2023 from https://dnsprivacy.org/public_resolvers/

[18] David Fifield, Chang Lan, Rod Hynes, Percy Wegmann, and Vern Paxson. 2015. Blocking-resistant communication through domain fronting. In *Proceedings of the 15th Privacy Enhancing Technologies Symposium* (Philadelphia, PA, USA) *(PETS '15)*. De Gruyter, Berlin, Germany, 46–64. https://doi.org/10.1515/popets-2015-0009

[19] Arturo Filasto and Jacob Appelbaum. 2012. OONI: open observatory of network interference. In *Proceedings of the 2nd USENIX Workshop on Free and Open Communications on the Internet* (Bellevue, WA, USA) *(FOCI '12)*. USENIX Association, Berkeley, CA, USA, 8 pages. https://www.usenix.org/conference/foci12/workshop-program/presentation/filast%C3%B2

[20] Christian Grothoff, Matthias Wachs, Monika Ermert, and Jacob Appelbaum. 2015. *NSA's MORECOWBELL: Knell for DNS.* Technical Report. GNUnet e.V., Halle, Germany.

[21] Paul Grubbs, Arasu Arun, Ye Zhang, Joseph Bonneau, and Michael Walfish. 2022. Zero-Knowledge Middleboxes. In *Proceedings of the 31st USENIX Security Symposium* (Boston, MA, USA) *(USENIX Security '22)*. USENIX Association, Berkeley, CA, USA, 4255–4272. https://www.usenix.org/conference/usenixsecurity22/presentation/grubbs

[22] Saikat Guha and Paul Francis. 2007. Identity Trail: Covert Surveillance Using DNS. In *Proceedings of the 7th International Symposium on Privacy Enhancing Technologies* (Ottawa, Canada) *(PET '07)*. Springer, Berlin, Germany, 153–166. https://doi.org/10.1007/978-3-540-75551-7_10

[23] Nguyen Phong Hoang, Michalis Polychronakis, and Phillipa Gill. 2022. Measuring the Accessibility of Domain Name Encryption and Its Impact on Internet Filtering. In *Proceedings of the 23rd International Conference on Passive and Active Network Measurement* (Virtual Event) *(PAM '22)*. Springer, Cham, Switzerland, 518–536. https://doi.org/10.1007/978-3-030-98785-5_23

[24] Paul E. Hoffman and Patrick McManus. 2018. DNS Queries over HTTPS (DoH). RFC 8484. https://doi.org/10.17487/RFC8484

[25] Amir Houmansadr, Giang T.K. Nguyen, Matthew Caesar, and Nikita Borisov. 2011. Cirripede: circumvention infrastructure using router redirection with plausible deniability. In *Proceedings of the 18th ACM Conference on Computer and Communications Security* (Chicago, IL, USA) *(CCS '11)*. ACM, New York, NY, USA, 187–200. https://doi.org/10.1145/2046707.2046730

[26] Zi Hu, Liang Zhu, John Heidemann, Allison Mankin, Duane Wessels, and Paul E. Hoffman. 2016. Specification for DNS over Transport Layer Security (TLS). RFC 7858. https://doi.org/10.17487/RFC7858

[27] Qing Huang, Deliang Chang, and Zhou Li. 2020. A Comprehensive Study of DNS-over-HTTPS Downgrade Attack. In *Proceedings of the 10th USENIX Workshop on Free and Open Communications on the Internet* (Virtual Event) *(FOCI '20)*. USENIX Association, Berkeley, CA, USA, 8 pages. https://www.usenix.org/conference/foci20/presentation/huang

[28] The Intercept_. 2014. *The NSA and GCHQ's QUANTUMTHEORY Hacking Tactics*. The Intercept_. Retrieved June, 2023 from https://theintercept.com/document/nsa-gchqs-quantumtheory-hacking-tactics/

[29] IPinfo. 2024. ASN API. Retrieved April, 2024 from https://ipinfo.io/products/asn-api

[30] Lin Jin, Shuai Hao, Haining Wang, and Chase Cotton. 2021. Understanding the Impact of Encrypted DNS on Internet Censorship. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*. ACM, New York, NY, USA, 484–495. https://doi.org/10.1145/3442381.3450084

[31] Knot DNS. 2018. *kdig – Advanced DNS lookup utility*. Knot DNS. Retrieved March, 2024 from https://www.knot-dns.cz/docs/2.6/html/man_kdig.html

[32] Chaoyi Lu, Baojun Liu, Zhou Li, Shuang Hao, Haixin Duan, Mingming Zhang, Chunying Leng, Ying Liu, Zaifeng Zhang, and Jianping Wu. 2019. An End-to-End, Large-Scale Measurement of DNS-over-Encryption: How Far Have We Come?. In *Proceedings of the Internet Measurement Conference* (Amsterdam, Netherlands) *(IMC '19)*. ACM, New York, NY, USA, 22–35. https://doi.org/10.1145/3355369.3355580

[33] MaxMind. 2023. GeoLite2 Free Geolocation Data. Retrieved June, 2023 from https://dev.maxmind.com/geoip/geolite2-free-geolocation-data

[34] Mozilla Corporation. 2023. Firefox for Desktop. Retrieved June, 2023 from https://www.mozilla.org/en-US/firefox/new/

[35] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nick Weaver, and Vern Paxson. 2017. Global Measurement of DNS Manipulation. In *Proceedings of the 26th USENIX Security Symposium* (Vancouver, Canada) *(USENIX Security '17)*. USENIX Association, Berkeley, CA, USA, 307–323. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/pearce

[36] PeeringDB. 2024. The Interconnection Database. Retrieved April, 2024 from https://www.peeringdb.com/

[37] Proxyrack. 2023. Become a Peer. Retrieved June, 2023 from https://www.proxyrack.com/become-a-peer/

[38] Proxyrack. 2023. Residential Proxies | Proxyrack. Retrieved May, 2023 from https://www.proxyrack.com/residential-proxies/

[39] Yuqi Qiu, Baiyang Li, Zhiqian Li, Liang Jiao, Yujia Zhu, and Qingyun Liu. 2023. Before Toasters Rise Up: A View into the Emerging DoH Resolver's Deployment Risk. In *Proceedings of the 2023 IEEE Symposium on Computers and Communications* (Tunis, Tunisia) *(ISCC '23)*. IEEE, New York, NY, USA, 1149–1155. https://doi.org/10.1109/ISCC58397.2023.10217976

[40] Eric Rescorla, Kazuho Oku, Nick Sullivan, and Christopher A. Wood. 2020. *Encrypted Server Name Indication for TLS 1.3*. Internet-Draft draft-ietf-tls-esni-06. Internet Engineering Task Force. https://datatracker.ietf.org/doc/draft-ietf-tls-esni/06/ Expired.

[41] Eric Rescorla, Kazuho Oku, Nick Sullivan, and Christopher A. Wood. 2023. *TLS Encrypted Client Hello*. Internet-Draft draft-ietf-tls-esni-17. Internet Engineering Task Force. https://datatracker.ietf.org/doc/draft-ietf-tls-esni/17/ Work in Progress.

[42] Yuval Shavitt and Noa Zilberman. 2011. A Geolocation Databases Study. *IEEE Journal on Selected Areas in Communications* 29, 10 (2011), 2044–2056. https://doi.org/10.1109/JSAC.2011.111214

[43] Sooel Son and Vitaly Shmatikov. 2010. The Hitchhiker's Guide to DNS Cache Poisoning. In *Proceedings of the 6th International ICST Conference on Security and Privacy in Communication Systems* (Singapore, Singapore) *(SecureComm '10)*. Springer, Berlin, Germany, 466–483. https://doi.org/10.1007/978-3-642-16161-2_27

[44] Mathy Vanhoef and Eyal Ronen. 2020. Dragonblood: Analyzing the Dragonfly Handshake of WPA3 and EAP-pwd. In *Proceedings of the 2020 IEEE Symposium on Security and Privacy* (Virtual Event) *(S&P '20)*. IEEE, New York, NY, USA, 517–533. https://doi.org/10.1109/SP40000.2020.00031

[45] Dmitrii Vekshin, Karel Hynek, and Tomas Cejka. 2020. DoH Insight: detecting DNS over HTTPS by machine learning. In *Proceedings of the 15th International Conference on Availability, Reliability and Security* (Virtual Event, Ireland) *(ARES '20)*. ACM, New York, NY, USA, Article 87, 8 pages. https://doi.org/10.1145/3407023.3409192

[46] Eric Wustrow, Colleen M. Swanson, and J. Alex Halderman. 2014. TapDance: End-to-Middle Anticensorship without Flow Blocking. In *Proceedings of the 23rd USENIX Security Symposium* (San Diego, CA, USA) *(USENIX Security '14)*. USENIX Association, Berkeley, CA, USA, 159–174. https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/wustrow

[47] Eric Wustrow, Scott Wolchok, Ian Goldberg, and J. Alex Halderman. 2011. Telex: Anticensorship in the Network Infrastructure. In *Proceedings of the 20th USENIX Security Symposium* (San Francisco, CA, USA) *(USENIX Security '11)*. USENIX Association, Berkeley, CA, USA, 30. https://www.usenix.org/conference/usenix-security-11/telex-anticensorship-network-infrastructure

[48] Liang Zhu, Zi Hu, John Heidemann, Duane Wessels, Allison Mankin, and Nikita Somaiya. 2015. Connection-oriented DNS to Improve Privacy and Security. In *Proceedings of the 2015 IEEE Symposium on Security and Privacy* (San Jose, CA,

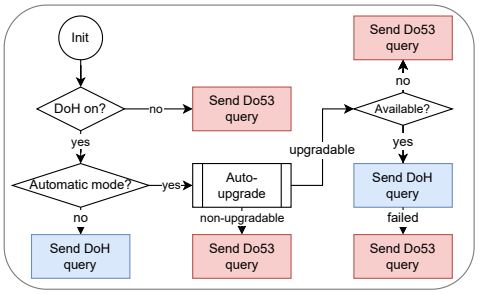

Fig. 7. Flow chart illustrating the implementation of DoH in Chromium. The secure mode ensures that the browser always sends DoH queries without falling back to plaintext DNS (Do53 in the diagram), while the automatic mode adopts opportunistic approaches.

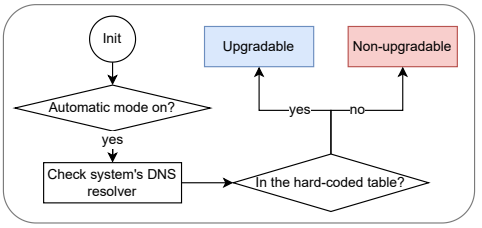

Fig. 8. Auto-upgrade process in Chromium. To upgrade the DNS resolver to a DoH resolver, Chromium maintains its table of DNS resolvers supporting DoH resolution.

USA) *(S&P '15)*. IEEE, New York, NY, USA, 171–186. https://doi.org/10.1109/SP.2015.18

A Detailed Analysis of Browsers

A.1 Justification of Browser Analysis

Instead of users directly constructing DoH packets, it is the responsibility of the browsers to handle this task on their behalf. Therefore, by inspecting the source code and monitoring the flow of DoH packets in selected browsers, we can gain insights into their DoH implementation and replicate their behavior in our experiments. This approach enables us to achieve a higher level of accuracy and realism in our investigations.

The browsers that have been inspected for this investigation can be classified into two categories: Chromium-based web browsers and the Mozilla Firefox web browser. These two categories have been selected to represent a significant portion of the browser market. It is important to note that Safari, despite its widespread usage globally, has been intentionally excluded from the analysis. This decision is based on the fact that Safari does not natively support DoH within the browser. Instead, DoH is configured and implemented at the operating system level. Given this study's specific focus on browsers that directly incorporate DoH functionality within their codebase, Safari is excluded from the analysis.

A.1.1 Chromium-Based Web Browsers. Chromium is an open-source web browser actively developed and maintained by the Chromium project [12]. Chromium is the foundation for numerous widely used browsers, including Google Chrome, Opera, Brave, and Avast Secure Browser. Given the widespread adoption of Chromium as the underlying framework for these browsers, our investigation begins by focusing on Chromium itself. We carefully analyze its source code to gain insights into Chromium's implementation of DoH. This analysis forms the baseline for comparing and contrasting the DoH packets generated by other Chromium-based web browsers. By conducting such comparisons, we aim to identify any variations or unique characteristics in the DoH packets generated by different Chromium-based browsers. To visually illustrate the implementation of DoH in Chromium, we provide a flow chart in Figure 7.

Modes of DoH in Chromium. The implementation of DoH in Chromium can be described in detail, focusing on its three modes: Off, Automatic, and Secure. It is important to note that users do

not directly determine the mode, as it is automatically set by the browser based on preferences. The following are the specifics of each mode.

- Off: In this mode, DoH is not activated. The browser does not utilize DoH for DNS resolution.
- Automatic: This mode follows the system configurations for DNS resolution. It employs a process called *auto-upgrade* to determine whether DoH should be used. If DoH is not available or encounters any issues, Chromium falls back to plaintext DNS. Additionally, Chromium performs *DoH probes* to check the status of the configured DoH resolver.
- Secure: In the Secure mode, Chromium utilizes a specified DoH resolver for DNS resolution. It exclusively relies on the DoH resolver and does not fall back to plaintext DoH, even if issues arise.

Respect local DNS resolvers. Chromium incorporates a unique process called *auto-upgrade* to ensure compatibility with local DNS resolvers while enabling DoH by default. As the name implies, this process aims to *upgrade* the system's existing DNS resolver, if feasible, by assessing its ability to support DoH resolution. The complete auto-upgrade process is illustrated in Figure 8.

Ensuring the reliability of DoH resolvers. Major browsers typically adopt opportunistic and best-effort approaches when it comes to DoH resolution. In line with this approach, Chromium incorporates its own algorithm, known as *DoH probes*, to verify the reliability of the currently used DoH resolver. The browser sends a DoH query to the configured resolver, explicitly requesting the IP address of www.gstatic.com. If the query succeeds and a valid response is received, the DoH resolver is marked as *available*. However, if the query fails for any reason, the browser retries the query until it reaches the maximum failure limit.

HTTP headers in DoH packets. Through extensive code inspection and packet monitoring, we have been able to identify various implementation details in Chromium, including the specific HTTP headers it includes in the DoH packets it sends. This level of analysis allows us to create and send DoH packets that closely resemble those constructed by Chromium.

However, it is important to acknowledge that there are certain limitations to our ability to fully replicate Chromium's behavior without developing the entire browser. Specifically, we may encounter challenges in accurately imitating the TLS/SSL-related details and the precise flow of packets, including subtle divisions of headers and data. These aspects of Chromium's implementation are intricately tied to the browser's underlying architecture and codebase, making them difficult to replicate in isolation.

After careful observation and analysis of packets from various Chromium-based web browsers, including Google Chrome, Microsoft Edge, Opera, Brave, Avast Secure Browser, Vivaldi, Epic Privacy Browser, Ungoogled Chromium, and Yandex, we have determined that these browsers exhibit the exact same behavior and implement DoH in identical ways. However, we have observed that Opera has a slight difference in its behavior, specifically related to the auto-upgrade feature, which does not work properly. Nevertheless, since our measurement does not involve the auto-upgrade procedure, we conclude that every Chromium-based web browser we have analyzed behaves in the same manner, including Opera.

A.1.2 Mozilla Firefox. Mozilla Firefox, often referred to as Firefox, is an open-source and free web browser developed and maintained by the Mozilla Foundation and Mozilla Corporation [34]. To gain a comprehensive understanding of Firefox's implementation of DoH, we conduct a thorough inspection of its source code and closely monitor its behavior through packet analysis. A detailed flow chart illustrating the implementation of DoH in Firefox is provided in Figure 9.

Modes of DoH in Firefox. Firefox introduces an additional mode compared to Chromium, resulting in a total of four modes. Despite the difference in the number of modes, their functionalities

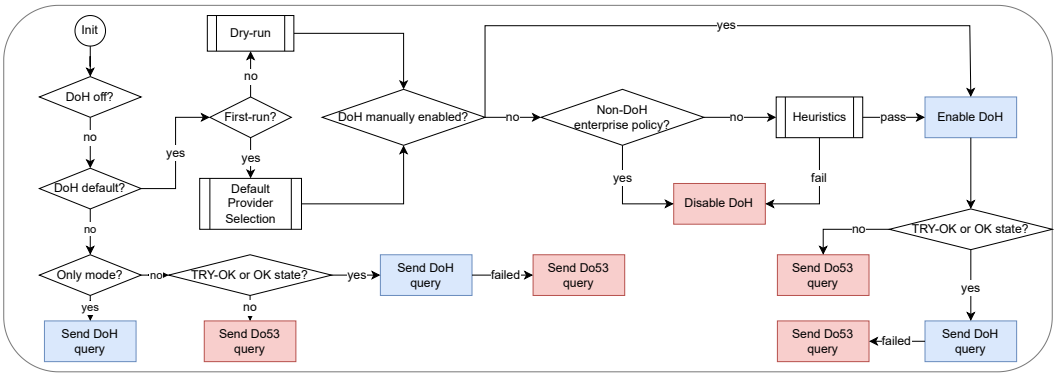

Fig. 9. Flow chart illustrating the implementation of DoH in Firefox. The only mode guarantees that the browser always sends DoH queries without falling back to plaintext DNS (Do53 in the diagram), while the first mode employs opportunistic approaches. Additionally, if DoH is enabled by default without the users' intention, the browser follows a specific logic to prioritize the local DNS resolvers.

are nearly identical. However, it is worth noting that Firefox provides users with the flexibility to manually set the mode of DoH, while also offering automatic mode selection based on user preferences.

- Off: DoH is not activated, automatically.
- First: Firefox utilizes the default DoH resolver, which is set through a special process known as *default provider selection*, or a specified DoH resolver if one is configured. In this mode, Firefox may fall back to plaintext DNS and verify the status of the DoH resolver through a process called *confirmation*.
- Only: Firefox exclusively resolves domain names using the default or specified DoH resolver and does not fall back to plaintext DNS.
- Off by choice: DoH is *manually* deactivated, allowing users to opt-out of its usage.

Default provider selection and dry-run. To determine the optimal DoH resolver for the user from the list of DoH resolvers maintained by Firefox, a comparison of resolver speeds is performed using a series of DoH requests. This process, known as a *dry-run*, is performed each time the browser is launched. The outcome of the dry-run is saved and subsequently utilized during the *default provider selection* process to determine the DoH resolver to be used.

Respect local DNS resolvers. Firefox takes a different approach to accommodate residential DNS resolvers. Instead of upgrading the system's DNS resolver like Chromium, Firefox maintains its own list of DoH resolvers. This approach necessitates stricter compatibility policies. As a result, Firefox chose not to enable DoH by default in many countries. In certain countries where DoH is enabled by default, Firefox employs a special process called *heuristics* considering local requirements.

Ensuring the reliability of DoH resolvers. Similar to Chromium's DoH probes process, Firefox implements its own algorithm called *confirmation* to ensure the reliability of the DoH resolver being used. The confirmation process in Firefox is more intricate, involving four distinct states and incorporating periodicity through a timer. The process starts in the *TRY-OK state* and transitions to the *OK state* if the test query is successful. However, if the test query fails, it transitions to the *FAIL state*. While in the *TRY-OK state* or *OK state*, Firefox utilizes DoH for name resolution. Conversely, if it is in the *TRY-FAIL state* or *FAIL state*, the browser avoids using DoH for name resolution. Transitions between states can occur by timer alarms or changes in network status.

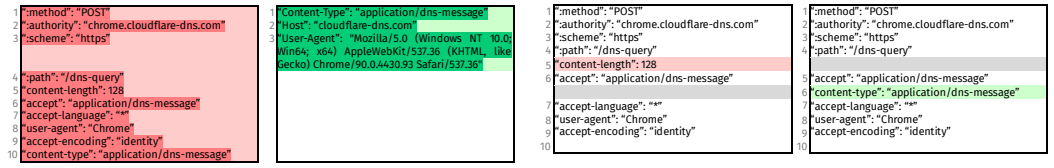

Fig. 10. Differences between HTTP headers included in DoH queries. The first and third boxes display the headers sent by Chromium, while the second and fourth show the headers sent by OONI and us, respectively.

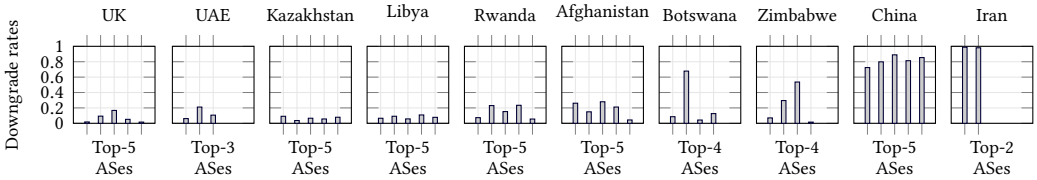

Fig. 11. DoH downgrade rates tend to vary depending on the geolocation of DoH query clients. We show the top 5 ASes (if possible) in each country on the x-axis, with their names redacted.

HTTP headers in DoH packets. During our meticulous analysis, we have successfully identified various implementation details in Firefox, including the HTTP headers present in DoH packets.

With the knowledge acquired about Firefox's DoH implementation, we are able to generate DoH packets that closely resemble those sent by Firefox. Still, certain limitations persist, such as the inability to replicate TLS/SSL-related details, as we are not developing the entire browser.

A.2 Comparison with OONI

To emphasize the significance of our browser analysis, we compare the HTTP headers in DoH queries. Figure 10 illustrates the disparity between Chromium and OONI's measurement, as well as the variance between Chromium and our measurement. OONI's query contains entirely distinct HTTP headers, while ours contains the same headers but in a different order.

B Detailed Analysis of DoH Downgrades

B.1 DoH Downgrades *by Geolocations*

We ask whether the current DoH downgrades are affected by the geolocation of the DoH query clients. We use an ASN as the minimum unit of geolocation for a DoH query client, obtained using the MaxMind GeoLite2 database [33].

Figure 11 shows how downgrade rates vary by the geolocation of DoH query clients in the same ten selected countries. We present the top five ASes (if possible) in each country by decreasing number of users, showing only those with more than 40 vantage points. The results reveal that DoH downgrade effectiveness varies by the geolocation of DoH query clients. For instance, in Botswana, the second-largest AS downgrades more than 60% of DoH queries, whereas others do not impose such severe downgrades. Similar disparities in DoH downgrade effectiveness are also observed in other countries.

Received June 2024; revised September 2024; accepted October 2024