Yesterday Once More: Global Measurement of Internet Traffic Shadowing Behaviors

Yunpeng Xing*
Tsinghua University
Beijing, China
xyp23@mails.tsinghua.edu.cn

Chaoyi Lu*
Tsinghua University
Beijing, China
luchaoyi@tsinghua.edu.cn

Baojun Liu
Tsinghua University
Beijing, China
lbj@tsinghua.edu.cn

Haixin Duan
Tsinghua University
Beijing, China
Quancheng Laboratory
Jinan, China
duanhx@tsinghua.edu.cn

Junzhe Sun
Tsinghua University
Beijing, China
sjz24@mails.tsinghua.edu.cn

Zhou Li
University of California, Irvine
Irvine, California, United States
zhou.li@uci.edu

Abstract

We present a global, large-scale measurement of Internet traffic shadowing, a less-studied yet covert format of on-path manipulation. As part of pervasive monitoring, data within packets is silently observed, retained, and then leveraged to produce additional, unsolicited requests. To depict the landscape of such behaviors, we generate a collection of decoy traffic that lures on-path exhibitors, spread them via 4,364 vantage points recruited from commercial VPN providers, and capture unsolicited requests triggered by them. We find traffic shadowing against DNS, HTTP, and TLS protocols; DNS queries to several public resolvers are most susceptible, by being observed on a wide range of Internet paths. Through hop-by-hop tracerouting, we find observers of DNS queries associated with destinations, while HTTP messages are mostly observed on the wire. User data can be retained for long, e.g., over 10 days, and can be leveraged for more than once. While a notable portion of unsolicited requests originate from addresses labeled by blocklists, we find most of them are performing reconnaissance, and we see no evidence of exploits attempted in the collected traffic.

CCS Concepts

• **Networks** → **Network measurement**; *Network monitoring*; • **Security and privacy** → **Network security**; • **Information systems**;

Keywords

Traffic Shadowing, Network Monitoring, Security Probing

ACM Reference Format:
Yunpeng Xing, Chaoyi Lu, Baojun Liu, Haixin Duan, Junzhe Sun, and Zhou Li. 2024. Yesterday Once More: Global Measurement of Internet Traffic Shadowing Behaviors. In *Proceedings of the 2024 ACM Internet Measurement Conference (IMC '24), November 4–6, 2024, Madrid, Spain.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3646547.3689023

*Both authors contributed equally to this research.

IMC '24, November 4–6, 2024, Madrid, Spain

ACM ISBN 979-8-4007-0592-2/24/11
https://doi.org/10.1145/3646547.3689023

1 Introduction

Traffic on the Internet has been subject to a wide array of manipulation. This paper looks into another more passive yet covert, less-studied format, termed as *traffic shadowing*: during transmission, packets (as a whole or data in selected fields) are sniffed and recorded by *on-path observers*; subsequently, they re-appear as additional, *unsolicited* requests when no clients are waiting for responses. Anecdotally, APNIC reported one in four DNS query names generated by their one-shot measurements was unexpectedly captured again, hours or even days after experiments concluded [35]. Considered as one possible format of pervasive monitoring (RFC 7258 [27]), traffic shadowing potentially poses privacy risks when exhibited without user awareness, as sensitive data can be silently retained and processed by parties other than its intended receivers (e.g., sniffing DNS query names enables user tracking [20, 62]).

We highlight that traffic shadowing differs from well-studied manipulation (e.g., censorship [30, 44, 52] and packet interception [16, 40, 59]) in that, communication between clients and servers is *not* tampered with. Clients initiate requests and *are* able to get authentic responses, otherwise than being blocked, discarded, or spoofed. Also note that other than purposes more sinister such as eavesdropping and surveillance [9, 19], traffic shadowing may also bear perfectly innocent purposes, e.g., serving as security features. For example, FireEye security appliances are known to perform additional scans for phishing websites from user traffic they monitor [43].

This paper presents a global, large-scale measurement of Internet traffic shadowing. We seek answers to: *i)* How broad is the landscape of traffic shadowing? *ii)* Where on-path does it occur? *iii)* What are the characteristics and incentives of different exhibitors? To this end, we generate a collection of *decoy* traffic (DNS, HTTP, and TLS messages for this pilot study) and spread them via a VPN-based measurement platform. Each decoy is sent only once and carries a unique domain name with decoy-specific identifiers, attempting to lure on-path observers. To capture unsolicited requests, honeypots are established in 3 different locations (US, DE, and SG). Our measurement platform recruits 4,364 vantage points (VPs) in

82 countries; from each VP we send decoys to a wide range of destination servers, including large public DNS resolvers, authoritative servers and Tranco top sites [49], such that our experiment covers sufficient Internet paths. Particularly, VPN-based VPs support varying the Time-to-Tive (TTL) values in the IP header of decoys, enabling us to locate traffic observers (i.e., at which hop of this client-server path does decoy trigger unsolicited requests). Finally, correlating unsolicited requests with triggering decoys, we perform behavioral analysis and summarize characteristics.

Our measurement is performed during Mar and Apr 2024 (2 months). We find DNS queries to large public resolvers are most susceptible: e.g., queries from 70% of global VPs to Yandex and OneDNS, and from 85% of CN VPs to 114DNS, trigger unsolicited requests after hours. A smaller portion (<10%) of HTTP (/TLS) decoys sent to Tranco Top sites are observed by devices on the wire. From temporal features, we find user data can be retained for long, sometimes leveraged multiple times by traffic shadowing exhibitors. For example, 40% of query names within DNS decoys sent to Yandex re-appear in unsolicited HTTP (/TLS) requests 10 days later; over 1 hour after emission, 51% of DNS decoys still produce over 3 unsolicited requests. From payloads of unsolicited requests, we find reconnaissance attempts, e.g., HTTP path enumeration. Alarmingly, 50% of them come from addresses labeled malicious by IP blocklists.

Contributions of this work include:

- We establish methodology and depict the global landscape of Internet traffic shadowing, a covert, less-studied format of on-path traffic manipulation.
- We perform behavioral analysis and report characteristics of different shadowing exhibitors.

2 Background and Related Work

Well-studied types of traffic manipulation. Studies have measured landscape and tactics of Internet censorship, either within global [15, 44, 48, 51] or per-country scopes [9, 33, 63, 66, 68]. Works also propose different censoring [65, 67] and evasion techniques [10, 17, 30, 61]. Others focus on packet interception of messages over selected protocols, including DNS [40, 53] and HTTP [16, 21, 59, 70]. Finally, messages can be directly tampered with on the wire, e.g., fields within IP and TCP layer [23, 24, 26, 32, 41, 57]. Works also find such tampering (e.g., TCP initial sequence number rewriting) poses adverse effects to protocol functions and scalability [31, 34, 60].

Traffic shadowing and potential exhibitors. As another type of less-studied manipulation, traffic shadowing appears more passive and covert. Data within user traffic is silently recorded by *on-path observers*; subsequently, additional, *unsolicited requests* (over the same or different protocols as original traffic) bearing such data are produced when no clients are waiting for the responses. A few cases of traffic shadowing have been reported. Within Tor, [19] injects decoy usernames and passwords (over FTP and Telnet) and finds Tor exit nodes sniffing the credentials, establishing unsolicited connections. APNIC also reported one in four DNS queries reaching their authoritative servers are repetitions of previous queries [35], unsolicited because the names were unique and only queried once for the experiment. While problematic resolver implementations (e.g., aggressively retrying) account for some, their blog presumably explains the remainder by stalking or tracking. Traffic shadowing has also been previously observed on network devices, e.g., FireEye security appliances [43] monitor and report back to the company selected URLs from user traffic, to which additional HTTP scans are scheduled from Internet proxies to detect phishing websites.

Note that by its definition (i.e., unsolicited requests bearing data of prior packets), traffic shadowing may be the results of diverse initiatives Other than surveillance [27], censorship [9], or security features as shown above, unsolicited requests may also be triggered by implementation choices (e.g., intentional retries) and performance enhancements (e.g., network connectivity monitoring tasks or automatic cache refreshing after expiration by destination server or local cache). While cases of traffic shadowing have been reported (e.g., by APNIC and researchers on the FireEye case), we scale the measurement of such behaviors to a broader scope by considering three critical protocols (DNS, HTTP, and TLS), attempting to locate observers over capabilities provided by VPN-based vantage points.

3 Methodology

Our methodology leverages a straightforward perception that if initiated only once from original clients, data embedded in user packets *should not re-appear* (because the original client is no longer waiting for a response) unless subject to traffic shadowing. Figure 1 overviews our methodology for detecting such behaviors. We first generate a collection of decoy traffic embedded with unique data, attempting to lure on-path observers. The decoys are then spread from a global VPN-based measurement platform to a wide range of destinations. Finally, if unsolicited requests are captured by our honeypots as a result of traffic shadowing, we may correlate them to the initial decoys and client-server paths.

Format of decoy traffic. We generate decoys over 3 protocols providing critical functions of the Internet: DNS, HTTP, and TLS (over port 443). They all contain fields that should be filled in with domain names (i.e., QNAME in DNS queries, `host` in headers of HTTP `GET` messages, and Server Name Indication in TLS `ClientHello` messages). We choose to lure traffic shadowing with domain names in these fields of decoys because they are in clear-text, exposed to on-path observers, and privacy-sensitive [62]. In addition, with this design, unsolicited requests can be diverted to our honeypots via DNS configuration. [1] The embedded domains are:

$$\underbrace{\text{g6d8jjkut5obc4-9982}}_{\text{identifier string (time, IP, TTL)}} \; . \; \underbrace{\text{www.experiment.domain}}_{\text{domain pointing to honeypots}}$$

The identifier string is decoy-specific, encoded from time of sent, source (i.e., VP) and destination addresses, and initial TTL in IP header (for tracerouting, see Phase II). We configure wildcard DNS records (TTL of DNS record=3,600) for all experiment domains to honeypots in US, DE, and SG.

Phase I: Finding paths subject to traffic shadowing. The first phase of the experiment spreads decoys from global vantage points

[1] Unsolicited requests are diverted to us because via wildcard DNS configuration, we resolve all experiment domains (which we use to lure traffic shadowing) to our honeypots. As a result, if a decoy is sniffed and unsolicited probes are performed against the embedded domain, they will arrive at our honeypots. [39] adopts a similar design to track certificate bots. For HTTP and TLS decoys, this design results in a mismatch between the host field and the destination address in the IP headers.

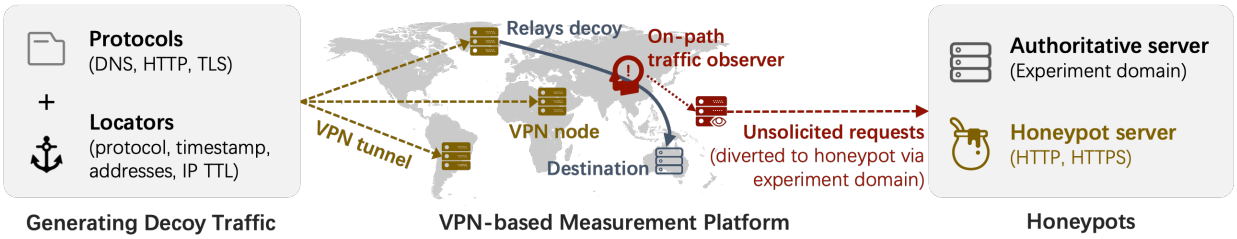

Figure 1: Overview of methodology

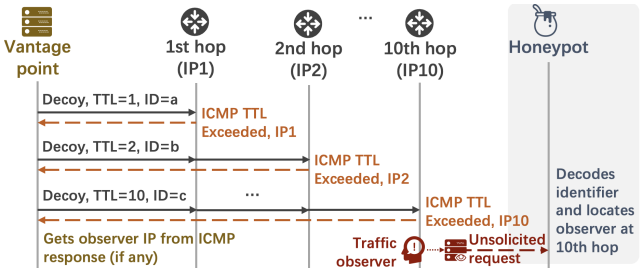

Figure 2: Locating on-path traffic observers hop-by-hop

(VPs, see below) to a wide range of destinations and finds client-server paths where decoys trigger unsolicited requests. For DNS decoys, we send them to primary addresses of 20 large public DNS services (e.g., Google and Cloudflare, see Appendix B), after consulting their use metrics [13]. We also include one self-built DNS resolver, 13 root servers, and 2 Top-Level Domain authoritative servers. For HTTP (`GET`) and TLS (`ClientHello`) decoys, on each VP we send them, after successful TCP handshakes, to IP addresses (2,325 in total, distributed in 234 ASes) behind Tranco top 1K sites [49]. We select IP addresses of popular services, instead of random addresses, as destinations of decoys, because we consider paths to popular servers more likely monitored by traffic shadowing exhibitors.

To identify unsolicited requests, for each arriving at our honeypots, we label its protocol combination: `Decoy-Request`. For example, an incoming HTTP request is labeled `DNS-HTTP` if it bears the unique experiment domain we embedded in a DNS decoy. An incoming request bearing decoy data is unsolicited if: *i)* request and decoy protocols are different (because we never send this particular data over the request protocol); or *ii)* request protocol is HTTP or TLS (because we never send HTTP or TLS decoys to our honeypots); or *iii)* request protocol is DNS, but the unique query name has appeared in another DNS query earlier (i.e., the initial decoy).

Phase II: Locating on-path traffic observers. After finding problematic client-server paths, the second phase of the experiment attempts to locate observers using a hop-by-hop traceroute technique, as also leveraged by one prior study to locate security devices monitoring traffic at finer granularity [43]. As shown by Figure 2, from the VP of a problematic path, we increase the initial TTL in the IP header of a decoy from 1 to 64 and send them consecutively. Note that changing TTL will result in a new identifier string in our experiment domain, such that our honeypots may decode and map to the exact decoy subject to traffic shadowing. If within one path, until we send decoy with initial TTL=t do we capture unsolicited requests bearing the same data, we conclude observers are t hops away from the VP. In addition, the ICMP TTL Exceeded messages returned to VPs expose IP addresses of observers (note that observers may not initiate unsolicited requests by themselves, hence cannot be revealed via source addresses of unsolicited requests). For HTTP (/TLS) decoys, we do not perform TCP handshakes with destinations before tracerouting, as this may keep the connection idle for long (i.e., until TTL is large enough for decoys to reach destination) and cause overhead to the servers.

Table 1: Capabilities of VPN measurement platform

	# Provider	IP	AS	Country/Province
Global (excl. CN)	6	2,179	74	81 (countries)
China (CN mainland)	13	2,185	47	30 (of 31 provinces)
Total	19	4,364	121	82 (countries)

VPN-based vantage points (VPs). We choose to build a new measurement platform, primarily because none of the existing, publicly-available platforms supports hop-by-hop traceroute over application protocols (required by Phase II; we present a survey of existing platforms in Appendix D). To address ethical concerns, we adopt the same strategy as ICLab [44], a well-established measurement platform built over *commercial VPN services*. As shown in Table 1, we recruit 2,179 VPs from 6 VPN providers with global accessibility. Additionally, we pay special attention to China, which has a large Internet population but lacks server nodes from global VPN providers. To add this perspective, we integrate another 2,185 VPs from 13 local providers in China. A particular note that because user-hosted (i.e., residential) nodes may pose ethical risks for their potential engagement in illicit activities [42, 69], we attempt to remove such providers from consideration and only recruit from datacenter VPN providers (see determination, statement, and listing of providers in Appendix C). [2] Finally, we do not use VP locations advertised by VPN providers, given they may be skewed [44]. Rather, we obtain VP addresses by directly establishing TCP connections from them to our honeypot and inspect the source addresses, then geo-locate them by looking them up in IP databases [11].

Comparison and limitations. Compared to works on specific cases of traffic shadowing (e.g., [43] detects FireEye shadowing by HTTP scanning from a single server), we send decoys from VPN platforms, enabling us to generate a broader set of decoys (over DNS, HTTP, TLS). Besides destination servers, our methodology may also

[2] The Chinese VPs can only be connected from measurement schedulers we establish in China. As a result, our VPN tunnels do not pass censorship devices located at international gateways [22], and thus are not intercepted.

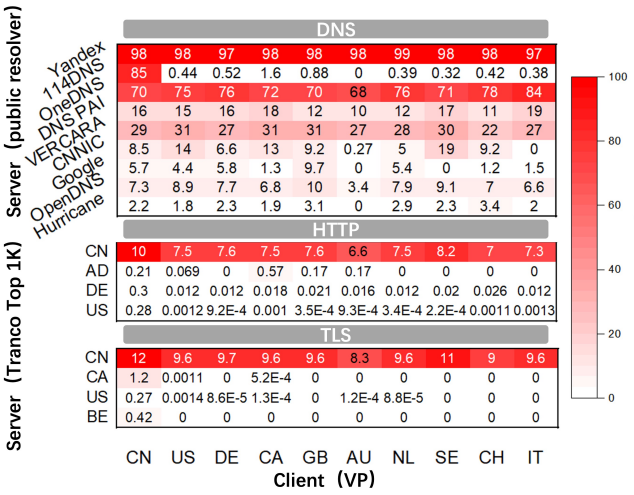

Figure 3: Ratio (%) of client-server paths subject to traffic shadowing behaviors

Table 2: Normalized location of traffic observers

Hops* from VP	1-2	3	4	5	6	7	8	9	10
DNS (% observers)	0.009	0.03	0.006	0.018	0.024	0.026	0.041	0.14	99.7
HTTP (%)	1.4	15	31	30	18	2.5	0	0.28	2.3
TLS (%)	0.14	0.68	0.78	1.1	26	6.0	0.15	0.013	65

* Hops are normalized into a scale of 1 to 10, where 10 indicates destination.

Table 3: Top networks of on-path traffic observers

DNS	AS203020 HostRoyale Technologies Pvt Ltd	4 (13%)
	AS4808 China Unicom Beijing Province Network	3 (10%)
	AS21859 Zenlayer Inc	3 (10%)
HTTP	AS4134 CHINANET-BACKBONE	172 (44%)
	AS58563 CHINANET Hubei province network	40 (10%)
	AS137697 CHINATELECOM JiangSu	24 (6.1%)
TLS	AS4134 CHINANET-BACKBONE	134 (54%)
	AS4812 China Telecom (Group)	16 (6.5%)
	AS23650 CHINANET jiangsu backbone	11 (4.5%)

locate traffic observers in ISP networks, given our VPs are globally distributed. That said, our decoys may not reflect real user traffic patterns due to the embedding of nonce locators, therefore we may only present a lower bound of traffic shadowing. We recruit VPs from datacenter VPN providers; while traffic shadowing (and other manipulation [64]) is presumably more prevalent within residential networks, we do not leverage user-hosted VPNs due to their ethical risks (e.g., they have been exploited for DDoS [42]). In addition, we acknowledge hops and addresses reported by traceroute are not always complete or reliable, when devices refuse to respond, respond with spoofed addresses, or co-locate at the same hop. Although we may find locations of observers (i.e., exact hops away from VPs) by decoding identifiers within unsolicited requests, if our decoys are redirected or replicated [40], responses with spoofed addresses from destinations are injected, and we may thus incorrectly locate observers at destinations. We attempt to introduce extra heuristics to remove noises in a best effort attempt (see Appendix E, e.g., under DNS interception); we also test VPN providers beforehand and do not integrate those resetting TTL in outgoing packets into our platform. Though we may not be able to reveal all observer IP addresses, we find hundreds of on-path observer IPs, accounting for over 80% of all shadowing behaviors (see Sections 4 and 5.2).

4 Traffic Shadowing Landscape

We run our experiment during Mar and Apr 2024 (2 months), switching between different VPs from VPN services continuously in a round-robin fashion without stop. In total, we send 46,613,616 DNS decoys, 1,694,109,438 HTTP decoys and 1,694,109,438 TLS decoys. This section reports the overall landscape, including problematic paths and location of on-path traffic observers.

Paths subject to traffic shadowing. Our measurement covers 157K client-server paths (4,364 VPs × 36 destinations) for DNS decoys, plus 10.1M paths (4,364 VPs × 2,325 destinations) for HTTP/TLS decoys. Figure 3 visualizes the ratio of paths (grouped by country) where decoys triggered unsolicited requests. We first find DNS decoys are more susceptible than HTTP and TLS, by associating with a wider range of problematic paths globally. HTTP and TLS decoys sent to servers in, or from VPs in China, are also more susceptible.

For DNS decoys, we only find those sent to popular public resolvers subject to traffic shadowing, while those to authoritative servers and our control resolver are not, suggesting observers exhibit preferences in traffic destination (similar to other types of manipulation, e.g., interception [40]). Particularly, a significant ratio (>70%) of problematic paths associate with Yandex, 114DNS, and OneDNS. Also interestingly, the ratio of problematic paths associated with 114DNS (a Chinese DNS vendor) is high only when VPs are also located in China (further investigated in Section 5). For HTTP and TLS decoys, we find problematic paths associated with several destinations (e.g., CN, AD, US, and CA), while the ratio in CN being slightly higher than the others.

Location of on-path traffic observers. Through hop-by-hop tracerouting, Table 2 shows the normalized location distribution of traffic observers found on problematic paths. For DNS decoys, we find on 99.7% of problematic paths, unsolicited requests are not triggered *until* they reach the destination (i.e., normalized TTL=10), indicating DNS traffic shadowing is mostly exhibited at the target resolvers (further investigated in Section 5). By contrast, within 97.7% (HTTP) and 35% (TLS) of problematic paths, on-path observers are capturing messages on-the-wire, particularly residing in the middle (i.e., normalized TTL=4, 5, 6) of problematic paths.

From the 572 IP addresses of traffic observers revealed by ICMP TTL Exceeded messages (i.e., of the exact hop triggering unsolicited requests), we find most are located in CN (448, 79%), echoing our prior finding that the ratio of problematic paths rises if associated with the two countries. Further, Table 3 shows top ASes where observer IP addresses reside. We find they are typically located in large ISPs (e.g., Chinanet) or cloud platforms (e.g., HostRoyale), both having strong paths to other networks. As a result, impact of such observers can be substantial, as in addition to traffic from networks and countries they reside, they may also have the capability to record traffic sent from and to servers in other countries.

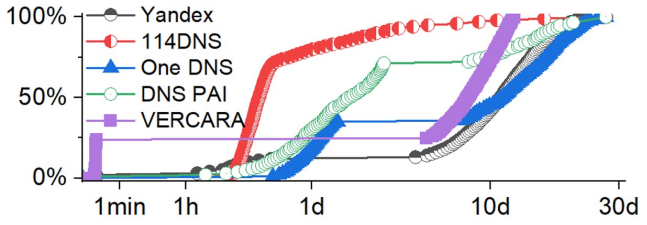

Figure 4: Cumulative distribution of time between unsolicited requests and initial DNS decoy (to $Resolver_h$).

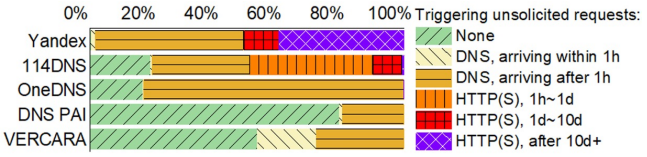

Figure 5: Breakdown of DNS decoys per destination

5 Behavioral Analysis

5.1 Traffic Shadowing against DNS Decoys

Temporal features. We use the time interval between sending one decoy and arrival of unsolicited requests bearing the same data as a quantitative proxy into how long (at least) user data is retained by traffic observers. For the top 5 destination DNS resolvers we find associated with the most ratio of problematic paths (i.e., Yandex, 114DNS, OneDNS, DNSPAI, and Vercara, see Figure 3; we term this resolver set as $Resolver_h$), Figure 4 shows the cumulative distribution of time interval between initial decoy and unsolicited requests. We find sizable proportion of unsolicited requests come either within 1 minute, or hours or even days after initial decoys. All unsolicited requests arriving within 1 minute are labeled as `DNS-DNS` (i.e., are repeated DNS queries for the same name), presumably resulting from benign implementation choices of DNS resolvers (e.g., retries; recall that 99.7% of DNS traffic shadowing occur at destination, see Section 4). While active cache refreshing mechanisms [36] and APIs (e.g., of OpenDNS [46]) may also produce unsolicited requests, we do not believe this is the major cause - we configure TTL=3,600 for wildcard DNS records of experiment domains (see Section 3) but do not find noticeable spikes (in Figure 4) around 1h or other hourly marks. Temporal features associated with Yandex, One DNS and DNSPAI are similar: unsolicited requests largely come either in 1 day or after days, suggesting widely-adopted shadowing patterns, or even possibility of the same exhibitors behind. All unsolicited HTTP(S) requests triggered by DNS decoys arrive at least 1h later. For the other 15 public resolvers beyond $Resolver_h$, 95% of unsolicited requests arrive within 1 minute.

We also find data in a significant portion of DNS decoys is leveraged multiple times: over 1 hour after emission from VP, 51% of DNS decoys *still* produce more than 3 unsolicited requests, and 2.4% produce more than 10. This outcome suggests data observed from traffic may not be removed after the first occurrence of shadowing and is retained (or even presumably stored) longer than expected.

Protocol combination. From `Decoy-Request` protocol combinations of unsolicited requests, we investigate how data embedded in decoys is leveraged. Figure 5 breaks down all DNS decoys per destination resolver, further grouped by time between their emission and unsolicited requests they trigger. For resolvers beyond $Resolver_h$, we only capture unsolicited DNS queries, most arriving within one hour. By contrast, 50% of DNS decoys we send to Yandex and 114DNS trigger unsolicited HTTP or HTTPS messages after hours or days, which falls beyond common implementation choices (e.g., cache refreshing) and shows probing incentives of traffic shadowing exhibitors (their payloads examined below).

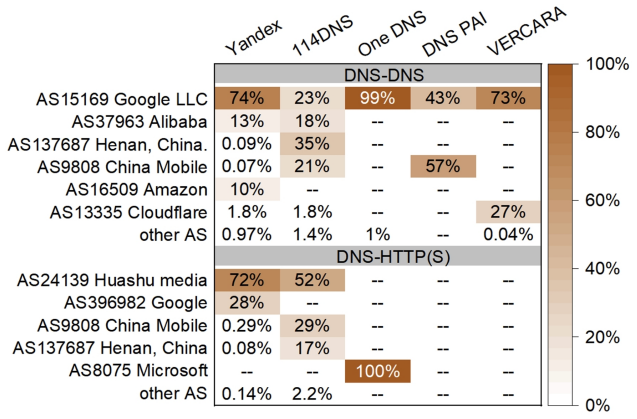

Figure 6: Origin ASes of unsolicited requests triggered by DNS decoys sent to $Resolver_h$.

Origin of unsolicited requests. While not necessarily related to traffic observers (or sometimes *not* associated with traffic shadowing exhibitors at all), the origin of unsolicited requests becomes processors of user data extracted from decoys. We inspect source IP addresses of unsolicited requests arriving at our honeypots and present top ASes in Figure 6. We first find for unsolicited DNS queries, Google (AS15169) becomes a significant origin, presumably because traffic shadowing exhibitors prefer querying Google Public DNS (the most popular public DNS service according to use metrics [13]) for the observed domain names. In addition, we find DNS decoys sent to one resolver may result in unsolicited requests originating from multiple ASes (e.g., DNS decoys to 114DNS primarily triggers unsolicited DNS queries from 4 ASes, including ISPs and Cloud platforms), suggesting diversified flows of data from resolvers to the origin servers. Through communication with operators of large resolvers, we learn that some resolvers may transfer their query data to other networks for operational considerations, which might have resulted in this outcome. Also notably, leveraging Spamhaus [8], a respected IP blocklist widely used [2, 3, 45], we find 5.2% of the origin IPs have been labeled as malicious.

HTTP and HTTPS probing incentives. Around 50% of DNS decoys sent to 114DNS and Yandex trigger unsolicited HTTP or HTTPS messages (see Figure 5), and we examine their payloads. From HTTP paths we learn that most requests (95%) are performing *path enumeration* that attempts to yield directories of our honey website (i.e., servers behind domain names in decoy DNS queries). By checking against the exploit-db database [6], we do not find requests with highly malicious payloads or vulnerability exploit codes, suggesting they may largely originate from innocent scans exhibited for security features. However, when checked against the Spamhaus IP blocklist, we find a larger portion of origin IP addresses

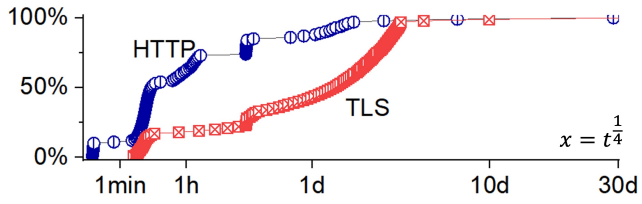

Figure 7: Cumulative distribution of time between unsolicited requests and HTTP (/TLS) decoy.

of the unsolicited requests (57% of HTTP and 72% of HTTPS) have been labeled as malicious. We presume this outcome reflects proxies which security organizations use to probe for malicious websites (e.g., as reported in [43]) may hit the blocklists.

Case study I: Yandex. Over 99% of DNS decoys sent to Yandex are subject to traffic shadowing (see Figure 5), data within which is typically retained for days (see Figure 4). 51% result in unsolicited HTTP/HTTPS requests, bearing clear probing and directory enumerating incentives targeting our honey websites. Though we actively contacted Yandex for reasons behind, they do not respond to our inquiry.

Case study II: 114DNS. From traceroute data (collected during Phase II of methodology), we attribute this outcome to different tactics of its anycast instances - decoys arriving at its CN instances trigger unsolicited requests, while its US instances do not (we position anycast instance from the geo-location of the nearest hop to the destination). We speculate that 114DNS conducts security analysis based on a large-scale passive DNS dataset they collect.

5.2 Traffic Shadowing against HTTP/TLS

Temporal features. Similarly, Figure 7 shows the cumulative distribution of time intervals between initial decoys and unsolicited requests. Compared to DNS decoys (Figure 4), data observed from HTTP and TLS decoys is retained for a shorter period of time, as a smaller portion of unsolicited requests arrives after days. We find some correlation between this finding and the location of traffic observers targeting different protocols (Table 2) - if observers are located in the middle of paths (e.g., 97.7% for HTTP decoys), then the time interval becomes generally shorter; the interval becomes larger if observers are at destination servers (e.g., 99.7% for DNS and 65% for TLS). This is possibly due to the limited storage capacity of routing devices serving as traffic observers.

Protocol combination and origin of unsolicited requests. We group HTTP (/TLS) decoys by observer ASes (derived from observer IPs in ICMP error responses, see Figure 2), and find the top 5 ASes account for over 80% of shadowing behaviors. They are located in CN (3 ASes, e.g., AS4134 Chinanet), US (AS40444 Constant Contact), and CA (AS29988 Rogers Communications). Protocol combinations differ among observer networks: when HTTP decoys are observed by devices within AS4134, 66% (17%) of them result in unsolicited HTTP(S) requests; all HTTP decoys observed by AS29988 produce unsolicited DNS requests only. Different from DNS shadowing (Figure 6), for HTTP (/TLS) decoys, we find a sizable ratio of unsolicited requests originating from the same networks as traffic observers (e.g., 100% of HTTP decoys observed by AS 40444 and AS29988 triggers unsolicited DNS queries from the same ASes).

Open ports of observers on the wire. By actively probing for their open ports and banners, we attempt to reveal what types of device traffic observers are. While, unfortunately, most (92%) observers do not have open ports, we find the most commonly open port among the remainder is 179 (BGP), indicating they are routing devices between networks.

HTTP and HTTPS probing incentives. Similar to unsolicited requests triggered by DNS decoys, more than 90% of HTTP paths yield directories of our honey website, while no malicious payloads or vulnerability exploit codes are found. When checked against the Spamhaus IP blocklist, we also find a significant portion of origin IP addresses (45% of HTTP and 55% of HTTPS) have been labeled as malicious.

Case III: HTTP (/TLS) observers in China. Most traffic observers we find are located in China (see Table 3), primarily in the populated provinces of Jiangsu, Shanghai, and Beijing. 85% unsolicited requests originate from local ISPs (e.g., AS4134 CHINANET-BACKBONE and AS140292 CHINATELECOM Jiangsu). We actively contact several Chinese ISPs and share our observations with them. From their feedbacks, we find most network operators are actually unaware of traffic observing and shadowing devices existing within their networks. Based on their operational experience, they speculate that such devices may serve security purposes, such as detecting phishing domains (e.g., monitoring and probing domains from DNS queries, especially newly-observed ones, which explain the outcome where our DNS queries for nonce domains were repeatedly queried). However, the operations potentially bear risks as clients are not aware of on-path data observing.

6 Discussion and Conclusion

We uncover the broad landscape of traffic shadowing against DNS, HTTP, and TLS messages. Alarmingly, data can be retained long, leveraged for multiple times, and triggers unsolicited requests from potentially abusive networks. For traffic observers on the wire, we find most of them located in ISP networks. We believe ISPs should learn about the risks of traffic shadowing and establish detection mechanisms to find unknown traffic shadowing exhibitors residing in their networks. We also try to connect to network operators about our results and receive some feedback. That said, the identities and purposes behind traffic shadowing exhibitors remain largely unknown, where future efforts still needs to be done.

Encryption (e.g., TLS and encrypted DNS) prevents data from being observed on the wire. However, note that encrypted protocols may also bear clear-text data of interest (e.g., data we embed in SNI is observed), stressing the need for deploying updated versions (e.g., TLS 1.3 with ECH [54]). In addition, encryption does *not* mitigate data collection by the destination server (especially for DNS), which decodes the message and sees everything. To enhance privacy, we recommend "oblivious" solutions splitting the visibility of message origin and content (e.g., OHTTP [58] and ODoH [37]).

Acknowledgments

This work is supported by the National Key Research and Development Program of China (No. 2023YFB3105600) and the National Natural Science Foundation of China (62102218). Baojun Liu is the corresponding author. Zhou Li is supported by NSF CNS-2047476.

References

[1] 2021. Speedchecker - Global Internet Testing. https://www.speedchecker.com/probeapi/.

[2] 2023. Postmaster. https://sendersupport.olc.protection.outlook.com/pm/troubleshooting.aspx.

[3] 2023. SMTP Error Codes. https://senders.yahooinc.com/smtp-error-codes/.

[4] 2023. Tor Metrics. https://metrics.torproject.org/.

[5] 2024. Bright data: Fast Residential Proxies. https://brightdata.com/proxy-types/residential-proxies.

[6] 2024. exploit-db. https://www.exploit-db.com/.

[7] 2024. Proxyrack: Premium Residential Proxies. https://www.proxyrack.com/premium-geo-residential/.

[8] 2024. Spamhaus. https://www.spamhaus.org/.

[9] Alice, Bob, Carol, Jan Beznazwy, and Amir Houmansadr. 2020. How China Detects and Blocks Shadowsocks. In *IMC '20: ACM Internet Measurement Conference, Virtual Event, USA, October 27-29, 2020*. ACM, 111–124. https://doi.org/10.1145/3419394.3423644

[10] Abderrahmen Amich, Birhanu Eshete, Vinod Yegneswaran, and Nguyen Phong Hoang. 2023. DeResistor: Toward Detection-Resistant Probing for Evasion of Internet Censorship. In *32nd USENIX Security Symposium, USENIX Security 2023, Anaheim, CA, USA, August 9-11, 2023*, Joseph A. Calandrino and Carmela Troncoso (Eds.). USENIX Association, 2617–2633. https://www.usenix.org/conference/usenixsecurity23/presentation/amich

[11] IP Geolocation API. 2024. IP API. https://ip-api.com.

[12] IP Information API. 2024. IP Info. https://ipinfo.io.

[13] APNIC Labs. 2024. Use of DNS Resolvers for World. https://stats.labs.apnic.net/rvrs.

[14] Michael D. Bailey, David Dittrich, Erin Kenneally, and Douglas Maughan. 2012. The Menlo Report. *IEEE Secur. Priv.* 10, 2 (2012), 71–75. https://doi.org/10.1109/MSP.2012.52

[15] Abhishek Bhaskar and Paul Pearce. 2022. Many Roads Lead To Rome: How Packet Headers Influence DNS Censorship Measurement. In *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, Kevin R. B. Butler and Kurt Thomas (Eds.). USENIX Association, 449–464. https://www.usenix.org/conference/usenixsecurity22/presentation/bhaskar

[16] Rui Bian, Lin Jin, Shuai Hao, Haining Wang, and Chase Cotton. 2024. Silent Observers Make a Difference: A Large-scale Analysis of Transparent Proxies on the Internet. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*.

[17] Kevin Bock, George Hughey, Xiao Qiang, and Dave Levin. 2019. Geneva: Evolving Censorship Evasion Strategies. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019, London, UK, November 11-15, 2019*, Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz (Eds.). ACM, 2199–2214. https://doi.org/10.1145/3319535.3363189

[18] CAIDA. 2024. Archipelago (Ark) Measurement Infrastructure. https://www.caida.org/projects/ark/.

[19] Sambuddho Chakravarty, Georgios Portokalidis, Michalis Polychronakis, and Angelos D. Keromytis. 2015. Detection and analysis of eavesdropping in anonymous communication networks. *Int. J. Inf. Sec.* 14, 3 (2015), 205–220. https://doi.org/10.1007/S10207-014-0256-7

[20] Deliang Chang, Joann Qiongna Chen, Zhou Li, and Xing Li. 2022. Hide and Seek: Revisiting DNS-based User Tracking. In *7th IEEE European Symposium on Security and Privacy, EuroS&P 2022, Genoa, Italy, June 6-10, 2022*. IEEE, 188–205. https://doi.org/10.1109/EUROSP53844.2022.00020

[21] Taejoong Chung, David R. Choffnes, and Alan Mislove. 2016. Tunneling for Transparency: A Large-Scale Analysis of End-to-End Violations in the Internet. In *Proceedings of the 2016 ACM on Internet Measurement Conference, IMC 2016, Santa Monica, CA, USA, November 14-16, 2016*, Phillipa Gill, John S. Heidemann, John W. Byers, and Ramesh Govindan (Eds.). ACM, 199–213. http://dl.acm.org/citation.cfm?id=2987455

[22] Jedidiah R. Crandall, Daniel Zinn, Michael Byrd, Earl T. Barr, and Rich East. 2007. ConceptDoppler: a weather tracker for internet censorship. In *Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, Alexandria, Virginia, USA, October 28-31, 2007*, Peng Ning, Sabrina De Capitani di Vimercati, and Paul F. Syverson (Eds.). ACM, 352–365. https://doi.org/10.1145/1315245.1315290

[23] Ryan Craven, Robert Beverly, and Mark Allman. 2014. A middlebox-cooperative TCP for a non end-to-end internet. In *ACM SIGCOMM 2014 Conference, SIGCOMM'14, Chicago, IL, USA, August 17-22, 2014*, Fabián E. Bustamante, Y. Charlie Hu, Arvind Krishnamurthy, and Sylvia Ratnasamy (Eds.). ACM, 151–162. https://doi.org/10.1145/2619239.2626321

[24] Gregory Detal, Benjamin Hesmans, Olivier Bonaventure, Yves Vanaubel, and Benoit Donnet. 2013. Revealing middlebox interference with tracebox. In *Proceedings of the 2013 Internet Measurement Conference, IMC 2013, Barcelona, Spain, October 23-25, 2013*, Konstantina Papagiannaki, P. Krishna Gummadi, and Craig Partridge (Eds.). ACM, 1–8. https://doi.org/10.1145/2504730.2504757

[25] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. 2013. ZMap: Fast Internet-wide Scanning and Its Security Applications. In *Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013*, Samuel T. King (Ed.). USENIX Association, 605–620. https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric

[26] Korian Edeline and Benoit Donnet. 2017. A First Look at the Prevalence and Persistence of Middleboxes in the Wild. In *29th International Teletraffic Congress, ITC 2017, Genoa, Italy, September 4-8, 2017*, Raffaele Bolla and Florin Ciucu (Eds.). IEEE, 161–168. https://doi.org/10.23919/ITC.2017.8064352

[27] Stephen Farrell and Hannes Tschofenig. 2014. Pervasive Monitoring Is an Attack. *RFC* 7258 (2014), 1–6. https://doi.org/10.17487/RFC7258

[28] Arturo Filastò and Jacob Appelbaum. 2012. OONI: Open Observatory of Network Interference. In *2nd USENIX Workshop on Free and Open Communications on the Internet, FOCI '12, Bellevue, WA, USA, August 6, 2012*, Roger Dingledine and Joss Wright (Eds.). USENIX Association. https://www.usenix.org/conference/foci12/workshop-program/presentation/filast%C3%B2

[29] Google. 2024. Google Ads. https://ads.google.com/.

[30] Michael Harrity, Kevin Bock, Frederick Sell, and Dave Levin. 2022. GET /out: Automated Discovery of Application-Layer Censorship Evasion Strategies. In *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, Kevin R. B. Butler and Kurt Thomas (Eds.). USENIX Association, 465–483. https://www.usenix.org/conference/usenixsecurity22/presentation/harrity

[31] Benjamin Hesmans, Fabien Duchene, Christoph Paasch, Gregory Detal, and Olivier Bonaventure. 2013. Are TCP extensions middlebox-proof?. In *Proceedings of the 2013 workshop on Hot topics in middleboxes and network function virtualization, HotMiddlebox 2013, Santa Barbara, California, USA, December 9, 2013*, Felipe Huici and Vyas Sekar (Eds.). ACM, 37–42. https://doi.org/10.1145/2535828.2535830

[32] Fahad Hilal and Oliver Gasser. 2023. Yarrpbox: Detecting Middleboxes at Internet-Scale. *PACMNET* 1, CoNEXT1 (2023), 4:1–4:23. https://doi.org/10.1145/3595290

[33] Nguyen Phong Hoang, Arian Akhavan Niaki, Jakub Dalek, Jeffrey Knockel, Pellaeon Lin, Bill Marczak, Masashi Crete-Nishihata, Phillipa Gill, and Michalis Polychronakis. 2021. How Great is the Great Firewall? Measuring China's DNS Censorship. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael D. Bailey and Rachel Greenstadt (Eds.). USENIX Association, 3381–3398. https://www.usenix.org/conference/usenixsecurity21/presentation/hoang

[34] Michio Honda, Yoshifumi Nishida, Costin Raiciu, Adam Greenhalgh, Mark Handley, and Hideyuki Tokuda. 2011. Is it still possible to extend TCP?. In *Proceedings of the 11th ACM SIGCOMM Internet Measurement Conference, IMC '11, Berlin, Germany, November 2-, 2011*, Patrick Thiran and Walter Willinger (Eds.). ACM, 181–194. https://doi.org/10.1145/2068816.2068834

[35] Geoff Huston. 2016. DNS Zombies. https://blog.apnic.net/2016/04/04/dns-zombies/.

[36] ICANN. 2024. ITHI Metric M5, Recursive Resolver Integrity. https://ithi.research.icann.org/graph-m5.html.

[37] Eric Kinnear, Patrick McManus, Tommy Pauly, Tanya Verma, and Christopher A Wood. 2022. RFC 9230: Oblivious DNS over HTTPS.

[38] Tadayoshi Kohno, Yasemin Acar, and Wulf Loh. 2023. Ethical Frameworks and Computer Security Trolley Problems: Foundations for Conversations. In *32nd USENIX Security Symposium, USENIX Security 2023, Anaheim, CA, USA, August 9-11, 2023*, Joseph A. Calandrino and Carmela Troncoso (Eds.). USENIX Association, 5145–5162. https://www.usenix.org/conference/usenixsecurity23/presentation/kohno

[39] Brian Kondracki, Johnny So, and Nick Nikiforakis. 2022. Uninvited Guests: Analyzing the Identity and Behavior of Certificate Transparency Bots. In *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, Kevin R. B. Butler and Kurt Thomas (Eds.). USENIX Association, 53–70. https://www.usenix.org/conference/usenixsecurity22/presentation/kondracki

[40] Baojun Liu, Chaoyi Lu, Haixin Duan, Ying Liu, Zhou Li, Shuang Hao, and Min Yang. 2018. Who Is Answering My Queries: Understanding and Characterizing Interception of the DNS Resolution Path. In *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, William Enck and Adrienne Porter Felt (Eds.). USENIX Association, 1113–1128. https://www.usenix.org/conference/usenixsecurity18/presentation/liu-baojun

[41] Alberto Medina, Mark Allman, and Sally Floyd. 2004. Measuring interactions between transport protocols and middleboxes. In *Proceedings of the 4th ACM SIGCOMM Internet Measurement Conference, IMC 2004, Taormina, Sicily, Italy, October 25-27, 2004*, Alfio Lombardo and James F. Kurose (Eds.). ACM, 336–341. https://doi.org/10.1145/1028788.1028835

[42] Xianghang Mi, Xuan Feng, Xiaojing Liao, Baojun Liu, XiaoFeng Wang, Feng Qian, Zhou Li, Sumayah A. Alrwais, Limin Sun, and Ying Liu. 2019. Resident Evil: Understanding Residential IP Proxy as a Dark Service. In *2019 IEEE Symposium on Security and Privacy (SP)*. 1185–1201. https://doi.org/10.1109/SP.2019.00011

[43] Ariana Mirian, Alisha Ukani, Ian D. Foster, Gautam Akiwate, Taner Halicioglu, Cynthia T. Moore, Alex C. Snoeren, Geoffrey M. Voelker, and Stefan Savage. 2023. In the Line of Fire: Risks of DPI-triggered Data Collection. In *2023 Cyber Security Experimentation and Test Workshop, CSET 2023, Marina del Rey, CA, USA, August*

7-8, 2023. ACM, 57–63. https://doi.org/10.1145/3607505.3607526
[44] Arian Akhavan Niaki, Shinyoung Cho, Zachary Weinberg, Nguyen Phong Hoang, Abbas Razaghpanah, Nicolas Christin, and Phillipa Gill. 2020. ICLab: A Global, Longitudinal Internet Censorship Measurement Platform. In *2020 IEEE Symposium on Security and Privacy, SP 2020, San Francisco, CA, USA, May 18-21, 2020*. IEEE, 135–151. https://doi.org/10.1109/SP40000.2020.00014
[45] Leo Oliver, Gautam Akiwate, Matthew Luckie, Ben Du, and kc claffy. 2022. Stop, DROP, and ROA: Effectiveness of Defenses through the lens of DROP. In *Proceedings of the 22nd ACM Internet Measurement Conference*. 730–737.
[46] OpenDNS. 2024. OpenDNS Cache Check. http://cachecheck.opendns.com.
[47] Angela Orebaugh and Becky Pinkard. 2011. *Nmap in the enterprise: your guide to network scanning*. Elsevier.
[48] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nicholas Weaver, and Vern Paxson. 2017. Global Measurement of DNS Manipulation. In *26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017*, Engin Kirda and Thomas Ristenpart (Eds.). USENIX Association, 307–323. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/pearce
[49] Victor Le Pochat, Tom van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019, San Diego, California, USA, February 24-27, 2019*. The Internet Society. https://www.ndss-symposium.org/ndss-paper/tranco-a-research-oriented-top-sites-ranking-hardened-against-manipulation/
[50] Mattew Prince. 2024. Introducing WARP: fixing mobile Internet performance and security. https://blog.cloudflare.com/1111-warp-better-vpn/.
[51] Ram Sundara Raman, Prerana Shenoy, Katharina Kohls, and Roya Ensafi. 2020. Censored Planet: An Internet-wide, Longitudinal Censorship Observatory. In *CCS '20: 2020 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, USA, November 9-13, 2020*, Jay Ligatti, Xinming Ou, Jonathan Katz, and Giovanni Vigna (Eds.). ACM, 49–66. https://doi.org/10.1145/3372297.3417883
[52] Ram Sundara Raman, Mona Wang, Jakub Dalek, Jonathan R. Mayer, and Roya Ensafi. 2022. Network measurement methods for locating and examining censorship devices. In *Proceedings of the 18th International Conference on emerging Networking EXperiments and Technologies, CoNEXT 2022, Roma, Italy, December 6-9, 2022*, Giuseppe Bianchi and Alessandro Mei (Eds.). ACM, 18–34. https://doi.org/10.1145/3555050.3569133
[53] Audrey Randall, Enze Liu, Ramakrishna Padmanabhan, Gautam Akiwate, Geoffrey M. Voelker, Stefan Savage, and Aaron Schulman. 2021. Home is where the hijacking is: understanding DNS interception by residential routers. In *IMC '21: ACM Internet Measurement Conference, Virtual Event, USA, November 2-4, 2021*, Dave Levin, Alan Mislove, Johanna Amann, and Matthew Luckie (Eds.). ACM, 390–397. https://doi.org/10.1145/3487552.3487817
[54] Eric Rescorla, Kazuho Oku, Nick Sullivan, and Christopher A. Wood. 2024. *TLS Encrypted Client Hello draft-ietf-tls-esni-18*. Technical Report. Internet draft.[Online]. Available:https://datatracker.ietf.org/doc/draft-ietf-tls-esni/.
[55] Will Scott, Thomas E. Anderson, Tadayoshi Kohno, and Arvind Krishnamurthy. 2016. Satellite: Joint Analysis of CDNs and Network-Level Interference. In *2016 USENIX Annual Technical Conference, USENIX ATC 2016, Denver, CO, USA, June 22-24, 2016*, Ajay Gulati and Hakim Weatherspoon (Eds.). USENIX Association, 195–208. https://www.usenix.org/conference/atc16/technical-sessions/presentation/scott
[56] RIPE NCC Staff. 2015. Ripe atlas: A global internet measurement network. *Internet Protocol Journal* 18, 3 (2015), 2–26.
[57] Valentin Thirion, Korian Edeline, and Benoit Donnet. 2015. Tracking Middleboxes in the Mobile World with TraceboxAndroid. In *Traffic Monitoring and Analysis - 7th International Workshop, TMA 2015, Barcelona, Spain, April 21-24, 2015. Proceedings (Lecture Notes in Computer Science, Vol. 9053)*, Moritz Steiner, Pere Barlet-Ros, and Olivier Bonaventure (Eds.). Springer, 79–91. https://doi.org/10.1007/978-3-319-17172-2_6
[58] Martin Thomson and Christopher A. Wood. 2022. *Oblivious HTTP draft-thomson-http-oblivious-02*. Technical Report. Internet draft.[Online]. Available:https://datatracker.ietf.org/doc/draft-thomson-http-oblivious/.
[59] Giorgos Tsirantonakis, Panagiotis Ilia, Sotiris Ioannidis, Elias Athanasopoulos, and Michalis Polychronakis. 2018. A Large-scale Analysis of Content Modification by Open HTTP Proxies. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018*. The Internet Society. https://www.ndss-symposium.org/wp-content/uploads/2018/02/ndss2018_04A-1_Tsirantonakis_paper.pdf
[60] Zhaoguang Wang, Zhiyun Qian, Qiang Xu, Zhuoqing Morley Mao, and Ming Zhang. 2011. An untold story of middleboxes in cellular networks. In *Proceedings of the ACM SIGCOMM 2011 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Toronto, ON, Canada, August 15-19, 2011*, Srinivasan Keshav, Jörg Liebeherr, John W. Byers, and Jeffrey C. Mogul (Eds.). ACM, 374–385. https://doi.org/10.1145/2018436.2018479
[61] Mingkui Wei. 2021. Domain Shadowing: Leveraging Content Delivery Networks for Robust Blocking-Resistant Communications. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael D. Bailey and Rachel Greenstadt (Eds.). USENIX Association, 3327–3343. https://www.usenix.org/conference/usenixsecurity21/presentation/wei
[62] Tim Wicinski. 2021. DNS Privacy Considerations. RFC 9076. https://doi.org/10.17487/RFC9076
[63] Mingshi Wu, Jackson Sippe, Danesh Sivakumar, Jack Burg, Peter Anderson, Xiaokang Wang, Kevin Bock, Amir Houmansadr, Dave Levin, and Eric Wustrow. 2023. How the Great Firewall of China Detects and Blocks Fully Encrypted Traffic. In *32nd USENIX Security Symposium, USENIX Security 2023, Anaheim, CA, USA, August 9-11, 2023*, Joseph A. Calandrino and Carmela Troncoso (Eds.). USENIX Association, 2653–2670. https://www.usenix.org/conference/usenixsecurity23/presentation/wu-mingshi
[64] Xueyang Xu, Zhuoqing Morley Mao, and J. Alex Halderman. 2011. Internet Censorship in China: Where Does the Filtering Occur?. In *Passive and Active Measurement - 12th International Conference, PAM 2011, Atlanta, GA, USA, March 20-22, 2011. Proceedings (Lecture Notes in Computer Science, Vol. 6579)*, Neil Spring and George F. Riley (Eds.). Springer, 133–142. https://doi.org/10.1007/978-3-642-19260-9_14
[65] Diwen Xue, Michalis Kallitsis, Amir Houmansadr, and Roya Ensafi. 2024. Fingerprinting Obfuscated Proxy Traffic with Encapsulated TLS Handshakes. In *33st USENIX Security Symposium, USENIX Security 2024, PHILADELPHIA, PA, USA, AUGUST 14–16, 2024*. USENIX Association. https://www.usenix.org/conference/usenixsecurity24/presentation/xue
[66] Diwen Xue, Benjamin Mixon-Baca, ValdikSS, Anna Ablove, Beau Kujath, Jedidiah R. Crandall, and Roya Ensafi. 2022. TSPU: Russia's decentralized censorship system. In *Proceedings of the 22nd ACM Internet Measurement Conference, IMC 2022, Nice, France, October 25-27, 2022*, Chadi Barakat, Cristel Pelsser, Theophilus A. Benson, and David R. Choffnes (Eds.). ACM, 179–194. https://doi.org/10.1145/3517745.3561461
[67] Diwen Xue, Reethika Ramesh, Arham Jain, Michalis Kallitsis, J. Alex Halderman, Jedidiah R. Crandall, and Roya Ensafi. 2022. OpenVPN is Open to VPN Fingerprinting. In *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, Kevin R. B. Butler and Kurt Thomas (Eds.). USENIX Association, 483–500. https://www.usenix.org/conference/usenixsecurity22/presentation/xue-diwen
[68] Diwen Xue, Reethika Ramesh, Valdik S. S, Leonid Evdokimov, Andrey Viktorov, Arham Jain, Eric Wustrow, Simone Basso, and Roya Ensafi. 2021. Throttling Twitter: an emerging censorship technique in Russia. In *IMC '21: ACM Internet Measurement Conference, Virtual Event, USA, November 2-4, 2021*, Dave Levin, Alan Mislove, Johanna Amann, and Matthew Luckie (Eds.). ACM, 435–443. https://doi.org/10.1145/3487552.3487858
[69] Mingshuo Yang, Yunnan Yu, Xianghang Mi, Shujun Tang, Shanqing Guo, Yilin Li, Xiaofeng Zheng, and Haixin Duan. 2022. An Extensive Study of Residential Proxies in China. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi (Eds.). ACM, 3049–3062. https://doi.org/10.1145/3548606.3559377
[70] Mingming Zhang, Baojun Liu, Chaoyi Lu, Jia Zhang, Shuang Hao, and Hai-Xin Duan. 2018. Measuring Privacy Threats in China-Wide Mobile Networks. In *8th USENIX Workshop on Free and Open Communications on the Internet, FOCI 2018, Baltimore, MD, USA, August 14, 2018*, Lex Gill and Rob Jansen (Eds.). USENIX Association. https://www.usenix.org/conference/foci18/presentation/zhang

A Ethics

The authors' institution does not subordinate an Institutional Review Board (IRB). We design and assess our experiments following ethical frameworks (including the Menlo Report [14] and one recent work establishing public ethical metrics for security and privacy esearch [38]), as well as prior works adopting similar VPN-based measurement methods [44, 51].

Usage of VPN-based vantage points. To spread decoy traffic, we recruit VPs from commercial VPNs, following the practice of ICLab [44], a well-established VPN-based measurement platform. We recriut VPs from datacenter VPNs and refrain from leveraging user-hosted VPN platforms (i.e., residential VPs) due to their potentially illegal methods of VP recruitment [42, 69] and potential usage in abusive activities (e.g., DDoS attacks [42]). While we attempt to remove residential and consider datacenter VPN providers only (see Appendix C for the determination process), we acknowledge

chances where we may not be able to eliminate every residential VP from our platform or experiment. In such cases, sending our decoys will lead to potential ethical implications.

We perform our measurements adhering to the terms of service of the VPN platforms: the decoys we send bear no sensitive keywords or domain names (e.g., prone to censorship), have valid destinations (e.g., public HTTP servers of top sites), and are not malformed. These measures ensure that we do not expose the operators to additional risks beyond what they would normally encounter when operating commercial services. Additionally, VPN providers are well-aware of the potential security risks associated with operating a VPN business. They are unlikely to deploy a server in a country where the company or its employees might face potential legal or extralegal consequences for the actions of its users. During our experiment, we do not observe any blockings by destination servers against our VPs (e.g., when further requests from the VP fail during our 2-month period), suggesting our decoys, though embedded with experiment data, do not trigger filtering or intrusion prevention mechanisms.

Generating decoy traffic. We select DNS, HTTP, and TLS as decoy protocols. The message format we send is regular (i.e., we do not generate malformed packets), ensuring that they do not yield transmission or processing problems for VPs, on-path devices, or destination servers. The decoys we send do not include personal data (all decoys are generated with rules documented in Section 3) or keywords prone to censorship, reducing the risk of VPN server nodes violating local laws. To avoid overloading VPN platforms, on-path devices, and destination servers, we employ a strict rate limit at which we send decoy packets. According to our calculations, throughout the experiment, we send no more than 2 decoy packets per second to a given target, several orders of magnitude lower than prior works based on scanning methods [25, 47]. All decoys are sent to publicly accessible servers on the Internet (i.e., large public DNS resolvers and Tranco Top 1K sites) and the honeypot DNS server we build. To inform accidental visitors and origins of unsolicited requests, we document the purpose of our experiment and contact information on the homepage of our honeypot website, whose URL is not publicly advertised. Until the end of our experiment, we have not received any complaints.

Diverting and processing unsolicited requests. We embed our domain name registered exclusively for this experiment in decoys (bearing no personal data, only the generated domain name) to attract traffic shadowing exhibitors to visit our servers instead of others. We resolve our wildcard domain names to our honeypots. As a result, unsolicited DNS queries will arrive at our authoritative server, while unsolicited HTTP(S) requests will arrive at our honey websites. This design ensures no other innocent devices are involved or unexpectedly requested and that unsolicited requests will not be processed by others.

Evaluation by ethical metrics. Additionally, we evaluate our experiment against recent public ethical metrics summarized by Kohno et al. for security and privacy research [38] and find that our experiment aligns with the principles. Principles they propose are based on consequentialism and deontological ethics, with the core ideas that benefits should outweigh costs, and that ethics, laws, and human rights should be respected during the experimental process. First, our experiment does not collect any personal information from users, thus eliminating human rights issues. Second, our experiment and results are benificial for the community in understanding Internet traffic shadowing, a less-noticed type of on–path manipulation, while also encouraging network operators to become aware of privacy risks, investigate and shut down unknown exhibitors. Our experiment sends decoys at a low rate, consuming only negligible public network resources (such as network bandwidth). These align with principles of consequentialist ethics, where the benefits outweigh the costs. Finally, organizations and companies (e.g., operators of public DNS and HTTP services) related to our experiment may also benefit: we detect and report shadowing behaviors against traffic to their services, inform them of the associated privacy risks, and encourage them to deploy updated versions of mitigation protocols (e.g., encryption). As a result, we believe our study is benificial in general, while bring minimal harms, and meet common ethical principles.

Table 4: DNS servers to which we send decoys

Type	Name	IP
Public resolvers	Cloudflare	1.1.1.1
	CNNIC	1.2.4.8
	DNS PAI	101.226.4.6
	DNSPod	119.29.29.29
	DNS.Watch	84.200.69.80
	Oracle Dyn	216.146.35.35
	Google	8.8.8.8
	Hurricane	74.82.42.42
	Level3	209.244.0.3
	VERCARA	156.154.70.1
	One DNS	117.50.10.10
	OpenDNS	208.67.222.222
	Open NIC	217.160.166.161
	Quad9	9.9.9.9
	Yandex	77.88.8.8
	SafeDNS	195.46.39.39
	Freenom	80.80.80.80
	Baidu	180.76.76.76
	114DNS	114.114.114.114
	Quad101	101.101.101.101
Self-built resolver	self-built	–
Authoritative servers	13 roots	198.41.0.4, and others
	.com	192.12.94.30
	.org	199.19.57.1

B List of Public DNS Servers

The DNS servers to which we send decoys are listed in Table 4. These DNS servers are 20 large public DNS resolvers (selected from their use metrics [13]), 1 self-built resolver, 13 root servers, and 2 TLD servers.

C VPN Providers Integrated into Measurement Platform

We recruit vantage points from 6 VPN providers with global accessibility and 13 VPN providers dedicated to the Chinese market. Table 5 provides a listing to VPN providers we integrate into our measurement platform. To determine whether a provider is from datacenter (rather than residential), we first check the providers' websites for descriptions implying data centers (e.g., "provider-owned data centers"). In addition, before purchasing, we consulted all providers' customer service to verify their IPs come from datacenters. Though

Table 5: Listing of VPN providers integrated into our platform

	VPN Provider	URL to provider
Global	Anonine	https://anonine.com/
	AzireVPN	https://www.azirevpn.com/
	Cryptostorm	https://cryptostorm.is/
	HideMe	https://hide.me/
	PrivateInt	https://www.privateinternetaccess.com/
	PureVPN	https://www.purevpn.com/
China	QiXun	https://www.ipkuip.com/product/Buy?id=3
	XunYou	https://www.ipkuip.com/product/Buy?id=6
	YOYO	https://www.ipkuip.com/product/Buy?id=51
	BeiKe	https://www.ipkuip.com/product/Buy?id=44
	SunYunD	https://www.ipkuip.com/product/Buy?id=92
	HuoJian	https://www.ipkuip.com/product/Buy?id=128
	DuoDuo	https://www.ipkuip.com/product/Buy?id=116
	MoGu	https://www.juip.com/product/Buy?id=1032
	QiangZi	https://www.juip.com/product/Buy
	XunLian	https://www.juip.com/product/Buy
	TianTian	https://www.juip.com/product/Buy?id=71
	JiKe	https://www.juip.com/product/Buy
	XiGua	https://www.juip.com/product/Buy

* ipkuip.com and juip.com are two hub websites to the VPN providers in China. The URLs are only accessible within mainland China.

the above information is self-reported by VPN providers, we consider it truthful - as residential IPs are more expensive to maintain (e.g., on proxy network platforms [7], residential nodes are rented at higher prices than datacenters), claiming residential nodes as data center nodes is not considered a wise business decision. Also to examine after purchasing, we check all ASes of recruited VPs against the IPinfo database [12] and find 71/74 (Global phase, 96%) ASes are labeled as "hosting".

That said, we acknowledge the method explained above is considered as a best effort approach, which may not eliminate using residential VPs completely from our experiment. As stated in Appendix A, we acknowledge the potential ethical implications resulting from our design.

D Survey on Prior Measurement Platforms

Before choosing to build our own VPN-based measurement platform, we perform a survey of existing works in Table 6. For platforms surveyed, whether their VPs are residential is determined from public material: on websites of commercial platforms (e.g., Proxyrack offers residential proxy nodes [7]) or in papers (e.g., the ICLab paper [44] explicitly removed residential VPNs from consideration). Crowdsourcing platforms recruit volunteers to set up VPs, thus they contain residential VPs.

Because our methodology leverages tracerouting via altering IP layer TTLs of application messages, only VPN-based VPs that do not require volunteer participation meet this requirement. We also need to recruit VPs from a wide range of geo-locations and networks, thus removing WARP (covering Cloudflare ASes only) and ICLab (not available for public research) from consideration.

E Mitigating Noises in Experiment

Bias caused by DNS interception during tracerouting. On-path DNS interception devices [40] may redirect or replicate DNS queries, force them to be handled by alternative DNS servers, and return DNS responses from spoofed addresses of their intended destinations. During hop-by-hop tracerouting (Figure 2), if intercepted by devices at some hop in the middle, our DNS decoys may trigger responses from spoofed resolver address, while in fact, they might not have reached the intended destination yet, causing our methodology to incorrectly report observers at destination resolver. While under request replication, multiple requests also arrive at our authoritative servers (i.e., from intended resolvers and alternative resolvers); we filter out this case from traffic shadowing, as communication between clients and servers *is* intercepted when clients *are* waiting for responses, as opposed to silent on-path observers.

We design a method called *pair resolver* to remove VPs affected by DNS interception (already removed from VPs counted in Table 1). A pair resolver of a target resolver is another IP address within the same /24 that does not offer public DNS service (e.g., 1.1.1.4 as to 1.1.1.1). Consider DNS queries sent from a VP to one target resolver and its pair, they will be transmitted over the same path (because they share the same /24). If a DNS query from a VP to pair resolver of any server listed in Table 4 triggers a DNS response (normally it should not, because the pair resolver does not offer DNS service), then DNS interception is exhibited on this path, and we remove this VP from consideration. Due to different choices of routing, this method cannot entirely rule out DNS interception, which is a limitation of our experiment. However, we compare the paths to all target servers and their pair resolvers and find no AS-level differences.

Bias caused by resolver-authoritative paths. While our method locates the vast majority (>99.7%, see Section 4) of on-path DNS observers at destination resolvers, there is a possibility that traffic shadowing is not exhibited by the resolvers but behind them (i.e., on resolver-authoritative paths, where our measurement platform has no visibility). However, we do not believe traffic shadowing on the resolver-authoritative path is noticeable in its landscape or may skew our results, due to the following observations. First, during our experiments, we already send DNS decoys from global VPs to a wide range of authoritative servers (see Table 4) and do not observe unsolicited requests triggered from any of the paths (see Section 4). Second, DNS queries transmitted over resolver-authoritative paths bear less privacy risks [62], as on-path observers may not correlate query names with client IP addresses (because DNS queries originate from resolver addresses), making such behavior less attractive to traffic shadowing exhibitors.

Bias caused by VPN nodes. For security reasons, some VPN providers may impose restrictions or manipulate user traffic, e.g., reset the TTL of every outgoing packet, rendering our tracerouting approach (phase II) inaccurate. To eliminate such nodes, when establishing the VPN-based measurement platform, we check whether the VPN services engage in such traffic manipulation, e.g., by directly sending packets to our controlled server and inspect whether contents or TTL fields have been tampered with. We only keep and use VPN services that do not perform traffic manipulation for this experiment.

Table 6: Capabilities and comparison of measurement platforms

Category	Platform	General Purpose?	Volunteer-Free?	VP & Coverage				Message Relay & Customizing Capabilities								
				Resi	#	CC	AS	DNS	HTTP	TLS	Other	TCP	UDP	Ping	Tracert	TTL
Crowdsourcing	Ark [18]	●	○	●	119	44	95	○	○	○	○	◐	◐	●	●	○
Crowdsourcing	Speedchecker [1]	●	●	●	?	170	?	●	●	○	○	◐	◐	●	●	○
Crowdsourcing	RIPE Atlas [56]	●	○	●	12,979	169	3,781	◐	◐	◐	○	◐	◐	●	●	○
Crowdsourcing	OONI [28]	○	○	●	?	113	670	●	●	●	●	●	●	●	●	●
Advertising	Google Ads [29]	●	●	●	?	?	?	○	○	○	○	○	○	○	○	○
Scanners	Satellite-Iris [48, 55]	○	●	–	–	–	–	●	○	○	○	○	●	○	○	○
Proxies	BrightData [5]	●	●	●	72M	195	?	○	●	●	○	●	○	○	○	○
Proxies	ProxyRack [7]	●	●	●	5M	140	?	●	●	●	●	●	●	○	○	○
VPN	WARP [50]	●	●	○	?	?	?	●	●	●	●	●	●	●	●	●
VPN	ICLab [44]	○	◐	○	281	62	234	●	●	●	●	●	●	●	●	●
Tor	Tor [4]	●	○	●	2,200	54	248	●	●	●	●	●	●	○	○	○
VPN	This work	●	●	○	4,364	82	121	●	●	●	●	●	●	●	●	●

●: Yes; provides full customizing capabilities ◐: Partial; provides protocol support with restricted customizing capabilities ○: No; provides no support
?: unknown from public material or field test –: not applicable
Resi: residential VP; CC: country-code; Other: FTP, SMTP, and Telnet; Tracert: ICMP traceroute; TTL: customizing IP header TTL