Tracking the Stray Sheep: Understanding DNS Response Manipulation in the Wild

Wenhao Wu
wuwenhao22s@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences, Beijing, China

Zhaohua Wang
wangzh@cnic.cn
Computer Network Information Center, Chinese Academy of Sciences
Beijing, China

Zihan Li
lizihan24z@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences, Beijing, China

Qinxin Li
liqinxin23s@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences, Beijing, China

Yiming Xia
xiayiming24s@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences, Beijing, China

Chuan Gao
gaochuan24s@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences, Beijing, China

Guangxing Zhang
guangxing@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
Beijing, China

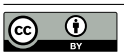
Zhenyu Li*
zyli@ict.ac.cn
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences, Beijing, China

Abstract

The Domain Name System (DNS) plays a crucial role in modern web applications; however, manipulations such as hijacking, tampering, and censorship can disrupt domain resolution, posing significant privacy and security risks. While such manipulations are prevalent across global DNS infrastructures, their scope and mechanisms remain poorly understood. Existing studies focus on country-level censorship or rely on authoritative data and passive traffic from selected domains, which prevents a comprehensive understanding. Moreover, the dynamic nature of modern DNS resolution, in which a single domain may resolve to thousands of edge servers, further complicates the detection of manipulated responses.

In this work, we propose a novel approach for measuring DNS manipulations based on resolution path analysis. Our method leverages CNAME chains and attributes of intermediate nodes in the DNS resolution process to link dynamic resolution results, enabling accurate detection of manipulation in highly dynamic DNS environments. We conduct large-scale measurements for 2,283 popular domains across global open DNS infrastructures. Measurement results reveal critical insights into DNS manipulation, uncovering the strategies and preferences of malicious manipulation operators and demonstrating how specific domains are exploited.

*Corresponding author: Zhenyu Li.



WWW '26, Dubai, United Arab Emirates.

ACM ISBN 979-8-4007-2307-0/2026/04
https://doi.org/10.1145/3774904.3792405

CCS Concepts

• **Networks** → **Network measurement**.

Keywords

Domain Name System; Network Measurement; DNS Manipulation

ACM Reference Format:
Wenhao Wu, Zhaohua Wang, Zihan Li, Qinxin Li, Yiming Xia, Chuan Gao, Guangxing Zhang, and Zhenyu Li. 2026. Tracking the Stray Sheep: Understanding DNS Response Manipulation in the Wild. In *Proceedings of the ACM Web Conference 2026 (WWW '26), April 13–17, 2026, Dubai, United Arab Emirates.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3774904.3792405

Resource Availability:
Our analysis and measurement data are publicly available at github.com/f-555/Tracking-Stray-Sheep (doi.org/10.5281/zenodo.18297300).

1 Introduction

The Domain Name System (DNS) supports modern web services by translating human-readable names into network addresses, thereby enabling access to resources ranging from static content to dynamic APIs, content delivery networks, and multimedia assets [6, 9, 33]. Yet, this importance also exposes DNS to manipulation. Adversaries can exploit DNS responses during interactions between users and DNS servers to redirect, intercept, or corrupt legitimate traffic flows. DNS manipulation spans a wide range of attack vectors [10, 11, 20], including cache poisoning [19], censorship, response injection, or hijacking. Each is capable of altering the intended mapping between domain names and IP addresses. Understanding the extent and occurrence of these manipulations is crucial for user security and

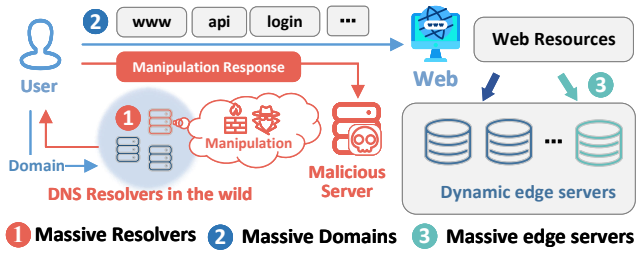


Figure 1: DNS manipulation: Scenario and challenges.

the reliable operation of web services. This motivates the need for large-scale measurements, particularly to analyze the behavior of globally distributed DNS servers.

With the growing scale and complexity of DNS infrastructures [35] and their integration into diverse web applications [6, 22, 23, 31], conducting large-scale measurements has become increasingly challenging. Specifically, as modern domains underpin heterogeneous infrastructures such as CDNs, APIs, and distributed backends [4, 34], naive methods based on IP or web content comparison fail to detect manipulations effectively. Meanwhile, web infrastructures have expanded dramatically, where a single domain may resolve to thousands of edge servers across diverse resolution paths and geographic regions [15]. Such challenges are compounded by the vast number of DNS resolvers in the wild [14, 24, 28], as the small fraction of manipulated responses can be easily concealed within predominantly legitimate traffic, making detection and analysis extremely difficult.

Given these challenges, existing approaches are ill-suited to the current DNS environment. Passive traffic analysis [3, 11, 21], while useful for post-incident forensics, introduces latency that hinders timely detection. Active measurement schemes are based on consistency-checking paradigms that compare DNS responses across vantage points [5, 25, 30]. These methods struggle with the inherent diversity of modern DNS, generating false positives in the presence of CDNs and geo-distributed edge servers, and failing to detect sophisticated manipulations that exhibit complex behaviors and objectives hidden in massive resolver populations. Collectively, these limitations highlight the need for a refined DNS manipulation measurement method that can handle today's complex, large-scale, and highly dynamic infrastructure.

To address these challenges, we propose a novel DNS manipulation measurement method that models DNS resolution as directed graphs. Each hop in the resolution process—including the original domain, CNAME chains, and A records—is represented as a node, while delegation and aliasing are used to link the diverse resolution paths observed across multiple resolvers. By analyzing these graphs, our method identifies paths that deviate from the expected patterns, revealing anomalies indicative of potential manipulation. Our method is resilient to the dynamics of the DNS resolution, operates on arbitrary domains, and requires neither cooperation from authoritative servers nor passive traffic collection.

Building on our methodology, we conduct large-scale measurements across the global open DNS infrastructure. We selected the top domains across major web application categories according to global popularity rankings [7, 18, 29], and further expanded these domains using semantically or popular prefix variations [36] (*e.g.,* `www`, `api`), resulting in a final set of 2,283 popular domains. These domains were measured across a representative set of DNS servers worldwide to assess manipulation. Our measurements reveal the widespread presence of DNS manipulation, yielding approximately one million manipulation reports, each representing the resolution outcome of a specific domain when queried through a particular resolver. We thoroughly analyzed all reports and carefully analyzed all destination IPs and around 7.6k web pages, categorizing all observed manipulations into six classes based on their intent.

Overall, we analyze the characteristics of the manipulation and make the following key observations:

- DNS manipulations are highly concentrated in a small set of destination IPs. High-risk categories, such as Abuse (redirect to abused IPs) and Redirect (redirect to third-party pages), rely on fewer than 20 IPs to handle over 90% of manipulation.
- Alterations in resolution chains reveal distinct operational strategies. Abuse manipulations typically shorten resolution chains to deliver malicious content, while Block (content blocking or censorship) lengthens chains via extra CNAMEs.
- Manipulations are always partial across domains. About 26% of DNS servers perform high-risk manipulation (*i.e.,* Abuse and Redirect) on only a subset of domains, enabling attackers to selectively target specific domains while evading detection.
- Manipulators exhibit clear domain-level preferences. High-risk manipulations (Abuse and Redirect) show strong, category-specific biases, with porn and government domains being the most heavily targeted by such manipulation.
- We conduct a case study to show manipulations targeting an abuse-related IP, illustrating how the macro patterns observed at scale translate into real-world behaviors and providing concrete validation for our large-scale measurement results.

2 Background and Related Work

2.1 DNS Manipulation and Measurement

Manipulation of DNS responses occurs in diverse contexts, ranging from traffic redirection for advertisement injection to state-level censorship, as well as malicious hijacking that routes queries to attacker-controlled infrastructure [27]. These manipulations demonstrate that DNS responses can be extensively altered during the recursive resolution process, with outcomes often dependent on the specific DNS resolvers used by end users. Interventions by open or local DNS resolvers have been widely reported [16, 26, 37]. Such behaviors are often localized, transient, and stealthy, making them difficult to observe without systematic measurement.

Passive Manipulation Analysis. Researchers have developed passive techniques to monitor DNS manipulation and provide valuable insights. For instance, large-scale network traces [3, 11, 12, 21] can be used to train AI-based detection models or to conduct passive Internet-wide measurements from embedded vantage points. Nosyk et al. [21] leveraged RIPE Atlas measurements to identify manipulation from the perspective of the root servers, while Houser et al. [11] examined historical traces from 2008 to 2020 to locate DNS hijacking activities. Although these approaches have advanced understanding of network and web behaviors, the scope of passive

measurements is inherently limited to specific domains or DNS infrastructures (*e.g.*, root servers).

Active Manipulation Measurement. Active measurement methods have proven effective in measuring global DNS manipulation, focusing on infrastructure status and the correctness of DNS responses. Existing approaches fall into two categories. The first leverages controlled authoritative servers and active probes, as in Liu et al. [17] and ODNS Clustering [35], revealing global interception and infrastructure interdependencies. However, these methods cannot generalize to arbitrary domains and may be evaded by the partial manipulation behavior of DNS resolvers (as Section 4.2). The second category relies solely on client-side vantage points, detecting manipulation by comparing responses across multiple responses [5, 25, 30]. This approach is commonly used to study state-level censorship, though its effectiveness depends heavily on the ability to cross-check response attributes.

2.2 Challenges and Motivation

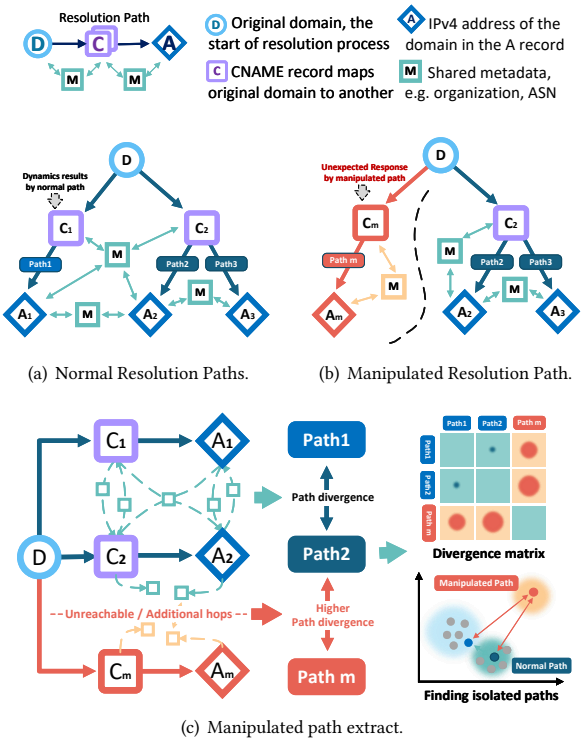
The increasing complexity of DNS infrastructure and applications underscores the urgent need for a refined measurement-based method to understand the manipulation behavior in the wild. Prior approaches face fundamental challenges as shown in Fig. 1. First, modern DNS resolution is highly dynamic: domains frequently resolve to massive shared, globally distributed edge nodes spanning multiple CDNs and autonomous systems [34]. Our measurements reveal that some domains can produce up to 10k distinct resolution paths (as discussed in Section 3.4). This dynamism blurs the line between normal diversity and malicious tampering. Second, the massive scale of the DNS ecosystem [24], encompassing millions of open DNS resolvers and a continuously expanding domain namespace, enlarges the attack surface beyond the reach of conventional monitoring. Third, adversaries exploit this scale and diversity of massive web domains to launch selective and stealthy manipulations, injecting forged responses only for targeted domains and specific resolver populations [16].

These limitations underscore the need for a measurement approach that can scale to today's DNS ecosystem and actively capture manipulation behaviors. Critically, such an approach must be applicable to the dynamics of DNS resolution.

3 Methodology

3.1 Preliminary and Overview

Modern web services increasingly rely on CDNs and edge infrastructures to deliver content with low latency. To efficiently steer traffic, domain owners often delegate domain resolution through mechanisms such as CNAME chains. While this deployment provides scalability and flexibility, it also introduces significant dynamics: queries to the same domain can yield different responses at different times or DNS resolvers. Luckily, these diverse responses are not arbitrary. Multiple IPs often cluster along a small set of recurring CNAME-based paths, whereas malicious manipulations typically disrupt these paths by introducing unexpected aliases or redirecting queries to unrelated IPs. This insight motivates a resolution-path-based methodology: rather than treating DNS resolution as a flat mapping from domain to IP, we model the process



(a) Normal Resolution Paths. (b) Manipulated Resolution Path.

(c) Manipulated path extract.

Figure 2: Overview of manipulation measurement.

as a structured graph that captures both legitimate dynamics and deviations indicative of manipulation.

Our approach consists following parts. First, we issue queries to multiple open recursive resolvers for the same domain and collect all intermediate `CNAME`, `NS`, and `A` records to reconstruct each resolver's complete resolution path. These paths are represented as directed graphs, with nodes denoting DNS records and edges encoding referral or alias relationships (Section 3.2). We further enrich the graph by embedding metadata derived from the resolution process, including authoritative `NS` domains, organizations, and corresponding ASNs. Nodes sharing identical metadata are indirectly linked through multi-hop connections, allowing structural correlation among different resolution paths. We identify potential manipulated paths that exhibit weaker inter-path associations, and distinguish legitimate path diversity from malicious manipulations (Section 3.3). Finally, we conduct large-scale in-the-wild measurements, applying our methodology to representative domains and open resolvers worldwide to ensure coverage and representativeness (Section 3.4). Our approach builds upon a key assumption that manipulated resolutions represent only a minor fraction of overall DNS traffic. This assumption naturally holds in large-scale measurements and has been widely corroborated by existing work [25] and our large-scale measurement in Section 4.1.

3.2 Resolution Path Collection

We conduct our measurements by simultaneously probing a large set of DNS resolvers, targeting one domain at a time. We issue

DNS queries of multiple record types (NS, CNAME, A) to each resolver from vantage points. By recursively querying until termination—obtaining an A record—the full resolution chain for the domain is revealed. For each resolver and each domain, we collect a sequence of responses forming a resolution chain:

$$\text{Domain} \rightarrow \text{CNAME}_1 \rightarrow \ldots \rightarrow \text{CNAME}_k \rightarrow \text{A}$$

The collected results are illustrated in Fig. 2(a) and Fig. 2(b). Our method models the DNS resolution process using four types of nodes: **D**, representing the original queried domain; **C**, representing all intermediate CNAME records; **A**, representing the final resolved IP addresses; and **M**, representing metadata associated with the other node types. Nodes **M** capture various attributes: for **D** and **C** nodes, **M** includes their authoritative NS records and associated organizations, while for **A** nodes, **M** contains the IP's ASN and organization. For example, if an **A** node belongs to a particular organization (ORG), ORG is added as an **M** node and linked to the corresponding **A** node. The NS records are obtained directly from our probes, and other information is retrieved via Whois [8] and IPInfo [1]. The **D**, **C**, and **A** nodes are connected as a directed chain capturing the main resolution process, while the **M** nodes are linked bidirectionally to all **D**, **C**, and **A** nodes, enriching the graph with authoritative and organizational context.

Aggregating all resolution chains of the same domain constructs a domain-level resolution graph $G = (V, E)$, which reveals both structural diversity across resolvers and anomalous branches caused by manipulation. As shown in Fig. 2(a), the normal path allows multiple resolution paths to coexist, but new CNAMEs often reconnect through metadata to existing nodes (e.g., within the same organization), reflecting legitimate DNS dynamics. In contrast, the manipulated path in Fig. 2(b) displays divergent structures, such as unexpected CNAME insertions or redirections to unrelated IPs, which stand out as clear deviations from the normal resolution process.

3.3 Manipulated Paths Extract

After constructing the resolution graphs, the next challenge is to identify manipulated paths. This is nontrivial, as the graph of a single domain may contain thousands of resolution paths (as Section 3.4). As shown in Fig. 2(c), manipulated paths often introduce A or CNAME nodes lacking connections to the original resolution chain, thereby reducing linkage with other paths through metadata nodes. In the graph, this is reflected by longer node-to-node distances or an increased number of unreachable nodes.

A key requirement is a quantitative metric for assessing the "outlier-ness" of each path. We introduce **Path Divergence**, which measures structural dissimilarity between paths. In the resolution graph G, we extract resolution paths P_i from the origin domain to the terminal IP nodes. The path divergence $D(P_a, P_b)$ between two paths $P_a = (u_1, \ldots, u_p)$ and $P_b = (v_1, \ldots, v_q)$ is defined as the average pairwise shortest-path length in G:

$$D(P_a, P_b) = \frac{1}{p \cdot q} \sum_{u \in P_a} \sum_{v \in P_b} \big(d_G(u, v) + d_G(v, u)\big),$$

where $d_G(x, y)$ denotes the shortest-path length (hops) from node x to node y[1]. When $P_a = P_b$, $D(P_a, P_b)$ is set as 0. Normalization by $p \cdot q$ ensures comparability across paths of different lengths. Furthermore, since our resolution chains are directional, changes closer to the terminal nodes induce larger distances, naturally emphasizing anomalies that occur near the end of the path. At this time, we have obtained the pairwise divergences between all paths in the graph, forming a divergence matrix as Fig. 2(c).

The next step is to identify outliers, which correspond to paths with a high likelihood of being manipulated. To enable large-scale and lightweight detection, we adopt a statistical approach. We define an anomaly score for each path P_i as the weighted distance to all other paths in the set $\mathcal{P} = \{P_1, \ldots, P_n\}$:

$$S_i = \sum_{j=1}^{n} N_j \cdot D(P_i, P_j),$$

where N_j corresponds to the number of resolvers N_i that share path P_j, the corresponding number of resolvers for each path is $\mathbf{N} = (N_1, N_2, \ldots, N_n)$. To identify outliers, we perform a weighted interquartile range (IQR) analysis. Let the anomaly score vector be $\mathbf{S} = (S_1, S_2, \ldots, S_n)$, we define the cumulative weight function as:

$$F(s) = \frac{\sum_{i:S_i \le s} N_i}{\sum_{i=1}^{n} N_i}.$$

Then the first quartile Q_1 and third quartile Q_3 are given by:

$$Q_1 = \inf\{s \mid F(s) \ge 0.25\}, \qquad Q_3 = \inf\{s \mid F(s) \ge 0.75\}.$$

A path is flagged as potentially manipulated if its score satisfies $S_i > Q_3 + \alpha \cdot (Q_3 - Q_1)$, where α is a tunable threshold. We set α as 1.5, following standard practice [32].

3.4 Large-Scale Measurement

To obtain a comprehensive view of DNS manipulation in the wild, we target global DNS infrastructures through active measurements. While probing all resolvers and domains would be ideal, such exhaustive exploration is infeasible due to the vast search space and ethical constraints. We therefore adopt a representative sampling strategy that balances coverage, diversity, and feasibility. Specifically, we select a representative set of DNS resolvers and a diverse set of popular domains, enabling us to capture the structural and operational characteristics of the global DNS while keeping large-scale active measurements practical.

Representative Domain Selection. Unlike prior studies that focus on censorship-sensitive domains, we target popular services with broad societal impact. We selected the top domains across web application categories according to global popularity rankings [7, 18, 29]. To enrich coverage, we expand beyond second-level domains by incorporating multiple subdomain prefixes, including the top-five [36] popular prefixes (*e.g.,* m, mynet) and five semantically meaningful variants (*e.g.,* login, mail), resulting in a final set of 2,283 popular domains. Fig. 3 illustrates the distribution of the number of distinct resolution paths observed across different categories and prefixes in our measurement. Most domains exhibit high dynamism and diversity, with individual domains yielding up to 10k distinct resolution paths. Further information on domain prefixes and categories is provided in Appendix B.

Representative DNS Resolver Selection. Our resolver selection strategy builds on recent insights into the client-side structure of

[1]We assign a large constant value when x is unreachable from y, set as 10 here.

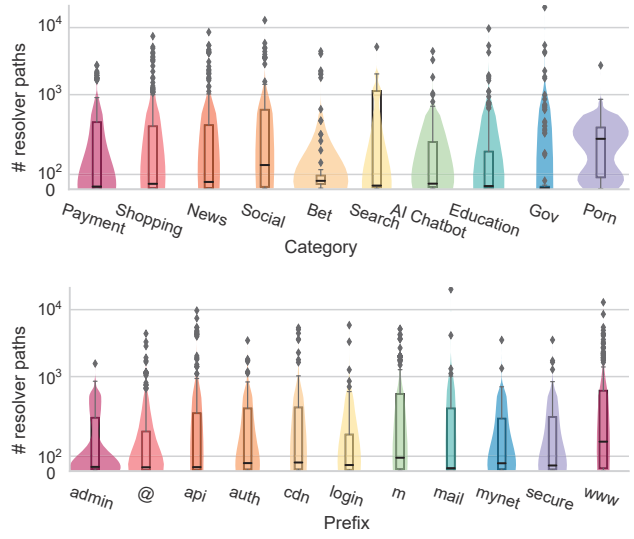
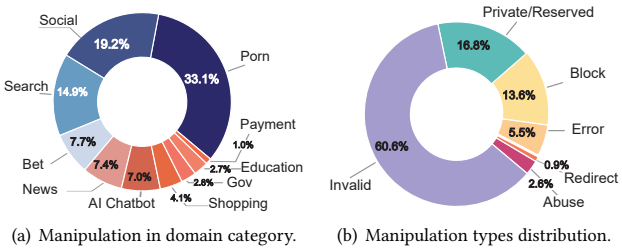


Figure 3: Domains and resolution path count. Most domains exhibit dynamic behaviors with multiple paths.



(a) Manipulation in domain category. (b) Manipulation types distribution.

Figure 4: Overview of DNS response manipulation.

open DNS resolvers [35]. DNS resolvers are unevenly distributed and form ODNS clusters representing upstream servers and their dependent forwarders. While prior work randomly sampled resolvers or directly selected resolvers from public DNS or ISP DNS [25], we sample proportionally according to ODNS clusters while guaranteeing at least two resolvers per large cluster. We focus on DNS clusters with at least 100 resolvers, as small clusters typically correspond to personal or small-scale deployments with limited global impact. Large clusters often span multiple ASNs or organizations and reflect operational infrastructures of ISPs or public DNS providers. This yields a representative set of approximately 10,000 resolvers, balancing coverage and efficiency. The details of our sample strategy and the mapping of selected resolvers to their location and ASN information are provided in the Appendix C.

Measurement Workflow. Based on the above design, we deploy vantage points on cloud servers in Malaysia, the United States, China, and Finland to issue probes to the selected set of 10k resolvers distributed worldwide. For each selected domain, queries are issued in parallel to all sampled resolvers to collect `CNAME`, `NS`, and `A` records. Finally, we identify potential DNS manipulation for each domain.

4 Measurements & Analysis

In this section, we conducted measurements to uncover the prevalence and characteristics of DNS manipulation in the wild. We verify our method with small-scale validation, the detailed description can be found in Appendix D. Comparing with existing works [5, 25, 30], our approach achieves fewer false positives and reducing the analysis burden for large-scale measurements. Then, we move toward a global view of DNS resolvers with popular domains.

4.1 Overview of Response Manipulation

Our large-scale measurement spans 2,283 domains across resolvers worldwide. We define the unit of manipulation as a *domain-resolver pair*, where one domain manipulated by one resolver contributes one manipulation. We carefully processed the measurement data and summarized the determined manipulations. Overall, we observed 1,042,360 manipulations, accounting about 4.5% of all probes issued. We further illustrate the distribution of manipulations across domain categories in Fig. 4(a). Specific domain categories, such as porn and social media, exhibit higher fraction of manipulations. These patterns indicate that manipulations are neither random nor uniform; Manipulation operators appear to selectively target domains based on perceived value, sensitivity, or risk.

We next analyze the results of our global-scale measurement. Based on the service hosted on the destination IPs in manipulated responses and corroborated via certificate checks, we classify manipulation behaviors into six distinct types:

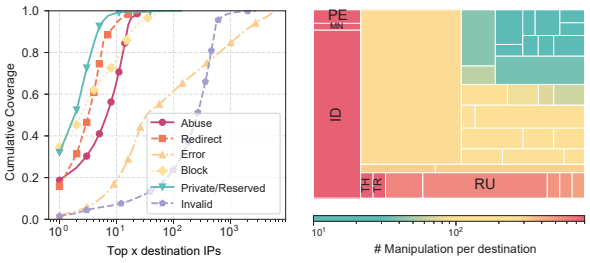
(1) **Abuse**: Response containing IPs reported for malicious activities [2], including malware, spam, or other abuse.
(2) **Redirect**: Response containing a third-party web interfaces that redirect users to unrelated services, such as captive portals or advertising pages.
(3) **Error**: Response containing IPs with incorrect service responses, e.g., expired domain pages, or default server pages like Ubuntu configuration page.
(4) **Block**: Response containing destination IPs directed to explicit blocking pages, such as security DNS warnings or national-level alert pages.
(5) **Private/Reserved**: Responses containing private or invalid IP addresses, corresponding to unallocated ranges.
(6) **Invalid**: IPs lacking any active web service, making the original service unavailable.

Fig. 4(b) presents the distribution of manipulation types, illustrating the prevalence of different manipulation behaviors in the wild. The most frequent category is **Invalid**, followed by **Private/Reserved** and **Block**, which aim to prevent service rather than executing targeted attacks. In contrast, **Error**, **Abuse**, and **Redirect** manipulations, though in smaller amounts, are of greater concern due to their potential risk to end users and the service.

4.2 Deep Dive into Manipulation

To better understand the strategies behind DNS manipulation, we organize our deep-dive analysis along four questions: where our queries are directed, how resolution chains are altered, which resolvers are involved, and which domain categories are most affected.

(A) Where do manipulations lead us: Destination Analysis.



(a) Destination IP coverage (b) Destination IP location

Figure 5: Destination IP concentration.

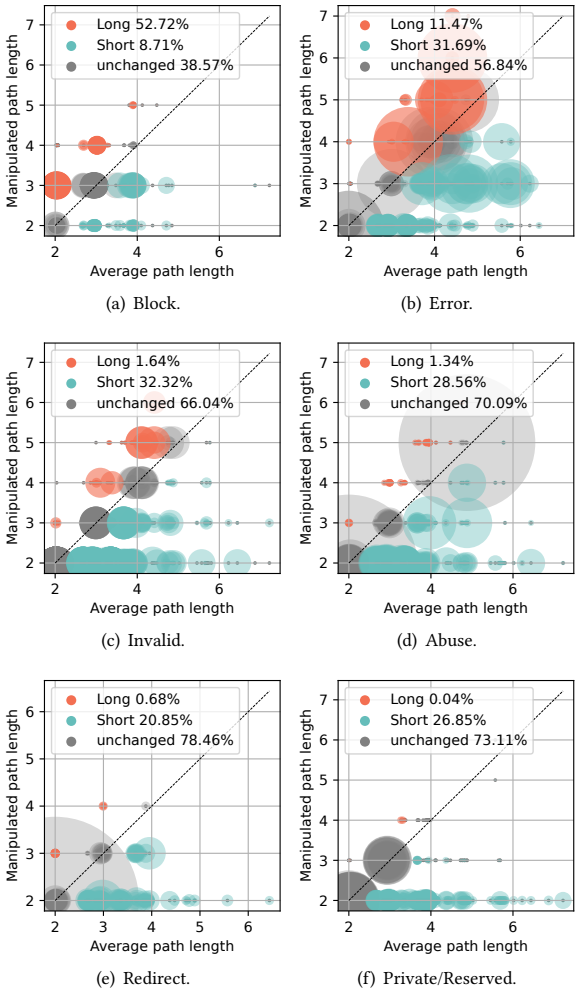
To understand the impact of DNS manipulation, we first examine the destination IPs of manipulated responses. We reveal how manipulations are concentrated across a small set of endpoints and highlight patterns in geographic deployment that shape the overall effect on users. Fig. 5(a) illustrates the concentration of destination IPs across different manipulation types. A clear preference for certain destination IPs is evident: **Abuse**, **Block**, and **Redirect** manipulations are highly concentrated, with nearly all affected queries directed to fewer than 100 IPs. In contrast, **Error** and **Invalid** destinations are more dispersed. Although **Private** addresses are also concentrated, their effect is limited and largely inconsequential for understanding targeted manipulation.

Fig. 5(b) summarizes the geographic distribution of the destinations for **Abuse**, **Block**, and **Redirect** manipulations. We provide detailed information on these destination IPs in Appendix E. We observe that **Abuse** and **Redirect** manipulation often concentrate on a single IP or very few IPs, indicating extreme centralization. By contrast, some **Block** destinations span multiple IPs, reflecting the deployment of explicit blocking interfaces in certain countries.

> **Takeaway:** High-risk manipulations (*i.e.,* Abuse and Redirect) concentrate responses on a small, centralized set of IP addresses, directing users to third-party websites or compromised hosts.

(B) How manipulations are carried out: Chain Alterations. To answer how manipulators implement their strategies, we examine the DNS resolution chains of manipulated responses. By comparing the length and structure of these paths with benign resolutions, we can identify systematic alterations that reveal both operational and strategic objectives. Fig. 6 shows the length of the resolution chain in the manipulated responses under six manipulation types and the average path length in the normal responses of this domain name. The larger the size of the point indicates the higher number of overlapping manipulations for that case.

A prominent pattern emerges in high-risk manipulation categories, such as Abuse IPs (Fig. 6(d)) and Redirect (Fig.6(e)). These manipulations typically shorten the resolution chain, bypassing intermediate CNAMEs and redirecting queries to malicious endpoints. Most shortened chains span only two to three hops. Manipulators generally avoid introducing new CNAMEs and instead replace an existing CNAME with their own A record. As a result, resolution often terminates prematurely, producing a shorter chain. A concrete



(a) Block. (b) Error.
(c) Invalid. (d) Abuse.
(e) Redirect. (f) Private/Reserved.

Figure 6: Chain Alterations.

case illustrating this behavior for an Abuse instance is presented in Section 4.3. Additionally, Private and Invalid responses also tend to shorten chains as shown in Fig. 6(f) and Fig. 6(c), reflecting a coarse-grained method of denying access. Such reductions in chain provide a potential signal for identifying such manipulation. In contrast, Block manipulations often lengthen the resolution chain by adding extra CNAME redirections as Fig. 6(a), while Error manipulations (Fig. 6(b)) exhibit irregular and less consistent patterns.

> **Takeaway:** High-risk manipulations (*i.e.,* Abuse and Redirect) tend to shorten resolution chains, funneling queries toward malicious endpoints.

(C) Who is carrying out manipulation: Resolver Behaviors. Identifying which resolver-level behavior is essential for understanding manipulation strategies. As shown in Fig. 8, our measurements reveal that manipulation is frequently partial, which means manipulation affects only a subset of domains handled by a given

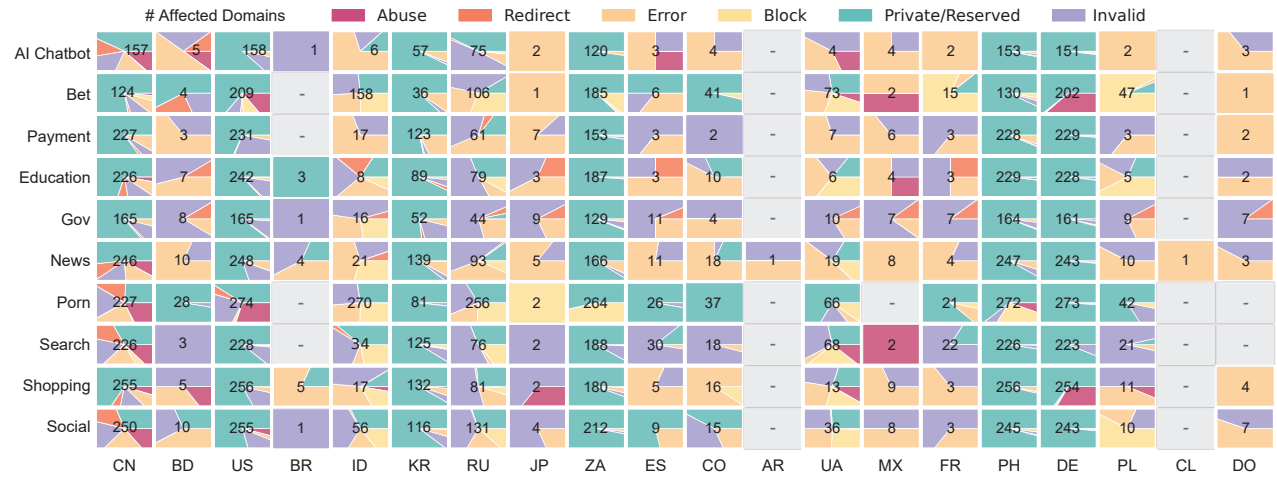


Figure 7: Geographic distribution of manipulating resolvers. Each cell represents a country–category pair, with the number is the count of affected domains, and the pie chart shows manipulation types. Countries are ordered by total number of resolvers.

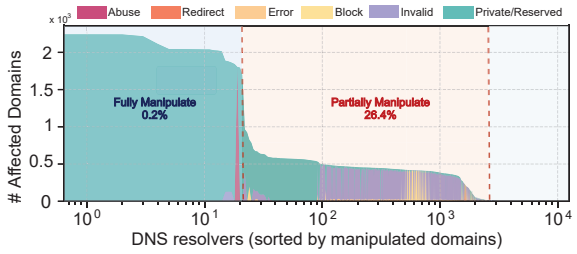


Figure 8: The number of affected domains per DNS server.

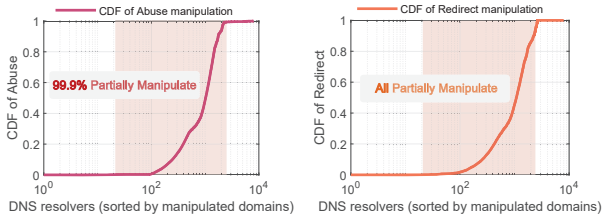


Figure 9: The number of affected domains per DNS server for Abuse and Redirect.

resolver. Approximately 26.4% of resolvers exhibit this behavior. Some of these patterns align with characteristics of national firewalls, yet similar features are also observed in resolvers conducting their own malicious manipulations as shown in Fig. 9, targeting only a very small set of domains (we present a concrete example in Section 4.3). We find that high-risk manipulation (Abuse and Redirect) almost universally exhibits partially manipulation behavior. Partial manipulation allows resolvers to selectively affect specific domains or categories while maintaining normal resolution for others, thereby reducing the likelihood of being detected.

At a broader scale, Fig. 7 illustrates the country-level manipulation of resolvers, revealing substantial heterogeneity across regions. In large, infrastructure-rich countries such as China (CN) and the United States (US), DNS responses exhibit complex manipulation behaviors. Content blocking is not limited to simple block pages but often involves more sophisticated tactics, such as returning invalid IP addresses or inconsistent responses [25]. In contrast, countries such as Bangladesh (BD) experience fewer manipulations. This can be attributed to the fact that, although Bangladesh hosts a large number of DNS resolvers, many of them do not perform recursion locally but instead forward queries to Google Public DNS [35]. In some regions (e.g., Indonesia and Russia), resolvers predominantly return blocking pages for regulated content, whereas in others (e.g., Mexico and Spain), a small number of domains are redirected to malicious IPs. These regional variations reflect the influence of local regulatory environments and the technical capabilities of infrastructure operators. Such patterns suggest that manipulation strategies dynamically adapt to both the targeted domains and the geographic or administrative context of the resolvers.

> **Takeaway:** Most resolvers selectively target specific domains to hide manipulation activity (especially in Abuse and Redirect), and behaviors vary significantly across countries.

(D) Who is affected by manipulation: Domain-Level Analysis. After examining the geographic scope of manipulation, we now turn to which domains are most at risk. To account for uneven domain volumes across categories, we normalize manipulations by domain counts (shown in Appendix B), so the results reflect the average number of manipulations per domain within each category and prefix. As shown in Fig. 10, manipulation behaviors are shaped more by domain category than by prefix. Within domain categories, Porn and Bet domains stand out due to extensive blocking, whereas in most other categories, manipulations primarily involve service denial through invalid or private IP responses. Prefix-level patterns, though less prominent, reveal nuanced distinctions: classic prefixes such as "www" remain the primary targets for redirection attacks.

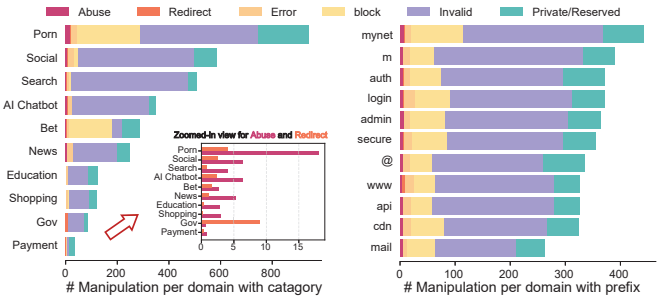


Figure 10: Manipulation distribution in domains.

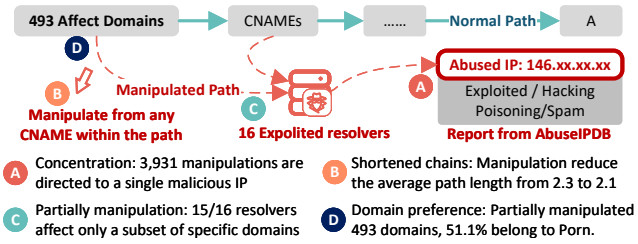


Figure 11: Case study: Manipulations of an abused IP.

High-risk manipulations (Abuse and Redirect), display distinct targeting preferences. As illustrated in the zoomed-in view of Fig. 10, our analysis shows that porn and government-related domains are the most frequently targeted. Manipulations directing domains to abuse-related IPs are predominantly concentrated on Porn sites, whereas Redirect exhibit a notable bias toward government domains. Together, these findings indicate that manipulations are applied selectively and strategically.

Takeaway: Manipulators focus primarily on domain category when targeting domains. High-risk manipulations (Abuse and Redirect) show strong, category-specific preferences.

4.3 Case Study

To show how macro patterns observed at scale translate into real-world behaviors, we conduct a detailed case study highlighting the role of abuse IPs in real-world DNS manipulations. Specifically, we identified an abused IP with the data from AbuseIPDB [2]. This abused IP has been repeatedly reported as an exploited host (18 reports), involved in hacking (25 reports), and associated with poisoning, spam, and brute-force activities (3 reports).

As shown in Fig. 11, we collected all manipulation cases in our measurement results where responses were redirected to this IP, enabling us to observe the following observations: (A) Concentration: 3,931 manipulations for domains converge on this single malicious IP, demonstrating the high concentration of targeted responses; (B) Shortened chains: Manipulators alter intermediate resolution steps so that queries ultimately point to such an abusive IP. Manipulation from an intermediate CNAME reduces the average resolution path length from 2.3 to 2.1; (C) Partially manipulation: Nearly all affected resolvers (15/16) perform manipulation selectively, returning altered responses only for a subset of domains while resolving others normally; (D) Domain preference: Of the 493 partially manipulated domains, 51.1% belong to the Porn category and 9.5% to Bet, reflecting deliberate targeting bias. These findings confirm that partial, selective manipulation is strategic rather than incidental, aligning with the global patterns illustrated in Section 4.2.

5 Discussion

DNS security in the wild. Open DNS resolvers continue to pose significant security concerns. We recommend prioritizing large, trusted public resolvers whenever possible, even though globally distributed open DNS resolvers may provide performance benefits. Many resolvers exhibit complex behaviors, including partial or selective manipulation, which complicates monitoring. These patterns introduce substantial risks to individual users and the broader DNS ecosystem, underscoring the need for careful assessment of resolver trustworthiness and systematic monitoring.

Challenges in detecting malicious behaviors. Detection approaches that rely on single-purpose or dedicated test domains may be insufficient. Malicious operators often design manipulations to evade large-scale measurement campaigns, selectively targeting specific domains or content categories. This selective targeting complicates comprehensive detection and results in variations when applying existing measurement results to new domains, underscoring the need for adaptive, representative measurement strategies and continuous reassessment of threat coverage.

Implications for manipulation monitoring. The concentration of malicious destinations enables us to effectively avoid known manipulations by blocking a small set of specific IPs. For unknown manipulations, shortened CNAME chains provide strong signals of potential high-risk manipulation, indicating redirection to attacker-controlled endpoints directly. Systematically monitoring and analyzing these structural features in resolution paths can improve the timely detection of high-risk manipulations and inform mitigation strategies for both operators and security researchers.

6 Conclusion

DNS manipulations pose critical security risks, yet their global scope and mechanisms remain poorly understood. We conduct large-scale measurements that uncover key insights into global manipulation. We identify patterns in high-risk DNS manipulations and highlight key findings, including their behavior in destination selection, chain alteration, partial manipulation, and domain preference. Our findings provide insight into their operational and strategic behaviors, informing future detection, monitoring, and mitigation efforts. Still, our results represent only the tip of the iceberg of global DNS manipulation, and we hope our work serves as a useful reference for those who come after.

Acknowledgment

We thank all the reviewers for their insightful comments. This work is supported by National Key R&D Program of China (Grant No. 2022YFB3103000).

References

[1] 2024. IP Info. https://ipinfo.io.

[2] AbuseIPDB. 2025. AbuseIPDB: Making the Internet Safer, One IP at a Time. https://www.abuseipdb.com/

[3] Gautam Akiwate, Raffaele Sommese, Mattijs Jonker, Zakir Durumeric, KC Claffy, Geoffrey M Voelker, and Stefan Savage. 2022. Retroactive identification of targeted DNS infrastructure hijacking. In *Proceedings of the 22nd ACM Internet Measurement Conference*. 14–32.

[4] Zakaria Benomar, Francesco Longo, Giovanni Merlino, and Antonio Puliafito. 2021. A cloud-based and dynamic dns approach to enable the web of things. *IEEE Transactions on Network Science and Engineering* 9, 6 (2021), 3968–3978.

[5] Andreas Borgwart, Spyros Boukoros, Haya Shulman, Carel van Rooyen, and Michael Waidner. 2015. Detection and forensics of domains hijacking. In *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 1–6.

[6] Taejoong Chung, Dave Levin, and Protick Bhowmick. 2025. Reliable and Decentralized Certificate Revocation via DNS: The Case for RevDNS. In *Proceedings of the ACM SIGCOMM 2025 Conference*. 882–895.

[7] Cloudflare, Inc. 2025. Cloudflare Radar. https://radar.cloudflare.com/. Accessed: 2025-10-04.

[8] Leslie Daigle. 2004. *WHOIS protocol specification*. Technical Report.

[9] José Luis García-Dorado, Javier Ramos, Miguel Rodríguez, and Javier Aracil. 2018. DNS weighted footprints for web browsing analytics. *Journal of Network and Computer Applications* 111 (2018), 35–48.

[10] Muks Hirani, Sarah Jones, and Ben Read. 2019. Global DNS hijacking campaign: DNS record manipulation at scale. *research/2019/01/global-dns-hijacking-campaign-dns-recordmanipulation-at-scale. html* (2019).

[11] Rebekah Houser, Shuai Hao, Zhou Li, Daiping Liu, Chase Cotton, and Haining Wang. 2021. A comprehensive measurement-based investigation of DNS hijacking. In *2021 40th International Symposium on Reliable Distributed Systems (SRDS)*. IEEE, 210–221.

[12] Jinyuan Jia, Zheng Dong, Jie Li, and Jack W Stokes. 2021. Detection of malicious dns and web servers using graph-based approaches. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2625–2629.

[13] Joseph B. Kruskal and Myron Wish. 1978. *Multidimensional Scaling*. Sage Publications.

[14] Marc Kührer, Thomas Hupperich, Jonas Bushart, Christian Rossow, and Thorsten Holz. 2015. Going wild: Large-scale classification of open DNS resolvers. In *Proceedings of the 2015 Internet Measurement Conference*. 355–368.

[15] James Larisch, Timothy Alberdingk Thijm, Suleman Ahmad, Peter Wu, Tom Arnfeld, and Marwan Fayed. 2024. Topaz: Declarative and verifiable authoritative DNS at CDN-scale. In *Proceedings of the ACM SIGCOMM 2024 Conference*. 891–903.

[16] Xiang Li, Chaoyi Lu, Baojun Liu, Qifan Zhang, Zhou Li, Haixin Duan, and Qi Li. 2023. The maginot line: Attacking the boundary of {DNS} caching protection. In *32nd USENIX Security Symposium (USENIX Security 23)*. 3153–3170.

[17] 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 18)*. 1113–1128.

[18] Majestic. 2025. *The Majestic Million: Top 1 Million Websites*. https://majestic.com/reports/majestic-million

[19] Keyu Man, Zhiyun Qian, Zhongjie Wang, Xiaofeng Zheng, Youjun Huang, and Haixin Duan. 2020. Dns cache poisoning attack reloaded: Revolutions with side channels. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. 1337–1350.

[20] Andrew McGregor, Phillipa Gill, and Nicholas Weaver. 2021. Cache me outside: A new look at DNS cache probing. In *Proc. Passive and Active Measurement Conference (PAM). Virtual*. 427–443.

[21] Yevheniya Nosyk, Qasim Lone, Yury Zhauniarovich, Carlos H Gañán, Emile Aben, Giovane CM Moura, Samaneh Tajalizadehkhoob, Andrzej Duda, and Maciej Korczyński. 2023. Intercept and inject: DNS response manipulation in the wild. In *International Conference on Passive and Active Network Measurement*. Springer, 461–478.

[22] John S Otto, Mario A Sánchez, John P Rula, and Fabián E Bustamante. 2012. Content delivery and the natural evolution of DNS: remote DNS trends, performance issues and alternative solutions. In *Proceedings of the 2012 Internet Measurement Conference*. 523–536.

[23] Jianping Pan, Y Thomas Hou, and Bo Li. 2003. An overview of DNS-based server selections in content distribution networks. *Computer Networks* 43, 6 (2003), 695–711.

[24] Jeman Park, Rhongho Jang, Manar Mohaisen, and David Mohaisen. 2021. A large-scale behavioral analysis of the open DNS resolvers on the internet. *IEEE/ACM Transactions on Networking* 30, 1 (2021), 76–89.

[25] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nick Weaver, and Vern Paxson. 2017. Global measurement of {DNS} manipulation. In *26th USENIX Security Symposium (USENIX Security 17)*. 307–323.

[26] 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 *Proceedings of the 21st ACM Internet Measurement Conference*. 390–397.

[27] Giovanni Schmid. 2021. Thirty years of DNS insecurity: Current issues and perspectives. *IEEE Communications Surveys & Tutorials* 23, 4 (2021), 2429–2459.

[28] Kyle Schomp, Tom Callahan, Michael Rabinovich, and Mark Allman. 2013. On measuring the client-side DNS infrastructure. In *Proceedings of the 2013 conference on Internet measurement conference*. 77–90.

[29] SimilarWeb. 2025. . https://www.similarweb.com/zh/top-websites

[30] Elisa Tsai, Deepak Kumar, Ram Sundara Raman, Gavin Li, Yael Eiger, and Roya Ensafi. 2023. Certainty: Detecting dns manipulation at scale using tls certificates. *arXiv preprint arXiv:2305.08189* (2023).

[31] Adrian-Victor Vevera, Andreea Cătălina CRĂCIUN, Mihail DUMITRACHE, Ionut SANDU, Carmen-Ionela ROTUNĂ, and Radu Alexandru BOSTAN. 2025. Cybersecurity Challenges in Managing Domain Names. From DNS to ENS in the Web3 Era. *Romanian Cyber Security Journal* 7, 1 (2025), 97–112.

[32] HP Vinutha, B Poornima, and BM Sagar. 2018. Detection of outliers using interquartile range technique from intrusion dataset. In *Information and decision sciences: Proceedings of the 6th international conference on ficta*. Springer, 511–518.

[33] Michael Walfish, Hari Balakrishnan, and Scott Shenker. 2004. Untangling the Web from DNS.. In *NSDI*, Vol. 4. 17–17.

[34] Yue Wang, Changhua Pei, Zexin Wang, Yingqiang Wang, Guo Chen, Yuchao Zhang, Yi Li, Jingjing Li, Jianhui Li, and Gaogang Xie. 2024. ActiveDNS: Is There Room for DNS Optimization Beyond CDNs?. In *2024 IEEE 49th Conference on Local Computer Networks (LCN)*. IEEE, 1–9.

[35] Wenhao Wu, Zhaohua Wang, Qinxin Li, Zihan Li, Yi Li, Jin Yan, and Zhenyu Li. 2025. ODNS Clustering: Unveiling Client-Side Dependency in Open DNS Infrastructure. In *Proceedings of the ACM on Web Conference 2025*. 4745–4754.

[36] Youwei Xu. 2024. *Understanding and Characterizing CDN Services and Paid Features*. Master's thesis. University of Twente.

[37] Yunyi Zhang, Mingming Zhang, Baojun Liu, Zhan Liu, Jia Zhang, Haixin Duan, Min Zhang, Fan Shi, and Chengxi Xu. 2024. Cross the Zone: Toward a Covert Domain Hijacking via Shared {DNS} Infrastructure. In *33rd USENIX Security Symposium (USENIX Security 24)*. 5751–5768.

A Ethics

All measurements were carefully designed to avoid disrupting normal resolver service. Each vantage point issued only a small number of queries each time. Specifically, measurement was completed over three days, ensuring that each resolver proceeded at most one domain per two minutes. Following best practices for responsible measurement, we deployed a web front at each vantage point with our contact information and the purpose of the study as an active opt-out mechanism. Additionally, any resolver that failed to respond consistently was excluded from further measurements. Throughout the study, we have not received any requests to quit our measurements or any abuse complaints.

B Domain Selection

We selected the top domains across web application categories according to global popularity rankings [7, 18, 29]. To enrich coverage, we expand beyond popular domains by incorporating multiple subdomain prefixes, including the top-five [36] popular prefixes (*e.g.,* m, mynet) and five semantically meaningful variants (*e.g.,* login, mail, auth), resulting in a final set of 2,283 popular domains. Table 1 summarizes the distribution of domains in prefixes and service categories included in our measurement.

C DNS Resolver Selection

Our resolver selection strategy builds upon recent insights into the client-side structure of open DNS resolvers (ODNS Clustering [35]). To the best of our knowledge, we are also the first to leverage this insight for measurement purposes. The client-side DNS structure is defined as *ODNS clusters*, referring to the collection of upstream servers and forwarders that have direct (or indirect) dependencies.

Table 1: Distribution of domains in prefixes and categories.

Prefix	# Domains	Category	# Domains
@	488	Shopping	264
www	467	Payment	231
secure	103	Government	167
m	236	News	251
api	272	Porn	274
login	114	Education	244
mail	168	Search	229
auth	127	Social medium	256
mynet	68	AI chatbot	158
cdn	130	Bet	209
admin	110	-	-

Table 2: Distribution of DNS resolvers in top-20 countries.

Country	# Resolvers	Country	# Resolvers
CN	1547	CO	209
BD	980	AR	194
US	927	UA	182
BR	709	MX	174
ID	650	FR	137
KR	591	PH	132
RU	479	DE	122
JP	303	PL	117
ZA	282	CL	113
ES	278	DO	113

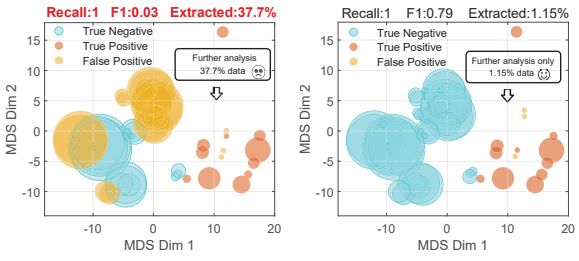
This study has revealed extensive dependencies within the DNS space, meaning that if we randomly sample DNS servers in the wild, we may end up selecting multiple servers that rely on the same upstream DNS: the scope may **remain limited** to the same upstream recursive resolver. By selecting representative resolvers from different ODNS clusters, we can achieve the maximum measurement coverage with minimal overhead. We replicated the ODNS clustering results and used Zmap to get global open resolvers, uncovering a total of 72,376 clusters worldwide.

While prior work randomly sampled popular resolvers or select DNS resolvers that identified as infrastructure DNS resolvers [25], we sample proportionally according to ODNS clusters while guaranteeing at least two resolvers per large cluster (with size over 100). We focus on clusters with at least 100 resolvers, as small clusters typically correspond to personal or small-scale deployments with limited global impact. Large clusters often span multiple ASNs or organizations and reflect infrastructures of ISPs or public DNS providers. This yields a representative set of approximately 10,000 resolvers, balancing coverage and efficiency. Table 2 shows the geographic distribution of DNS resolvers included in our study, which covers a broad range of 120 countries worldwide. China (CN) has the largest number of resolvers (1,547), followed by Bangladesh (BD, 980) and the United States (US, 927). Other countries, including Brazil (BR), Indonesia (ID), South Korea (KR), and Russia (RU), also contribute notable numbers.

D Method Validation

We validated our methodology by two approaches. First, we manually annotated DNS responses for `www.baidu.com`. Baidu directs queries to multiple edge servers in ASNs belonging to different Chinese ISPs, balancing dynamicity and the feasibility of manual annotation, making it an ideal choice for verification. Annotated data included 59 resolution paths shared by 99.2% normal resolvers and 11 manipulated paths corresponding to 0.8% resolvers. Fig. 12 illustrates the comparison between our method and Iris-style [5, 25, 30] approaches. Iris-style approaches provide multiple attributes for comparison, including HTTP content and TLS certificates. However, we find that most of these attributes are not universally applicable in practice. In our measurements, 15% of domains expose neither a TLS certificate nor an HTTP interface due to diverse usage purposes or custom service protocols. This limitation is more pronounced at the prefix level, where 57.7% of domains under the `mail.` prefix and 19.4% under the `secure.` prefix cannot be validated. Consequently, we adopt ASN-based consistency as the comparison baseline, as it relies solely on DNS resolution results and is universally available.

Due to Baidu's deployment across multiple edge servers, correct responses exhibit inherent diversity, leading to high false positives under the Iris-style method, which outputs 37.7% of measurement results as potential manipulation with a low F1-score. Our approach achieves fewer false positives and only extracts 1.15% of the responses as potential manipulation, reducing the analysis burden for large-scale measurements. We further examined the sensitivity of the IQR parameter α under above measurement results. We varied α with a step size of 0.01 and observed that our method exhibits a wide tolerance range: values of α between 1.21 and 2.29 achieve identical performance with an F1-score of 0.79. When α falls below this range, the performance degrades slightly due to increased false positives (e.g., F1= 0.76 at $\alpha = 0.45$), whereas overly large values introduce more false negatives. We therefore adopt a moderate value of $\alpha = 1.5$ to balance sensitivity and robustness. This tolerance arises from our path divergence design and weighted IQR formulation, where legitimate DNS resolution paths are shared by many resolvers and exhibit low divergence, while manipulated paths are purposefully constructed and thus form a clear separation, especially in large-scale measurements.



(a) Iris-style method. (b) Our method.

Figure 12: Comparison with existing methods. Resolution paths are visualized in 2D via MDS [13].

For broader validation across our measurement results, obtaining ground truth in real-world DNS resolution is inherently challenging. To address this, we first identified four IP addresses (Cases

Table 3: Manipulated cases used for validation.

Case	Behavior	Affected / Detected
Case A	Private address	2283/2283
Case B	Abused IP address [2]	219/219
Case C	Block with a block page	308/308
Case D	Block with a block page	417/417

Table 4: Top-10 countries of manipulation.

Country	# Manipulation	# IP	Category
Indonesia	91034	25	Block
Mongolia	1491	1	Block
Peru	2235	2	Redirect
Thailand	942	1	Block
Turkey	857	1	Redirect
Cyprus	1744	3	Block
Russia	5613	10	Block
Ireland	508	1	Abuse
Spain	497	1	Abuse
Austria	463	1	Abuse

A–D) that could be unambiguously attributed to manipulated responses and verified whether these IPs appeared in our large-scale measurement results. Using these four addresses as ground truth, we then evaluated whether our method successfully detected the corresponding manipulations during measurement. As summarized in Table 3, all manipulations associated with these ground truth IPs were correctly identified, demonstrating that our methodology can effectively captures real-world manipulation events.

E Destination IP Analysis

Following the analysis in Section 4.2, Table 4 provides detailed information on destinations IPs for **Abuse**, **Block**, and **Redirect** manipulations. We observe that **Abuse** and **Redirect** manipulation often concentrate on a single IP or very few IPs, indicating extreme centralization. By contrast, some **Block** destinations span multiple IPs, reflecting the deployment of explicit blocking interfaces in certain countries.