# Global-Scale Measurement of DNS Manipulation

PAUL PEARCE, BEN JONES, FRANK LI, ROYA ENSAFI, NICK FEAMSTER, NICHOLAS WEAVER, AND VERN PAXSON

Paul Pearce is a senior PhD student at UC Berkeley advised by Vern Paxson and is a member of the Center for Evidence-based Security Research (CESR). His research brings empirical grounding to Internet security problems, including censorship, cyber crime, and advanced persistent threats (APTs). pearce@cs.berkeley.edu

Ben Jones is a Software Engineer at Google and a PhD candidate at Princeton University. His research is in the area of Internet censorship measurement. He holds a BS in computer science from Clemson University and was an Open Technology Fund Senior Fellow in Information Controls. bj6@cs.princeton.edu

Frank Li is a PhD student at the University of California, Berkeley. His research mainly focuses on improving the remediation process for security issues such as vulnerabilities and misconfigurations. More broadly, he is interested in large-scale network measurements and empirical studies in a computer security context. frankli@cs.berkeley.edu

Roya Ensafi is a Research Assistant Professor in Computer Science and Engineering at the University of Michigan, where her research focuses on computer networking and security. She pioneered the use of side-channels to remotely measure network interference and censorship of Internet traffic. Prior to joining Michigan, she was a postdoc at Princeton University. ensafi@umich.edu

Despite the pervasive and continually evolving nature of Internet censorship, measurements remain comparatively sparse. Understanding the scope, scale, and evolution of Internet censorship requires global measurements, performed at regular intervals. We developed Iris, a scalable, accurate, and ethical method to continually measure global manipulation of DNS resolutions. Iris reveals widespread DNS manipulation of many domain names across numerous countries worldwide.

Anecdotes and reports indicate that Internet censorship is widespread, affecting at least 60 countries [5]. Despite pervasive Internet censorship, empirical Internet measurements revealing the scope of that censorship remain sparse. A more complete understanding of Internet censorship around the world requires *diverse* measurements from a wide range of geographic regions and ISPs, not only across countries but also within regions of a single country.

Unfortunately, most mechanisms for measuring Internet censorship rely on volunteers who run measurement software deployed on their own Internet-connected devices (e.g., laptops, phones, tablets) [9, 10]. Because these tools rely on individuals performing actions, their scale and diversity are fundamentally limited.

Although recent work has developed techniques to continuously measure widespread manipulation at the transport [4, 7] and HTTP [1] layers, a significant gap remains in understanding global information control via manipulation of the Internet's Domain Name System (DNS). Towards this goal, we developed and deploy Iris [8], a method and system to ethically detect, measure, and characterize the manipulation of DNS responses within countries across the entire world—without involving users or volunteers.

Iris aims to provide sound assessments of potential DNS manipulation indicative of an intent to restrict user access to content. To achieve high detection accuracy, we rely on a set of metrics that we base on the underlying properties of DNS domains, resolutions, and infrastructure. Using our implementation of Iris, we performed a global measurement study that highlights the heterogeneity of DNS manipulation across resolvers, domains, and countries—and even within a country.

One significant design challenge concerns ethics. In contrast to systems that explicitly involve volunteers in collecting measurements, methods that perform censorship measurement without volunteers raise the issue of user risk. To this end, Iris ensures that, to the extent possible, we only involve Internet *infrastructure* (e.g., within Internet service providers or cloud hosting providers) in an attempt to minimize the risk to individual users.

## How and What to Measure?

Detecting DNS manipulation is conceptually simple: perform DNS queries through geographically distributed DNS resolvers and analyze the results for "incorrect" responses that indicate manipulation of the answers. Despite this apparent simplicity, realizing a system to scalably collect DNS data and analyze it for manipulation poses significant technical and ethical challenges. Technically, how do we acquire or find global vantage points? Once we

Nick Feamster is a Professor in the Computer Science Department at Princeton University and the Deputy Director of the Princeton University Center for Information Technology Policy (CITP). He received his PhD in computer science from MIT in 2005 and his SB and MEng degrees in electrical engineering and computer science from MIT in 2000 and 2001, respectively. His research focuses on many aspects of computer networking and networked systems, with a focus on network operations, network security, and censorship-resistant communication systems. feamster@cs.princeton.edu

Nicholas Weaver is a Computer Security Researcher at the International Computer Science Institute and a Lecturer in the Computer Science Department at UC Berkeley. nweaver@icsi.berkeley.edu

Vern Paxson is a Professor of Electrical Engineering and Computer Sciences at UC Berkeley and leads the Networking and Security Group at the International Computer Science Institute in Berkeley. His research focuses heavily on measurement-based analysis of network activity and Internet attacks. He works extensively on high performance network monitoring, detection algorithms, cybercrime, and countering censorship and abusive surveillance. vern@berkeley.edu

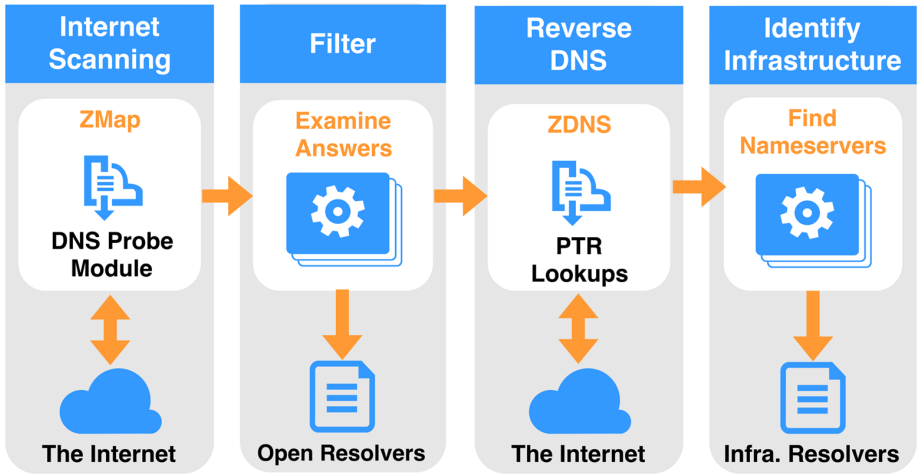


**Figure 1:** Overview of Iris's DNS resolver identification and selection pipeline

have them, what DNS names should we measure? Ethically, how do we conduct wide-ranging third-party measurements without implicating innocent citizens? What steps should we take to ensure that Iris does not induce undue load on the DNS resolution infrastructure?

### *Finding Vantage Points*

To obtain a wide range of measurement vantage points, we leverage *open DNS resolvers* deployed around the world; such resolvers will resolve queries for any client.

Measurement using open DNS resolvers entails complex ethical considerations. Given their prevalence and global diversity, open resolvers offer a compelling resource, providing researchers with extensive volume and reach. Unfortunately, open resolvers also pose risks not only to the Internet but to individual users. For example, resolvers may operate in an open fashion as a result of configuration errors; they frequently operate on end-user devices such as home routers [6]. Using these devices for measurement can impose monetary costs and, if the measurement involves sensitive content or hosts, can also expose their owners to harm.

Due to these ethical considerations, we restrict the set of open resolvers that we use to a small subset of resolvers for which we have high confidence they are part of the Internet infrastructure, as opposed to attributable to particular individuals. Figure 1 illustrates the process by which Iris finds safe open DNS resolvers.

Conceptually, the process of finding infrastructure resolvers has two steps: (1) scanning the Internet for open DNS resolvers and (2) pruning the list of open DNS resolvers that we identify to limit the resolvers to a set that we can plausibly attribute to Internet infrastructure.

**Step 1: Scanning the Internet's IPv4 space for open DNS resolvers.** Scanning the IPv4 address space provides a global perspective on open resolvers. To do so, we developed an open-source module for the ZMap network scanner [3] to enable Internet-wide DNS resolutions. This module queries port 53 of all IPv4 addresses with a recursive DNS A record query. We use a purpose-registered domain name we control for these queries to ensure there is a known correct answer. From these scans, we conclude that all IP addresses that return the correct answer to this query are open resolvers.

**Step 2: Identifying infrastructure resolvers.** We prune the set of all open DNS resolvers on the Internet to include only those resolvers that appear to function as authoritative nameservers for some DNS domain. Iris uses the ZDNS tool to perform reverse DNS PTR lookups

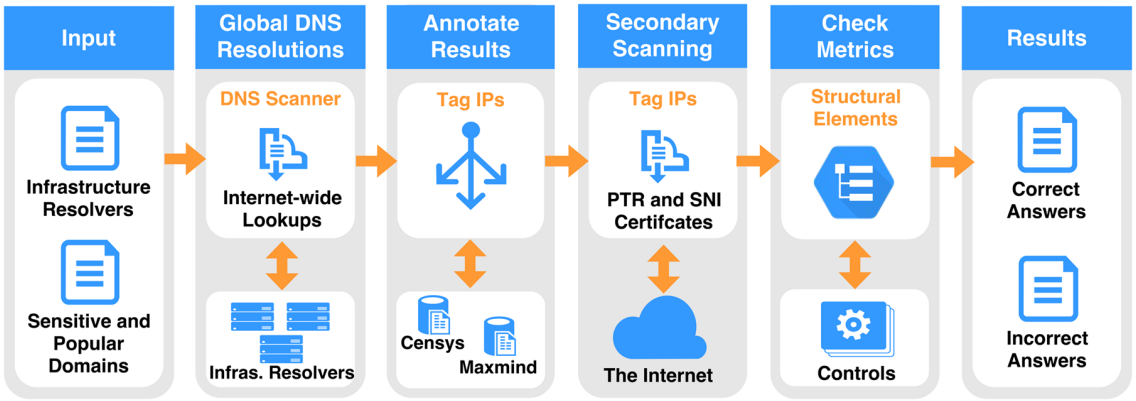


**Figure 2:** Overview of DNS resolution, annotation, filtering, and classification. Iris takes as input a set of domains and DNS resolvers and outputs results indicating manipulated DNS responses.

for all open resolvers and retains only the resolvers that have a valid PTR record beginning with the subdomain `ns[0-9]+` or `nameserver[0-9]*`. This filtering step reduces the number of usable open resolvers from millions to thousands; fortunately, even the remaining set of resolvers suffices to provide broad country- and network-level coverage.

Because we conduct our measurements using resolvers we do not control, we cannot differentiate between countrywide or state-mandated censorship and localized manipulation at individual resolvers (e.g., captive portals or malware [6]). To mitigate this uncertainty, we aggregate our measurements to ISP or country granularity.

### *Ethical Reasoning*

Our primary ethical concern is minimizing the risks associated with issuing DNS queries via open resolvers for potentially censored or manipulated domains, as these DNS queries could implicate innocent users. A second concern is the query load that Iris induces on authoritative nameservers. With these concerns in mind, we use the ethical guidelines of the Belmont Report and Menlo Report to frame our discussion. One important ethical principle is *respect for persons*; essentially, an experiment should respect the rights of humans as autonomous decision-makers. In lieu of attempting to obtain informed consent, we draw upon the principle of *beneficence*, which weighs the benefits of conducting an experiment against the risks associated with the experiment. We note that the benefit of issuing DNS queries through tens of millions of resolvers has rapidly diminishing returns, and that using only open resolvers that we can determine are unlikely to correspond to individual users greatly reduces the risk to any individual without dramatically reducing the benefits of our experiment.

An additional guideline concerns *respect for law and public interest*, which essentially extends the principle of beneficence to all relevant stakeholders, not only the experiment participants. To abide by this principle, we rate-limit our queries for each domain to ensure that the owners of the domains do not face significant expenses as a result of the queries that we issue.

### *Which DNS Domains to Query*

Iris queries a list of sensitive URLs published by Citizen Lab, known as the Citizen Lab Block List (CLBL). This list of URLs is compiled by experts based on known censorship around the world, labeled by category. We distill the URLs down to domain names and use this list as the basis of our data set. We then supplement the list by adding additional domain names selected at random from the Alexa Top 10,000. These additional domain names help address geographic or content biases in the CLBL while maintaining a manageable total number of queries.

## Iris: Systematic Detection of DNS Manipulation

We describe Iris's process for issuing queries for the domains to the set of usable open resolvers. Figure 2 provides an overview of the process. Iris resolves each DNS domain using the global vantage points afforded by the open DNS resolvers; annotates the responses with information from both auxiliary data sets as well as additional active probing; and uses consistency and independent verifiability metrics to identify manipulated responses. A more in-depth treatment of this topic appeared at USENIX Security 2017 [8].

### *Collecting Annotated DNS Responses*

#### Performing Global DNS Resolutions

Iris takes as input a list of suitable open DNS resolvers as well as the combined CLBL and Alexa domain names. We also include three DNS domains under our control to allow us to compute consistency metrics. Querying tens of thousands of domains across tens of thousands of resolvers required the development of a new DNS query tool, because no existing DNS measurement tool supported this scale. We implemented this aspect of Iris

in 649 lines of Go. The tool takes as input a set of domains and resolvers and coordinates randomized querying of each domain across each resolver. To prevent overloading resolvers and domains, we randomize domain order and maintain strict upper bounds on how fast Iris queries individual resolvers.

**Annotating DNS Responses with Auxiliary Information**
Our analysis ultimately relies on characterizing both the *consistency* and the *independent verifiability* of the DNS responses that we receive. To enable this classification, we first must gather additional details about the IP addresses returned in each of the DNS responses. Iris annotates each IP address returned in the set of DNS responses with additional information about geolocation, autonomous system (AS), port 80 HTTP responses, and port 443 HTTPS X.509 certificates. We rely on Censys [2] for daily snapshots of this auxiliary information, and further annotate IP addresses with AS and geolocation information from the MaxMind service, as applicable.

**Additional PTR and TLS Scanning**
For each IP address, we perform a DNS PTR lookup to facilitate additional consistency checks. Complicating inference further, a single IP address might host multiple Web sites via HTTP or HTTPS (virtual hosting). To mitigate this effect, we perform an active HTTPS connection to each returned IP address using the Server Name Indication (SNI) to specify the name originally queried.

### *Identifying DNS Manipulation*
To determine whether a DNS response is manipulated, Iris relies on two metrics: *consistency* and *independent verifiability*.

Unmanipulated access to a domain should manifest some degree of *consistency*, even when accessed from varying global vantage points. The consistency may take the form of network properties, infrastructure attributes, or content. We leverage these attributes, both in relation to control data as well as across the data set itself, to classify DNS responses.

Our consistency metric relies on access to a set of geographically diverse resolvers that we control and are presumably not subject to manipulation. These control resolvers return a set of answers that we can presume are correct and thus can use to identify consistency across a range of IP address properties. Geographic diversity helps ensure that area-specific deployments do not cause false positives. We also use control domains to test whether a resolver reliably returns unmanipulated results for non-sensitive content (e.g., not a captive portal).

For each domain name measured, we create a set of consistency metrics by taking the union of each metric across all of our control resolvers. For example, we consider an answer *consistent* if the IP address matches at least one seen by any of our controls.

In addition to consistency metrics, we also define a set of metrics based on HTTPS certificate infrastructure that we can independently verify using external data sources. This data is collected both from both auxiliary annotations and active HTTPS SNI scans.

We say that a response is correct if it satisfies *any* consistency or independent verifiability metric; otherwise, we classify the response as *manipulated*.

## Global Measurement Study
Using Iris, we performed an end-to-end global measurement study of DNS manipulation during January 2017. Here we describe the basic composition and statistics of this measurement study.

### *Resolvers*
We initially identified a large pool of open DNS resolvers through an Internet-wide ZMap scan using a custom DNS measurement extension to ZMap that we developed. In total, 4.2 million open resolvers responded with a correct answer to our scan queries. This number excludes 670K resolvers that replied with correctly formed DNS responses but with either a missing or an incorrect answer for the scan's query domain.

The degree to which we can investigate DNS manipulation across various countries depends on the geographic distribution of the selected DNS resolvers. By geolocating this initial set of resolvers, we observed that the pool spanned 232 countries and dependent territories, with a median of 659 resolvers per country. Abiding by our ethical considerations reduced this set to 6,564 infrastructure resolvers in 157 countries, with a median of six resolvers per country. Finally, we removed unstable or otherwise errant resolvers, further reducing the set of usable resolvers to 6,020 in 151 countries, again with a median of six resolvers in each. While our final set of resolvers is a small fraction of all open DNS resolvers, it still suffices to provide a global perspective on DNS manipulation.

### *Domains*
We began with the CLBL, consisting of 1,376 sensitive domains. We augmented this list with 1,000 domains randomly selected from the Alexa Top 10,000, as well as the three control domains that we manage that we do not expect to be manipulated. Due to overlap between the two domain sets, our combined data set consisted of 2,330 domains. We excluded 27 problematic domains (e.g., domains that had expired or had broken authoritative name servers) that we identified through our data collection process, resulting in a final set of 2,303 domains.

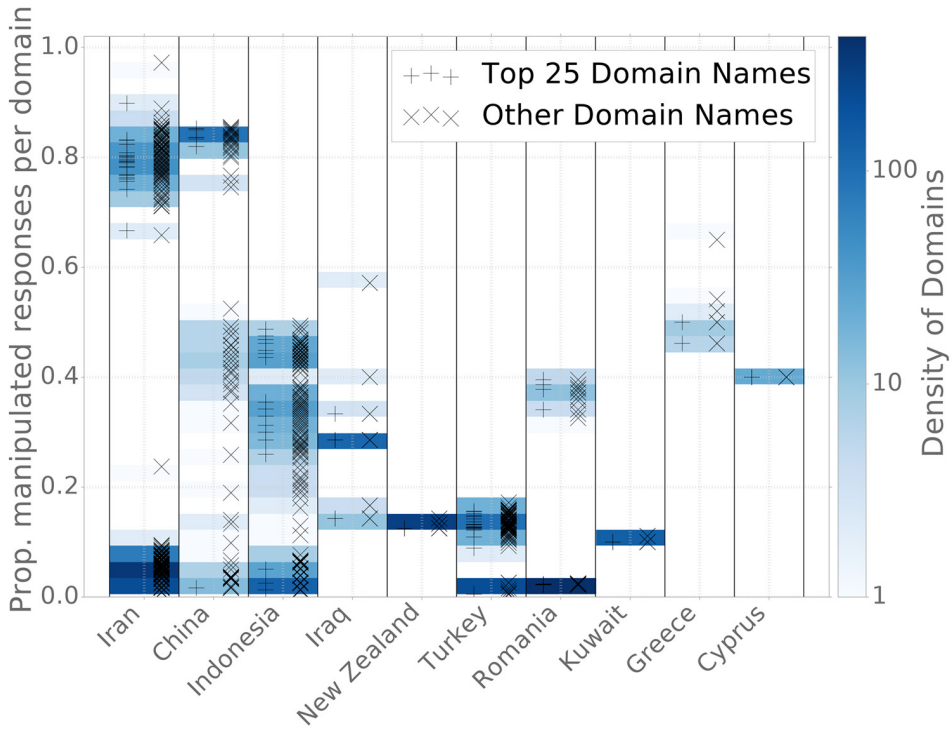


**Figure 3:** The fraction of resolvers within a country that manipulate each domain

### *Total Queries*

We issued 14.5 million DNS A record queries during a two-day period in January 2017. After removing problematic resolvers, domains, and failed queries, the data set had 13.6M DNS responses. Applying our *consistency* and *independent verifiability* metrics, we identified 42K manipulated responses (0.3% of all responses) for 1,408 domains, spanning 58 countries (and dependent territories).

## Manipulation by Country

Previous work has observed that some countries deploy nationwide DNS censorship technology; therefore, we expect to see groups of resolvers in the same countries, where each group of resolvers manipulates similar sets of domains. Table 1 shows the median percentage of manipulated responses per resolver, aggregated across resolvers in the top censored country.

### *Which Countries Experience the Most DNS Manipulation?*

Resolvers in Iran exhibited the highest degree of manipulation, with a median of 6.02% manipulated responses per Iranian resolver; China follows with a median value of 5.22%. The relative rankings of countries depend on the domains in our input data set.

For example, if sites censored in Iran and China are overrepresented in the CLBL, the overall rankings will skew towards those countries. Creating an unbiased globally representative set of test domains remains an open research problem.

### *Are There Outliers within Countries?*

Yes. Each country shown in Table 1 had at least one resolver that did not manipulate *any* domains. This effect likely results from IP address geolocation inaccuracies. For example, resolvers in Hong Kong (which are not subject to Chinese Internet censorship) were incorrectly labeled by MaxMind as Chinese. Additionally, for countries that rely on individual ISPs to implement government censorship, the actual manifestation of manipulation can vary across ISPs within the country. Localized manipulation by resolver operators in countries with few resolvers can also influence these results. Similarly, most countries had at least one resolver that showed DNS manipulation significantly greater than the median. This again points to localized manipulation, such as corporate networks deploying firewall products that block content unrelated to state-mandated censorship.

## Consistency within Countries

Figure 3 shows the DNS manipulation of each domain by the fraction of resolvers *within* a country, for the 10 countries with the greatest (normalized) level of manipulation. Each point represents a domain; the vertical axis represents the fraction of resolvers in that country that manipulated it. Shading shows the density of points for that part of the distribution. We only plot domains that experience blocking by at least one resolver within a given country.

| Country | Number Resolvers | Median Manipulation |
|---|---|---|
| Iran | 122 | 6.02% |
| China | 62 | 5.22% |
| Indonesia | 80 | 0.63% |
| Greece | 26 | 0.28% |
| Mongolia | 6 | 0.17% |
| Iraq | 7 | 0.09% |
| Bermuda | 2 | 0.04% |
| Kazakhstan | 14 | 0.04% |
| Belarus | 18 | 0.04% |

**Table 1:** Top countries by median percent of manipulated responses per resolver

Heterogeneity across a country suggests that different ISPs may implement filtering with different block lists; it might also indicate variability of blocking policies across geographic regions within a country.

### *Is There Heterogeneity Within Countries?*

Yes. For example, one group of domains was manipulated by about 80% of resolvers in Iran, and another group was manipulated by fewer than 10% of resolvers. Similarly, one set of domains in China experienced manipulation by approximately 80% of resolvers, and another set experienced manipulation only about half of the time.

### *Is There Non-Determinism in Censorship?*

Yes. Effects that appear as smearing across the vertical axis, such as for Iran and China, denote instances where individual domains were not manipulated by an identical set of resolvers but rather by an *almost* identical set. These phenomena can arise as the result of censorship systems using DNS *injection*, where a race condition exists between the injected and the original responses. It can also occur if systems under load fail open, or if the censor operates its manipulations in a probabilistic manner.

### *Is There Geolocation Inaccuracy?*

Yes. Upper bounds on the proportion of resolvers within a country performing manipulation suggest IP geolocation errors are common. For example, no domain in China experienced manipulation across more than approximately 85% of resolvers. This is also supported by the frequency of outliers within countries performing no manipulation, as discussed earlier.

| Rank | Domain Name | Category | Countries |
|---|---|---|---|
| 1 | www.*pokerstars.com | Gambling | 19 |
| 2 | betway.com | Gambling | 19 |
| 3 | pornhub.com | Pornography | 19 |
| 4 | youporn.com | Pornography | 19 |
| 5 | xvideos.com | Pornography | 19 |
| 6 | thepiratebay.org | P2P sharing | 18 |
| 7 | thepiratebay.se | P2P sharing | 18 |
| 8 | xhamster.com | Pornography | 18 |
| 9 | www.*partypoker.com | Gambling | 17 |
| 10 | beeg.com | Pornography | 17 |
| 80 | torproject.org | Anon. & cen. | 12 |
| 181 | twitter.com | Twitter | 9 |
| 250 | www.*youtube.com | Google | 8 |
| 495 | www.*citizenlab.org | Freedom expr. | 4 |
| 606 | www.google.com | Google | 3 |
| 1086 | google.com | Google | 1 |

**Table 2:** Domain names manipulated in the most countries, ordered by number of countries with manipulated responses

## Manipulation by Domain

Table 2 highlights which specific domains experienced manipulation in numerous countries, ranked by the number of countries.

### *Which Domains Are Most Frequently Manipulated?*

The two most manipulated domains were both gambling Web sites, each experiencing censorship across 19 different countries. DNS resolutions for pornographic Web sites were similarly manipulated, accounting for the next three most commonly affected domains. Peer-to-peer file sharing sites were also commonly targeted, particularly The Pirate Bay.

### *Are Commonly Measured Sites Manipulated by the Most Countries?*

No. Sites such as The Tor Project, Citizen Lab, Google, and Twitter are common censorship measurement targets. Yet our results show these sites experienced manipulation by significantly fewer countries than others (bottom half of Table 2). The Tor Project DNS domain, manipulated by 12 countries, was the most widely interfered with among anonymity and censorship tools; Citizen Lab experienced manipulation by four countries. Such disparity points to the need for a diverse domain data set.

| Rank | Domain Category | Countries |
|---|---|---|
| 1 | Alexa Top 10K | 36 |
| 2 | Freedom of expr. | 35 |
| 3 | P2P file sharing | 34 |
| 4 | Human rights | 31 |
| 5 | Gambling | 29 |
| 6 | Pornography | 29 |
| 7 | Alcohol and drugs | 28 |
| 8 | Anon. & censor. | 24 |
| 9 | Hate speech | 22 |
| 10 | Multimedia sharing | 21 |
| 20 | Google (All) | 16 |
| 34 | Facebook (All) | 10 |
| 38 | Twitter (All) | 9 |

**Table 3:** Top 10 domain categories, ordered by number of countries with manipulated answers

## Manipulation by Category

Table 3 shows the prevalence of manipulation by CLBL categories. We consider a category as manipulated within a country if any resolver within that country manipulated a domain of that category.

### *Which Types of Domains Are Most Frequently Manipulated?*

Domains in the Alexa Top 10K experienced the most manipulation; these domains did not appear in the CLBL, which highlights the importance of measuring both curated lists from domain experts as well as broad samples of popular Web sites. Although no single domain experienced manipulation in more than 19 countries, several categories experienced manipulation in more than 30 countries, indicating that while broad categories appear to be commonly targeted, the specific domains vary country to country.

### *Are Commonly Measured Sites in the Most Frequently Manipulated Categories?*

No. As was the case with domain ranking, common platforms such as Google, Facebook, and Twitter witnessed manipulation across significantly fewer countries than other categories.

### *Are the Top Manipulated Sites from the Top Manipulated Categories?*

No. While eight of the top 10 most frequently manipulated sites were in the Gambling and Pornography categories, those categories ranked 5th and 6th overall when aggregated. This difference highlights the need for measurement studies to consider multiple aggregation metrics when reporting ranked censorship results.

## Conclusion

Internet censorship is widespread, dynamic, and continually evolving; understanding the nature of censorship thus requires techniques to perform continuous, large-scale measurements.

Iris supports regular, continuous measurement of DNS manipulation, ultimately facilitating global tracking of DNS-based censorship as it evolves over time. Our first large-scale measurement study using Iris highlights the heterogeneity of DNS manipulation across countries, resolvers, and domains, and demonstrates the potential of operationalizing such measurements for longitudinal analysis.

### *Acknowledgments*

The authors are grateful for the assistance and support of Manos Antonakakis, Randy Bush, Jed Crandall, Zakir Durumeric, and David Fifield. This work was supported in part by National Science Foundation Awards CNS-1237265, CNS-1406041, CNS-1518878, CNS-1518918 CNS-1540066 and CNS-1602399.

***References***

[1] S. Burnett and N. Feamster, "Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM '15)*, pp. 653-667: http://bit.ly/2yw5OHZ.

[2] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman, "A Search Engine Backed by Internet-Wide Scanning," ACM Conference on Computer and Communications Security (CCS '15): https://censys.io/static/censys.pdf.

[3] Z. Durumeric, E. Wustrow, and J. A. Halderman, "ZMap: Fast Internet-Wide Scanning and Its Security Applications," in *Proceedings of the 22nd USENIX Security Symposium (Security '13)*: http://bit.ly/2wHvyE7.

[4] R. Ensafi, J. Knockel, G. Alexander, and J. R. Crandall, "Detecting Intentional Packet Drops on the Internet via TCP/IP Side Channels," in *Proceedings of the International Conference on Passive and Active Network Measurement (PAM '14)*, pp. 109-118.

[5] Freedom House, Freedom on the Net, "Silencing the Messenger: Communication Apps under Pressure," 2016: https://freedomhouse.org/report/freedom-net/freedom-net-2016.

[6] M. Kührer, T. Hupperich, J. Bushart, C. Rossow, and T. Holz, "Going Wild: Large-Scale Classification of Open DNS Resolvers," ACM Internet Measurement Conference (IMC '15): http://bit.ly/2xoVG3T.

[7] P. Pearce, R. Ensafi, F. Li, N. Feamster, and V. Paxson, "Augur: Internet-Wide Detection of Connectivity Disruptions," IEEE Symposium on Security and Privacy (S&P '17): http://bit.ly/2nlR1cY.

[8] P. Pearce, B. Jones, F. Li, R. Ensafi, N. Feamster, N. Weaver, and V. Paxson, "Global Measurement of DNS Manipulation," in *Proceedings of the 26th USENIX Security Symposium (Security '17)*: http://bit.ly/2xA6Qoy.

[9] A. Razaghpanah, A. Li, A. Filastò, R. Nithyanand, V. Ververis, W. Scott, and P. Gill, "Exploring the Design Space of Longitudinal Censorship Measurement Platforms," Technical Report 1606.01979, ArXiv CoRR, 2016: https://arxiv.org/pdf/1606.01979.pdf.

[10] The Tor Project, "OONI: Open Observatory of Network Interference": https://ooni.torproject.org/.