Understanding Routing-Induced Censorship Changes Globally

Abhishek Bhaskar
abhaskar@gatech.edu
Georgia Institute of Technology
Atlanta, GA, USA

Paul Pearce
pearce@gatech.edu
Georgia Institute of Technology
Atlanta, GA, USA

ABSTRACT

Internet censorship is pervasive, with significant effort dedicated to understanding what is censored, and where. Prior censorship measurements however have identified significant inconsistencies in their results; experiments show unexplained non-deterministic behaviors thought to be caused by censor load, end-host geographic diversity, or incomplete censorship—inconsistencies which impede reliable, repeatable and correct understanding of global censorship. In this work we investigate the extent to which Equal-cost Multi-path (ECMP) routing is the cause for these inconsistencies, developing methods to measure and compensate for them.

We find ECMP routing significantly changes observed censorship across protocols, censor mechanisms, and in 17 countries. We identify that previously observed non-determinism or regional variations are attributable to measurements between fixed end-hosts taking different routes based on Flow-ID; i.e., choice of intra-subnet source IP or ephemeral source port changes observed censorship. By developing new route-stable censorship measurement methods that allow consistent measurement of DNS, HTTP, and HTTPS censorship, we find ECMP routing yields censorship changes across 42% of IPs and 51% of ASes, but that impact is not uniform. We also develop an application-level traceroute tool to construct network paths using *specific* censored packets, thus identifying numerous causes of differences, ranging from likely failed infrastructure, to routes to the same end-host taking geographically diverse paths which experience differences in censorship *en-route*. Finally, we examine our results in the context of prior global measurement studies, exploring the applicability of our findings to prior observed variations, and then demonstrating how specific experiments from two studies could be impacted by, and specific results are explainable by, ECMP routing. Our work points to methods for improving future studies, reducing inconsistencies and increasing repeatability.

CCS CONCEPTS

• **Security and privacy** → **Network security**.

KEYWORDS

Internet Censorship; Censorship Measurement; Routing

ACM Reference Format:
Abhishek Bhaskar and Paul Pearce. 2024. Understanding Routing-Induced Censorship Changes Globally. In *Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (CCS '24), October 14–18, 2024, Salt Lake City, UT, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3658644.3670336

CCS '24, October 14–18, 2024, Salt Lake City, UT, USA

ACM ISBN 979-8-4007-0636-3/24/10
https://doi.org/10.1145/3658644.3670336

1 INTRODUCTION

Internet censorship impacts the lives of 72% of people [32], with governments and ISPs using sophisticated in-network capabilities to manipulate and disrupt DNS [53, 59, 61], HTTP [22, 52, 61], and HTTPS [42, 55, 61]. The growing prevalence of Internet censorship has given rise to significant measurement efforts focused on understanding the scale and scope of censorship globally [42, 52, 53, 61].

The challenge of obtaining globally distributed hosts has made *outside-in* measurement [7] an appealing alternative method of understanding censorship. Outside-in measurement leverages the symmetric nature of many countries' censorship infrastructure to send measurements *to* a vantage point in a censored area (instead of originating from) and then observing any actions taken against that flow. Outside-in measurements have demonstrated censorship globally, across DNS manipulation [53], packet drops [3, 22], RST injection [22, 52, 63], and block-page injection [55].

At the same time, prior censorship studies found inconsistency in DNS, HTTP, and HTTPS measurements, globally [2, 17, 23, 28, 45, 53, 55, 66, 68, 70–72]. These appear in the form of non-uniform censorship across a country/ISP [28, 37, 45, 53, 71–73], variations in results over time [55], and RST injection not observed for certain experiments [66]. Inconsistency makes understanding censorship and reproducing results challenging, impacting our ability to develop technical and policy interventions.

Integral to the notion of outside-in censorship measurement is the idea that measurements *to* a specific vantage point will traverse at least some of same set of the network infrastructure performing censorship as if *from* a vantage. Deeply embedded in that assumption is the concept of Equal-cost Multi-path (ECMP) routing. Routers on the Internet use various fields of a packet to construct a flow identifier (Flow-ID), and use that Flow-ID to assign the packet to a flow for load balancing [4]. Flow-ID is influenced by "ephemeral" fields such as source port, thus communication between a single source IP and destination IP/port may take numerous possible routes through the network [64]. While ECMP routing is well understood [4, 64], the extent to which ECMP routing influences censorship measurement globally across protocols is unknown.

Our work seeks to understand the extent of ECMP-routing's impact on outside-in censorship measurement across protocols, mechanisms, and countries, with an eye toward *why* changes in route impact observed censorship. Prior work exploring China's Great Firewall's (GFW) [7] DNS injection identified that some source parameters resulted in variations in injected DNS censorship. Their study is limited; they explore only China, and only DNS censorship; both of which are known to exhibit unique behaviors among the worlds' censors [31]. Given such unique characteristics, it is unclear if ECMP-induced censorship measurement variations are an artifact of the GFW, or if such phenomena generalize across countries and disparate censorship deployments and protocols. It also remains unclear *why* such variations exist, and their impact on measurement studies.

We ask: Does ECMP routing influence outside-in censorship measurement globally, is it a generalized phenomenon of censorship infrastructure, has it impacted prior studies, and why? This problem is challenging as prior tools do not allow control of the parameters that influence route, and prior traceroute tools either produce stable-routes of packets that are not of interest (e.g., ICMP ECHO), or produce unstable routes of application level protocols (e.g., HTTP packets). To these ends we develop *Monocle*, a new route-stable censorship measurement and traceroute platform able to understand ECMP-induced censorship changes not only across DNS, but also HTTP and HTTPS. *Monocle* expands prior DNS tooling [7] while also developing new methods to measure and traceroute RST packet injection, packet drops, and censor blockpages, across protocols.

We conduct a global study of 21 countries, 3 network protocols, and 4 censorship mechanisms, aimed at quantifying the effect of ECMP routing on both current remote censorship measurement as well as on prior studies. **We find that ECMP routing has significant impact on outside-in censorship measurement across countries, affecting 17 of 21 countries, as well as all types of protocols and mechanisms explored.** Our results illustrate a complex entanglement of end-to-end censor activity with low-level network behaviors that were previously thought unrelated and not considered. We find 42% of IPs and 51% of ASes show ECMP-induced changes in measured censorship, with that impact unevenly spread across 17 countries. We also find that some source IP and ephemeral source port combinations detect up to 2x more censorship than others, between the same end-points.

We also utilize *Monocle* to conduct censorship-and-path aware traceroutes of *specific* censored packets, enabling us to reconstruct network graphs, and explain *why* variations exist. We find a diverse set of explanations, ranging from routing within ASes sending some packets to potentially failed infrastructure, to routes to a single end-host spanning geographic regions with diverse censorship *on-path*. We also explore the different forms of observed variation in prior work, contextualizing when ECMP routing is potentially applicable. Finally, we compare our results to 2 prior outside-in studies, showing that previously observed non-determinism [53] is explainable by source-parameter selection, and a reproduction of the selection method of a prior censorship study [55] using our experiments shows as many as 35% end-points could experience ECMP-induced variations.

Contributions. Our contributions include:

- Designing and deploying *Monocle*, a platform able to quantify the effects of ECMP routing on censorship measurement across DNS, HTTP, and HTTPS protocols, and across DNS injection, RST injection, packet drops, and blockpage censorship methods.
- Finding that 17 of 21 countries explored with externally measurable censorship show censorship differences by varying intra-subnet source IP and/or source port, with the impact ranging from 100% to <1% of destinations showing differences. We also find all types of protocols and censor methods impacted.
- Finding that ECMP-induced differences are due to a diverse set of properties, ranging from route differences within ASes performing censorship likely having failed infrastructure on some paths, to routes to the same host having geographic diversity, and that geography demonstrating non-uniform censorship.
- Finding that previously observed non-determinism in a prior study [53] is explainable by ECMP routing, and replication of the selection method of a prior study [55] that experienced non-determinism shows up to 35% of its hosts could be impacted.

We thus argue ECMP routing must be taken into account when measuring in-network phenomena, and that measurement knowledgeable of these properties can remain a valuable research method.

2 RELATED WORK

Censorship measurement has evolved in the past decade to understand how censorship works, what is censored, and how censorship changes over time. In order to comprehensively measure censorship, studies built techniques to perform measurements on various protocols: DNS [2, 24, 39, 43, 53, 59], HTTP [16, 23, 41, 63, 67], HTTPS [8, 10, 12, 25, 55], HTTP/3 [21], and echo [55, 63].

Most state-sponsored censorship is deployed at the ISP or network backbone [38, 72, 75] in the form of network middle-boxes that intercept packets and perform actions based on them; such behavior affords measurement, whereby sensitive packets are sent across the middle-box, and behaviors are studied. Measurement typically takes two forms: 1) Outside-In (Remote) measurement, where packets are sent from *outside* a country being measured towards points *inside* the country [3, 10, 23, 24, 31, 43, 44, 52, 53, 55, 57, 61, 63, 68, 72], or 2) Inside-Out measurement, where measurements use observation points *inside* the country from volunteers, VPN servers, etc, to send sensitive censorship triggering packets [30, 42].

Both measurement types have trade-offs in deployability, ethical considerations, and scalability. Several global censorship measurements employ remote measurement techniques as it eliminates the need to have volunteers inside all the countries being measured [7, 44, 46, 53–55, 57, 58, 61, 63]. While these systems are scalable and reduce the need for volunteers, they in-turn can only measure censorship that is symmetric. Measurements are generally conducted at the DNS, HTTP or HTTPS layer with countries potentially performing censorship at any of these layers. At the HTTP layer, studies generally use the `Host:` header [22, 55, 63] to include a sensitive payload, and at the HTTPS layer a sensitive payload is encoded in the `SNI:` field [8, 12] of the HTTPS `client-hello`, that triggers various forms of censorship like RST-injection, blockpages, packet drops etc. At the DNS layer the sensitive payload (in the form of a domain) is issued as a DNS A? query, that elicits DNS injection that is then used to study censorship [24, 31, 43, 53, 68]. *Monocle* studies these protocols globally in the context of ECMP routing.

Equal-cost Multi-path (ECMP) Routing. Load Balancing is widespread on the Internet. Augustin et al. [4] explored multi-path routing in traceroute measurement, and subsequently [5] found that close to 72% of the (source, destination) pairs experienced some form of load balancing. Recently Vermeulen et al. [64] found 18% of ECMP routing path divergences span multiple ASes. Routers use different components of Internet Layer (IP) and Transport Layer (TCP and UDP) like the source and destination IPs, source and destination ports, etc, to perform these routing decisions. Routers are also known to use various bits of the source/destination IP and port to make decisions on which up-link to send packets [13, 65]. Thus any IP measurements are subject to load balancing.

Variance in Censorship Results. Prior censorship studies have noted country/ISP level inconsistency in DNS, HTTP, and HTTPS measurements [2, 17, 23, 28, 37, 53, 55, 66, 68, 70–73]. Pearce et al. [53] observed differences in DNS manipulation across resolvers *within* a country. Raman et al. [55] noticed sporadic blockpage injection within ISPs and organizations. Wang et al. [66] noticed that for a small percent of their experiments, they were successful in bypassing the GFW without any evasion strategy. Crandall et al. [17] in as early as 2007 observed that on 28.3% of destinations (in China) they observed no filtering, and even on the paths they *did*, filtering appeared to be volatile during high-load periods of the day. Rambert et al. [68] noticed differing levels of censorship depending on the source and destination of the probes (independent of geography). Anonymous [2] showed the presence of different injecting interfaces with changing source IPs. Wright et al. [71] and Xu et al. [72] both highlight geographic variation in censorship implementation across the country of China, with different provincial ISPs performing their own filtering (in addition to filtering at the border). Aryan et al. [3] speculate that individual ISPs can potentially implement their own blocking mechanisms in addition to centralized blocking (in Iran). Nisar et al. [45] showed differing blocking implementation by ISPs and differences even within ISPs in Pakistan. Both Yadav et al. [73] and Katira et al. [37] observed differences in censorship (in terms of domains censored) across different ISPs tested in India. Gill et al. [28] observed changes in measured censorship: across time, ISPs (within the same country), and even URLs within the same ISP in several countries. Winter et al. [70] found that certain Tor relays remained reachable from VPSes in China even after several days of the first Tor connection request from the VPS while a majority of them were blocked. Ensafi et al. [23] when studying China's GFW behavior with respect to Tor found that GFW's failure (to block Tor) were both persistent with routes *and* sporadic.

While country and ISP-level variation in censorship has been globally observed, there are numerous suspected caused of such variation. These causes include (but are not limited to): geographical differences in blocking [53, 71, 72], constantly changing blocking methods [55], differing ISP implementations [28, 37, 45, 73], network load (e.g., time-of-day) on censoring devices [17, 23], and selection of source/destinations [68]. We stress that while the goal of this work is to understand the influence of ECMP routing on censorship variation, we do not believe that all previously observed censorship variation is ECMP-induced.

Bhaskar et al. [7] first explored the role of routing on the Chinese DNS censorship system. Their study found that varying source IP and port has an impact on the path taken by packets and subsequently influenced the measurement of China's Great Firewall (GFW). What remains unclear from prior work is if the observed behaviors are an isolated artifact of the GFW and DNS injection, or a broader behavior across the Internet and censorship measurement at-scale. Our work seeks to expand and generalize this prior work by systematically exploring the wider effects of ECMP routing on censorship globally, across both countries and various protocols. We seek to quantify *why* this behavior exists pervasively across the Internet, occurring in disparate countries that lack coordination, consistent network topologies, or common technical measures.

3 METHOD

We seek to understand the impact of ephemeral parameters such as source IP (with a subnet) and source port—and hence Flow-ID—on censorship measurement globally [52, 53, 55, 61, 63]. *Exhaustively* identifying censorship is neither ethical nor our objective, but rather our goal is replicating established methods from prior studies and understanding whether outside-in censorship measurement methods are affected by ECMP routing.

We develop *Monocle* (Figure 1) to explore 3 commonly censored protocols [27]: DNS, HTTP, and HTTPS. Across these protocols *Monocle* looks at multiple known censorship techniques: DNS manipulation [53], RST packet injection [52], packet drops [3, 22], and blockpages [35, 55]. *Monocle* extends measurement methods from prior studies to create route-stable censorship measurements, enabling us to measure the impact of packet parameters on censorship.

3.1 Approach

Censorship measurement involves sending sensitive payloads between hosts that trigger a censor, and *inferring* censorship from responses (or lack thereof). *Monocle* implements outside-in censorship measurements by sending route-stable sensitive TCP and UDP packets, varying source IP across a single /24 network as well as source ports, and observing responses.

TCP-based censorship. TCP censorship works by identifying *connections* to censor, then leveraging termination or hijacking to disrupt communications. Censors use `Host:` header in HTTP or the SNI field (as part of the `client-hello`) in HTTPS connections to perform censorship, both of which are well studied [8, 10, 12, 16, 23, 25, 41, 55, 63, 67]. To disrupt communications, TCP censorship techniques include RST injection, packet drops, and blockpages [36].

Monocle establishes TCP connections with destination IPs inside a censored country from a vantage point outside the country and sends a packet with keywords—either embedded in the `Host:` header of a HTTP GET request or SNI field of a HTTPS `client-hello` request—and recording the responses to measure censorship. **A key aspect of *Monocle* not seen in prior work is controlling all aspects of generated packets that are known to change Flow-ID, thus allowing us to perform route-stable censorship measurements.** We perform trials with control and sensitive domains to establish measurement reliability. We test destinations that have ports 80 for HTTP and 443 for HTTPS, open. We identify candidate IPs via Censys [20] and then confirm their status.

To identify RST based censorship we look for the presence of RSTs in our response packets in addition to *no* RSTs from control measurements. To identify packet drop based censorship we look for the absence of a payload response in addition to the presence of a payload response for the controls. Differentiating true censorship from network phenomena for both these forms of censorship can be challenging as: 1) RST packets can occur for reasons unrelated to censorship, 2) lack of RST packets or payload response can occur due to packet loss, and 3) transient censorship failures can cause RST packets to be missing or payload responses to appear.

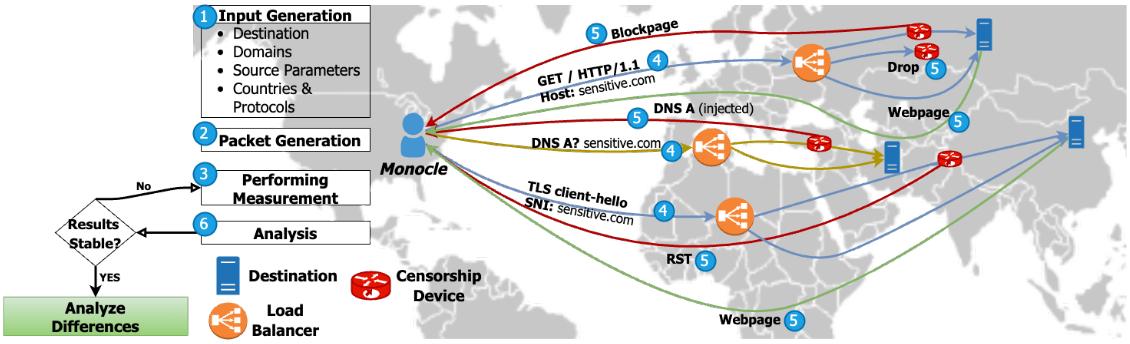


Figure 1: *Monocle*, our system to understand impact of routing on censorship measurement across countries and protocols.

To disambiguate censorship from network behavior we perform repeated measurements separated by a fixed time interval. We note a significant time interval needed to ensure the tuple can be reused by the remote operating system. We use 30 minutes, which is 15x longer than the RFC recommendation [62] and significantly longer than needed based on our experiments. We note long-lived residual censorship [9] beyond 30-minutes is not a concern, since our goal is to study routes, not censor activity. We only consider a result censorship for a particular (destination, source) parameter if: 1) we obtained a RST for all the repetitions and *no* RST for any of the control measurements, or 2) we obtained *no* payload responses for any of the repetition and responses for all the control measurements. We only consider a result *not* censorship if across all repetitions: 1) no RST packet was observed for RST based censorship, or 2) payload responses were obtained for packet drop based censorship. We exclude all other results. This approach is *conservative*, as it gives us a lower bound of possible true-positive results, discarding potential differences due to network changes, packet loss, etc. i.e., this approach gives us a set of differences we have high confidence are routing-induced, rather than transient effects. Other phenomena and instances of routing-induced censorship differences may very well exist and be excluded by this method due to transient network effects, but our results will still provide a lower-bound.

Blockpages are detected by conducting a manual study of censorship in each country to extract blockpage templates that are then matched against responses. This manual operation ensures matches with no false positives but possible false negatives. This is acceptable, as it provides a lower bound on our findings on routing-induced censorship changes, rather than exhaustive censorship measurement. The lack of a blockpage could also be attributed to similar causes as the lack of a RST packet. To account for this we perform similar repetitions and consider a result valid when there is consistency across all iterations.

UDP-based censorship. For UDP censorship, *Monocle* extends prior work [7]. We send route-stable sensitive and control DNS A? queries to destinations globally that *do not* operate DNS resolvers and record all responses. Censorship occurs when we obtain a response to a sensitive query, but not the control.

Improving Result Reliability. Across all experiments we use a combination of control measurements and repetition separated by fixed time intervals to obtain reliability. We only consider censorship present or absent for a particular experiment if the relevant behavior is present **across all** repetitions in addition to consideration of control experiments. This stabilizes our results by eliminating factors like packet loss, residual censorship, and transient behaviors that may appear as censorship changes. It is possible that within the time-frame of our repetitions network topologies may change, resulting in a missed non-censorship result. This is acceptable as our goal is to have a lower bound on censorship changes.

Route-Stable, Per-Protocol Traceroute. Beyond simple observed censorship differences, we also seek to understand *why* these differences occur, which requires constructing the routes taken by *specific* packets with *specific* parameters. Common prior traceroute techniques are built on UDP, TCP-based methods use SYN packets, and purpose built censorship traceroute techniques [56] do not control for Flow-ID. For our usage we need to measure the path of *specific* sensitive packets with full control over all aspects of the packets; no tools allow reconstructing routes of a specific user-controlled packet. Further, for TCP, the sensitive packet is sent after establishing a TCP connection, thus we must iterate the TTL of the payload packet of an active connection. This is challenging as we cannot control retransmission or TCP state from user space. Thus we develop a route-stable traceroute tool to measure the path taken by packets for all protocols we explore. We control all fields in the packet that influence routing, varying the TTL of a given sensitive packet, and build the path using ICMP responses.

In lieu of implementing a TCP stack in user-space to vary TTLs of mid-flow sensitive packets, *Monocle* leverages a combination of Netfilter nfqueues [69] and firewall rules. Our firewall redirects outbound packets we seek to traceroute to an nfqueue, and our nfqueue hook sends copies of the packet at incrementing TTLs. Kernel retransmission attempts are dropped until the traceroute is complete or the connection is terminated by the endpoint or censor. This technique ensures: 1) we can capture the actual path taken by the packets, and 2) it can be applied to any censorship measurement technique. With this technique the connection does not need to be torn down after each TTL attempt, as the packet does not reach the end-host until the end of the trace. We embed the TTL in the IP ID field of the packet to disambiguate responses. We note the IP ID field is not known to be used for Flow-ID [4], and does not influence route. We extract the source parameters and TTL from the packet embedded in the ICMP response to build the network path.

3.2 Ethical Considerations

Measuring Internet censorship requires careful consideration of the ethical implications of all experiments, and weighing those implications against the potential benefits of the understanding gained from the experiment. We build our ethical framework around the models from prior censorship measurement works [7, 52, 53], who in turn modeled their work after the Belmont [6] and Menlo [18] reports. Namely, we consider the concepts of *justice*, *respect for persons*, *beneficence*, and *respect for law and public interest.* Broadly speaking, these principles dictate that censorship measurement researchers should strive to: 1) Ensure those who bear the risk of the work are also the work's beneficiaries, 2) Given the impossibility of obtaining informed consent, seek to *minimize* risk, 3) Ensure that no experiments stress the infrastructure or users' machines.

We call attention to *beneficence* which deals with experimentation that has inherent risks and speaks to the need to reduce risk to the extent such that the benefits of conducting the measurements outweigh the risk. Prior censorship measurement [7, 52, 53] discussed this concept. e.g., "In lieu of attempting to obtain informed consent, we turn to the principle of beneficence, which weighs the benefits of conducting an experiment against the risks associated with the experiment. Note that the goal of beneficence is not to eliminate risk, but merely to reduce it to the extent possible." [53].

Guided by these principles, we reduce risk by: 1) significantly down sampling measurement vantage points to a minimum set per autonomous system necessary to show the effect of ECMP on censorship measurement, 2) measuring only a single censored domain per host (the same domain across an entire country) which we manually selected to minimize potential harm (e.g., by excluding terrorism or similar categories), 3) only conducting remote measurements that do not result in follow-on host-initiated communication with censored domains or IPs, and 4) we rate-limit and randomize experiments to minimize load on any machine.

We note that the benefits of this work include providing the community with the knowledge of how to conduct sound censorship measurement, which will enable us to develop tools that better understand censorship globally, with fewer measurements. Such understanding in-turn enables the development of better circumvention technologies, and aids policy makers and activists; all benefits which impact the population that bears the risk of the experiments.

4 RESEARCH QUESTIONS AND EXPERIMENTS

We begin by enumerating our research questions, and then defining the experiments we designed and datasets we collected to answer them. We seek to answer these questions across protocols, censorship techniques, and countries:

- **RQ1:** What is the path diversity of censorship measurements?
- **RQ2:** What is the impact of ECMP routing on censorship measurement results?
- **RQ3:** Why and how much do different packet parameters influence censorship measurement?
- **RQ4:** What are the underlying network structures that cause changes in censorship?
- **RQ5:** How do these results contextualize with specific prior works?

These research questions build on one another, beginning with an exploration of the impact of varying source parameters (and thus Flow-ID) on the path measurement packets take to vantages globally, regardless of censorship, and then observing this impact on censorship results. We then perform a deeper analysis to understand the different reasons that cause such variation. We end with trying to understand the applicability of such variation on prior censorship measurement studies, and its potential impact on two specific studies. All experiments were conducted from a purpose built /24 scanning network at an academic institution in North America, using the methods from Section 3.

4.1 RQ1: What is the path diversity of censorship measurements?

In **RQ1** we explore how varying censorship measurement packet source IP and source port changes paths across countries and protocols, absent censorship. The goal in exploring this question is quantifying the extent to which different parameters result in different routes, which is a prerequisite for ECMP censorship measurement differences. To answer this, we conduct route-stable HTTP, HTTPS, and DNS traceroute measurements (Section 3.1) with a control domain to a geographically diverse set of vantage points in various countries. Across measurements we either fix source parameters, vary an individual parameter, or vary both parameters [7].

Destination IPs. To understand the impact of packet parameters on route, we select a diverse set of measurement IPs within each country. Starting with a list of all possible destinations for each protocol, we use Censys [20] data to extract destinations with port 80 open for HTTP, port 443 open for HTTPS and port 53 *not* open for DNS). We then pick one destination per AS for each country tested. The number of destinations selected will vary for each country.

Domain. As our goal in this scenario is to measure the variation in packet path in different censorship measurement scenarios and not to perform censorship measurement, we perform measurements with a known benign domain (example.com). The advantage of this is twofold: (1) keeping with our ethical considerations, we reduce risk to individuals by using benign domains when possible, and (2) we avoid any side-effects from censorship activity that could possibly alter the state of open connections used for traceroute (e.g., RST packets that tear down active connections).

Parameter Variation. When selecting source IPs and ports, we conduct four variations of the following experiments:

- **Everything Constant.** We repeat the experiment 144 times, replicating experiments from censorship measurement tools across DNS, HTTP, and HTTPS [10, 53, 63].
- **Varying Source Port.** Varying the source port of the measurement packet with 144 randomly selected ports from the ephemeral range and fixing the source IP.
- **Varying Source IP.** Varying the source IP of the measurement packet with 144 randomly selected IPs from a */24* subnet and using a fixed source port.

- **Varying Source IP and Source Port.** Varying both the source IP and source port of the packet simultaneously. We pick 12 randomly selected source IPs and source ports, ensuring that we obtain 144 measurements in total.

This four-fold approach gives us the ability to understand the variation of the path taken by the packet across different individual dimensions and with them combined. Of note is the need to ensure a consistent total number of experiments (e.g., 144 experiments vs 12 experiments 12 times), to ensure comparisons *between* experiments are apt. We pick 12 based on prior work [7] which found that results stabilized by iteration 12 (Section 4.5 [7]).

4.2 RQ2: What is the impact of ECMP routing on censorship measurement results?

Next **RQ2** seeks to explore the impact of ECMP routing on the *results* of outside-in censorship measurement across protocols, censorship mechanisms, and countries. i.e., do different routes lead to different censorship results? We thus conduct an experiment utilizing the methods described in Section 3, varying different source parameters, across numerous censorship methods and countries. We do not perform traceroutes for this experiment, reducing packets and risk.

Country and Domain Selection. For this experiment our goal is to find all potential countries with state-sponsored and ISP censorship mechanisms amenable to outside-in measurement, and an associated sensitive domain that can be used to measure the same. We do this by performing a preliminary manual exploration on *all* countries and dependent territories we could identify as having IP addresses geolocated too. This totaled 249 countries and dependent territories. For each, we: 1) pick potential sensitive domains using a combination of OONI [47] and CLBL [15], and sample destination IPs (open on 80 and 443, and closed on 53, similar to RQ1) from Censys [20], and 2) perform a handful of external censorship measurements to these destination IPs on the different potential sensitive domains. As a result of this exploration, we find 21 countries from Europe, Asia, Africa, and Middle East that were able to be remotely measured. We explore this RQ with the associated domains that are confirmed to be censored in the target country. We note that such an approach may not exhaustively find *all* countries and protocols that are amenable to outside-in measurement, but this is acceptable, as our goal is to understand and provide a lower bound on the phenomenon's generalizability globally, not exhaustively enumerate all possible scenarios.

Destination IPs. For this experiment, for each country, we select a set of geographically diverse destination IPs to get an accurate picture of the change in censorship due to ECMP routing. The criteria we look for in a destination IP are unchanged from RQ1: open on 80 and 443, and closed on 53. These criteria allow us to perform measurements across protocols `HTTP`, `HTTPS`, and `DNS` on the same IP. From this usable set of IPs (identified from Censys [20]), we sample up to 60 destinations per AS, depending on the volume of ASes and available IPs for each country. AS mapping is performed with the RouteViews prefix-to-as dataset [11].

Source Parameters. We use 208 source IPs selected from a single `/24` subnet (excluding `.0` and `.255`), and 8 source ports randomly selected from the ephemeral port range. We ensure that the source IPs are normally distributed with respect to their lowest 3 bits, which is relevant for quantifying some load balancing [14].

4.3 RQ3: Why and how much do different packet parameters influence censorship measurement?

RQ3 seeks to understand whether particular source parameters (IP or port) have a noticeable impact on variation in censorship results caused by ECMP routing, i.e., do some source IPs/ports have a greater impact than others? Exploring this begins to sheds light *why*, from a packet perspective, differences occur, and potentially identifies different ECMP algorithms that use different parts of the packet to perform routing. We use measurements from RQ2's experiments to answer this question and explore impact across source IPs, source ports and their combinations.

4.4 RQ4: What are the underlying network structures that cause changes in censorship?

Next **RQ4** explores different underlying causes for ECMP routing induced censorship differences. i.e., is ECMP routing causing sensitive packets to route around censorship, traverse through failed censoring nodes, or pass through completely different geographical locations with different censorship behavior? To answer this question we conduct route-stable traceroute measurements (Section 3.1) to produce network graphs which we use to answer this question. We describe the dataset used subsequently:

Destinations and Sources. We use the measurements of **RQ2** to sample destinations for all countries/protocols that exhibit variation and produce network graphs for those source parameters (Section 3.1). We select destination and source parameter combinations that exhibited consistent, repeatable censorship change.

Our goal is to: 1) qualitatively demonstrate that the differing observed censorship results are a direct consequence of varying network paths, and 2) understand and quantify the different underlying effects that contribute to such variation. We perform this exploration across all censorship methods and countries from RQ2.

4.5 RQ5: How do these results contextualize with specific prior works?

The goal of **RQ5** is twofold: 1) to understand the applicability of ECMP routing on the variation observed in prior studies, and 2) to understand the potential impact ECMP routing had on 2 specific prior global outside-in censorship measurement studies. For the later, we focus on censorship differences unattributed in prior work that could be ECMP routing induced. We stress that these 2 prior works are not *incorrect*, but rather effects observed in such work may be attributable to ECMP routing, rather than other effects. We seek to understand the impact of ECMP routing induced censorship differences in two dimensions: 1) can unattributed ambiguity in prior results be potentially attributed to ECMP routing, and 2) are end-to-end results from prior work impacted by ECMP routing? We use the measurements of **RQ2** to answer this questions.

When comparing to the 2 prior studies, we make a best-effort to use their published measurement methods to sample the same amount of destinations from the same locations, and compare those

	DNS		HTTP		HTTPS	
Country	Methods &	Affected?	Methods &	Affected?	Methods &	Affected?
Algeria (DZ)	-	-	Drop+RST	✓	Drop	✓
Bangla. (BD)	-	-	BPage+Drop	✓	Drop+RST	✓
Belarus (BY)	Inject.	✗	BPage	✓	RST	✓
China (CN)	Inject.	✓	RST	✓	RST	✓
India (IN)	Inject.	✗	BPage+RST	✓	RST	✓
Indonesia (ID)	-	-	BPage	✓	RST	✓
Iran (IR)	Inject.	✓	BPage+RST	✓	Drop+RST	✓
Jordan (JO)	-	-	RST	✗	RST	✓
Kuwait (KW)	-	-	Drop+RST	✓	Drop+RST	✓
Oman (OM)	Inject.	✗	BPage+Drop	✗	RST	✓
Pakistan (PK)	Inject.	✗	Drop+RST	✓	Drop+RST	✓
Qatar (QA)	-	-	BPage	✗	RST	✗
Russia (RU)	Inject.	✓	BPage+Drop	✓	Drop+RST	✓
Rwanda (RW)	-	-	RST	✓	RST	✓
S. Korea (KR)	-	-	BPage	✓	RST	✓
Syria (SY)	-	-	RST	✓	RST	✓
Turkey (TR)	Inject.	✗	RST	✓	RST	✓
Turkmen. (TM)	Inject.	✓	RST	✗	-	-
UAE (AE)	-	-	BPage	✗	-	-
Uzbek. (UZ)	-	-	BPage	✗	Drop+RST	✗
Yemen (YE)	-	-	RST	✗	RST	✗

Table 1: Impact of ECMP routing on censorship measurement. 17 out of the 21 countries show changes in due to route. ✓ denotes changes, ✗ denotes no changes, and - denotes no such externally measurable censorship.

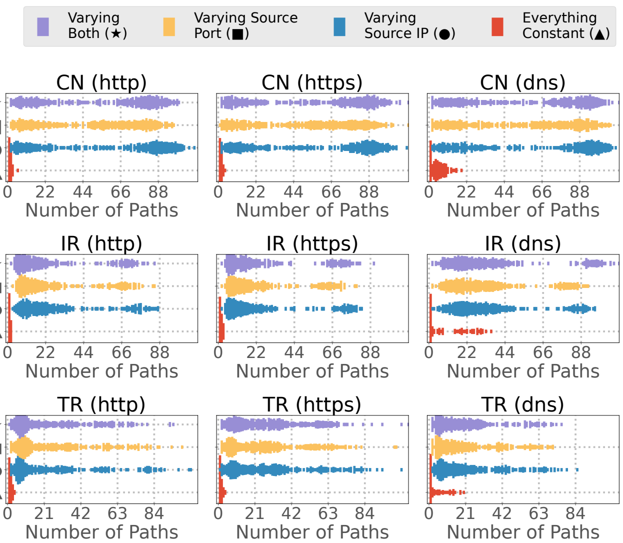


Figure 2: *Normalized Distribution of Number Of Paths for all destinations.* Marker size represents number of experiments across all destinations that had the particular *Number of Paths*. We observe that the variation in path has two observable modes, constant parameters vs varied parameters, and that source IP has a greater impact on network path variations than varying source port. We also observe different modes within a particular country (e.g., Iran) potentially indicating destination port-specific based load balancing.

results with our own. We do not replicate their *specific* destinations, due to both the significant time since their studies were collected during which routes likely changed, and dataset of specific IPs not being publicly available. Our formulation is meant to understand potential impact and scope, rather than identify precise results.

5 RESULTS

We now explore results from the experiments outlined in Section 4, focused on answering the five motivating research questions. Table 1 provides an overview of results from our study detailing the censorship explored for each country. We find 17 out of 21 countries explored that perform externally measurable state-level censorship are subject to variation in observed censorship results due to ECMP.

5.1 RQ1: What is the path diversity of censorship measurements?

We aim to understand if varying networking paths leads to changes in censorship measurement globally, across protocols and ports. We employ the metric of *Number of Nodes* and *Number of Paths* to capture variation in network path. *Number of Nodes*, for a destination is a *set* of all unique router hops in the network path across all experiments. A *path* is defined to be a set of all nodes for a particular combination of destination and source parameters. *Number of Paths* is thus a count of the set of all paths for a destination.

This form of ECMP routing within a country potentially depends on the network infrastructure of the country and its ISPs. Additionally, as packets pass through several different infrastructures to reach a destination, they pass through several different ASes [64] with potentially different network infrastructures that can influence their path. Similarly routing may differ based on destination port [4], causing different protocols to result in different routing.

Figure 2 shows results from these experiments (Section 4.1) and the variation in network path as a function of varying different source parameters part of the Flow-ID. Results are represented as the (normalized) distribution of the *Number of Paths* for the experiments conducted for each country and protocol. For space we abbreviate the number of countries shown to a characteristic subset. Appendix A in the extended version [76] shows similar results for *Number of Nodes*.

For the experiment where we fix source parameters and repeat the measurement (144 times), we see that there is very little to no variation in the network path. The mean *Number of Paths* remained ≤ 3 across all the different protocols tested in different countries. This observation aligns with expectations and prior work [7] and holds for a larger set of protocols and countries. Although the path remains constant across repetitions most of the time, there can be subtle variation over time in some cases (e.g., TR and IR with DNS). We find that repeating the experiments with the same source parameters does not frequently exercise different network paths.

Figure 2 shows that varying the different source parameters (independently and together) yields notably different behavior in the network path variation among different countries. For a (country, protocol), we observe two general modes in the observed variation: 1) varying source IP has a greater effect on the network path than varying source port (e.g., CN & IR), and 2) varying source IP/port seems to have a similar effect on resulting path density (e.g., TR). In addition, we see modes in the observed variation pointing to different infrastructure (e.g., IR, TR, CN), highlighting the need for both source IP *and* port variation.

Apart from variation between countries, we also observe variation in path diversity across *different protocols* within the same country. In the case of Iran, we see that the distribution of *Number of Nodes* varies between HTTP, HTTPS & DNS. We hypothesize this steams from destination port also being used for ECMP routing, opening the door to differences in censorship measurement based solely on port. All told, we note the significance in path diversity in censored countries due to source parameters indicates censorship variations are possible, explored next.

5.2 RQ2: What is the impact of ECMP routing on censorship measurement results?

RQ1 established that varying fields contributing to Flow-ID impacts the network path taken by a censorship measurement packet for numerous protocols across various countries. We now look at ECMP routing's impact on observed censorship results globally. For this experiment we selected a small number of destinations per AS per country (Section 4.2). Per our ethical framework (Section 3.2), our goal is to limit our measurements while still confirming the existence of the phenomenon. We then conduct measurements across different protocols with the same set of destinations and source parameters, for one control domain and one sensitive domain specifically picked to exercise censorship in the country. We use the methods described in Section 3 to identify the presence or absence of RSTs, intentional packet drops, blockpage, or DNS based censorship. We then identify destinations that produce different observed censorship for different sets of source parameters.

We find that censorship measurement from 17 out of 21 countries are affected by ECMP routing, with 34 out of 49 contexts impacted. Table 1 provides results across all countries and protocols. We find that the extent of impact varies significantly with country *and* protocol, ranging from 99% to 1% of destinations (depending on country) affected for HTTP, from 100% to <1% of destinations affected for HTTPS, and from 65% to 5% of destinations affected for DNS. Table 2 provides a breakdown of these results. Routing impact not only varies for different countries, but also for different protocols within a country. e.g., Iran shows differences across 2%, 9%, and 8% of destinations for HTTP, HTTPS, and DNS respectively. Such comparison among protocols supports the hypothesis of different underlying censorship infrastructures possibly deployed at different locations in addition to ECMP routing affecting protocols differently. We speculate our observed lower Chinese DNS censorship impact compared to prior work [7] is due to our use of fewer parameters. Appendix B in the extended version [76] shows overlap across protocols.

Country	Protocol	Destinations (ASes)	Destinations (ASes) Affected
Algeria	HTTP	311 (4)	2% (25%)
	HTTPS	288 (5)	3% (40%)
Bangladesh	HTTP	743 (119)	41% (58%)
	HTTPS	750 (122)	37% (54%)
Belarus	HTTP	1094 (13)	12% (31%)
	HTTPS	1049 (10)	11% (20%)
China	DNS	2501 (289)	25% (46%)
	HTTP [1]	210 (64)	23% (21%)
	HTTPS [1]	948 (217)	69% (71%)
India	HTTP	3289 (87)	99% (99%)
	HTTPS	3316 (90)	100% (100%)
Indonesia	HTTP	974 (78)	84% (71%)
	HTTPS	979 (80)	84% (68%)
Iran	DNS	1774 (127)	9% (26%)
	HTTP	1351 (142)	2% (2%)
	HTTPS	1347 (139)	9% (20%)
Jordan	HTTPS	231 (4)	< 1% (25%)
Kuwait	HTTP	1188 (15)	91% (80%)
	HTTPS	70 (11)	90% (82%)
Oman	HTTPS	842 (15)	2% (7%)
Pakistan	HTTP	373 (52)	15% (18%)
	HTTPS	704 (93)	44% (62%)
Russia	DNS	105 (33)	65% (46%)
	HTTP	1808 (327)	66% (47%)
	HTTPS	2526 (362)	15% (20%)
Rwanda	HTTP	23 (1)	92% (100%)
	HTTPS	54 (1)	75% (100%)
South Korea	HTTP	951 (44)	17% (44%)
	HTTPS	391 (29)	2% (4%)
Syria	HTTP	154 (1)	14% (100%)
	HTTPS	152 (1)	15% (100%)
Turkey	HTTP	526 (36)	21% (34%)
	HTTPS	410 (33)	18% (31%)
Turkmenistan	DNS	643 (1)	5% (100%)

Table 2: RQ2 Summary. Given are number of censored destinations studied (number of ASes in parentheses) and percent impacted by ECMP; impact varies significantly by context.

5.2.1 *Prevalence of Source Parameters Yielding Changes.*

We now explore measurements for each of the censorship protocols across the different countries, by number of destinations and source parameters. We seek to understand the prevalence of parameters leading to changes in censorship responses in order to gauge the expected impact of ECMP routing on censorship measurement. e.g., do many, or only a few parameters, impact results?

Figure 3 shows a CDF of (Source IP, Source Port) combinations that demonstrate no censorship, across varying destinations, broken out by country and protocol. We limit this measurement to destinations that experience variation in Table 2. We observe that the per destination impact of ECMP routing varied broadly by country. With the percent of (source IP, source Port) that observed no censorship for the median *affected destination* ranging from 93% in the case of KW (HTTPS) to as low as 1% in the case of CN (HTTP). This calls for careful selection of measurement parameters.

[1] During our study a significant change to China's GFW was deployed, resulting in a new form of reactive blocking [26], impacting our ability to measure ECMP routing in China. Thus we present partial results for HTTP and HTTPS predating the change.

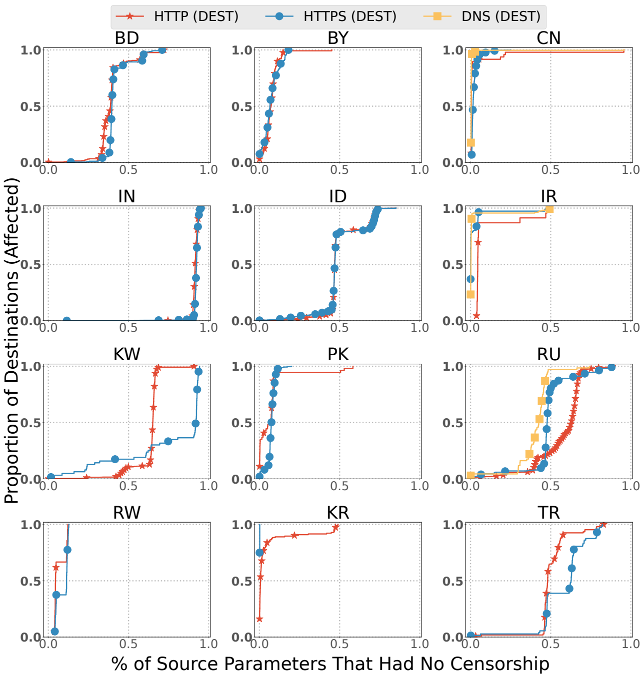


Figure 3: CDF of percent of (source IP, source Port) combinations per destination for which we observed no censorship. Measurements are limited to the subset of destinations where variation is observed in Table 2. We observe that 1) the percent of source parameters that produce no censorship for the median affected destination varies significant by country and protocol. 2) `HTTP` and `HTTPS` follow very similar trends in how source parameters affect their results in some countries (BD, BY, IN, & ID) while in others, they vary significantly (RU, KW) and 3) `DNS` is affected least by source parameters.

We also observe modal phenomena. For example, in ID (HTTP and HTTPS) for nearly ~70% of the destinations that are affected, ~48% of the (source IP, port) combinations produced no censorship. We see similar patterns in BD with the split being at ~40% of the source parameters, in BY with a split at around 10%, in RU (HTTP) with a split at 50%, and also patterns in IN and RW. We speculate this could be routing algorithms that divide traffic based on some fixed function of the packet, with some paths lacking censorship.

We also observe that while for some countries the pattern of per-destination impact follow similar trends for protocols within the country (e.g., HTTP & HTTPS with BD, BY, IN, & ID), in others, the pattern *varies even among protocols* (e.g., HTTP vs HTTPS RU). It is interesting to note that while HTTP & HTTPS follow similar trends for most countries, DNS varies the most from the other two protocols. We speculate that this behavior could be caused by either: 1) different censorship infrastructure (possibly at different locations) for each of the protocols, or 2) the destination ports of these protocols themselves being used in load balancing [4], yielding different routes and changing censorship infrastructure.

5.3 RQ3: Why and how much do different packet parameters influence censorship measurement?

We now seek to understand how particular source IPs, source ports or the combination thereof impact censorship variation. To achieve this we perform analysis on the low-order bit-patterns (known to be used for routing [14]) of both source and destination IPs compared to observed censorship variation. Figure 4 shows a breakdown of five modal behaviors, discussed subsequently. Results are colored by the lowest 3 bits of either source or destination IP.

1. Particular source IPs see more censorship than others, associated with specific bit patterns in the source IP. Figure 4a shows that based on the source IP selected, the number of destinations that are impacted by ECMP routing more than double. It was interesting to note that: 1) we only observed this particular effect with China and the Great Firewall's censorship, and 2) the magnitude of lower order bits' impact flipped in China (HTTPS), compared to China (DNS) from prior work [7]. i.e., 0b000 had the least impact in DNS previously versus most impact in HTTPS now. This finding is consistent with and further supports the variation observed by Anonymous [2] where they found different injecting interfaces simply by changing the source and destination of the probes (with the GFW in China).

2. Particular source IPs have higher impact than others with no known associations with source IP bit patterns. In this case we still see significant changes in the impact on particular source IPs but cannot directly attribute them to any source IP bit patterns. We see this particular effect in Figures 4b and 4c where the impact varies significantly by source IP, but each country denotes a distinct pattern with respect to destination IP bits. In addition we observe modal patterns in Figure 4b which indicate some source parameters are used differently for ECMP routing, changing censorship.

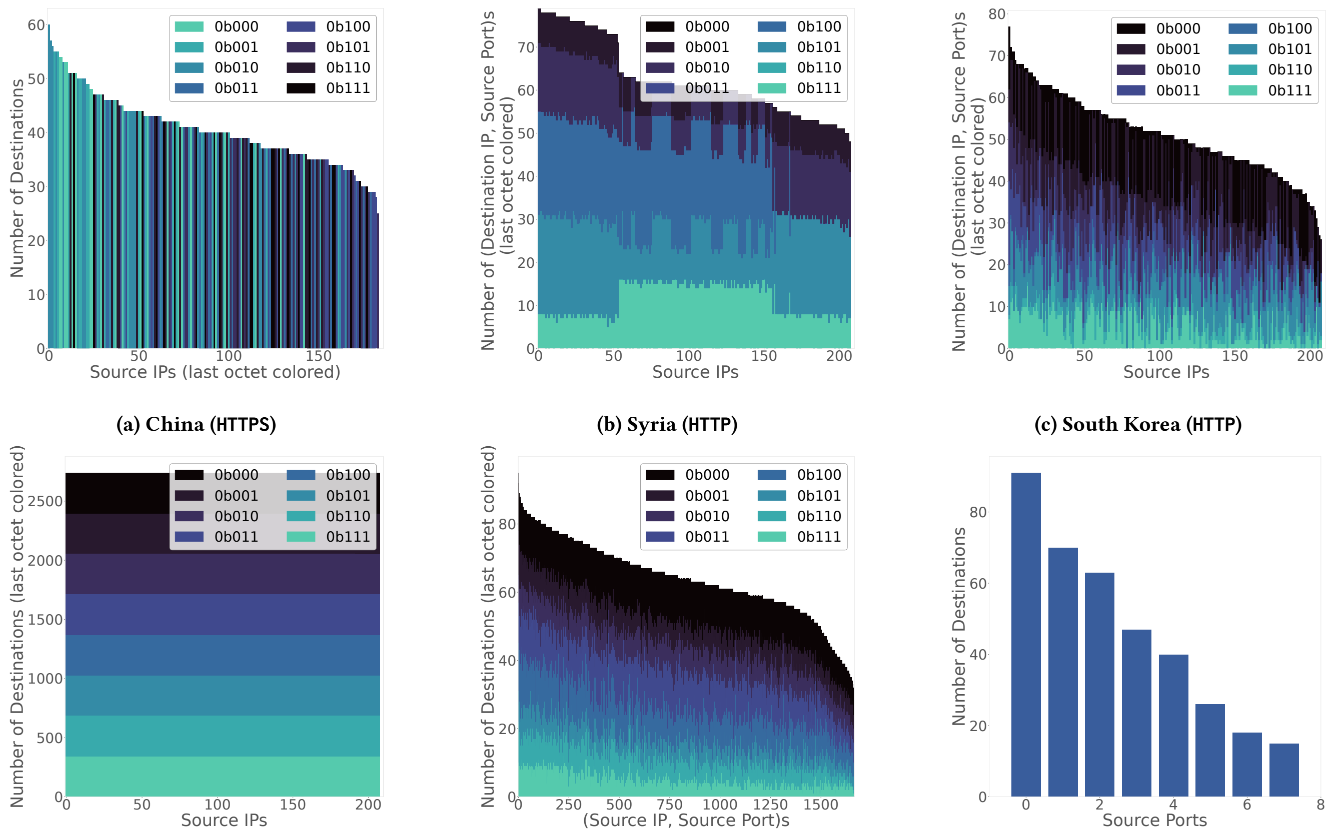
3. Uniform distribution of observed changes across the source IPs. In this case the destination IPs yield uniformly distributed changes across all source IPs tested, as seen in Figure 4d. This likely indicates hashing to perform ECMP routing, vs specific bits.

4. Particular Source IP + Source Port combinations have higher impact than others. In this case a particular (source IP, source port) combination produces significantly more variation than others. Figure 4e shows this, where the number of destinations impacted more than doubles between two (source IP, port) combinations. This further points to the need to control both parameters.

5. Particular Source Ports have greater impact than others. In this particular case, simply choosing different source ports has significantly different impacts on the variation in censorship measurement. We can observe this in Figure 4f where the number of destinations impacted by a particular source port can more than double based on port selection. We note that: 1) we only observed this behavior in the context Iran, and 2) this directly contrasts prior work [7] which did not observe any variation due to source port.

5.4 RQ4: What are the underlying network structures that cause changes in censorship?

We have established that routing has a direct impact on the network paths of packets and subsequently the results of common



(a) China (HTTPS)

(b) Syria (HTTP)

(c) South Korea (HTTP)

(d) India (HTTP)

(e) Bangladesh (HTTP)

(f) Iran (DNS)

Figure 4: *Influence of Source IPs and/or Ports on Changes in Censorship*. Sources and destinations are sampled uniformly across the lowest 3 bits, and colored based on those bits for either source or destination IPs. X-Axis sorted.

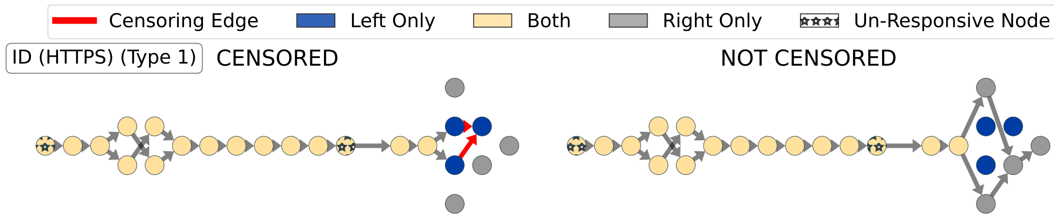
censorship measurement techniques. In this section we seek to make qualitative associations between differing censorship results and potentially deviating network paths as a result of ECMP routing in an attempt to understand the underlying effects that cause variation in censorship. To achieve this we perform a deep dive to understand the concrete causes for such variation across the different countries and protocols. We find that *such variation is not due to a single global effect but rather a collection of different effects* ranging from routing that appears to exercise failed/misconfigured censoring nodes to routing that goes through completely different geographical regions that produce different censorship *en-route*.

We utilize *Monocle* to perform traceroutes (Section 3.1) to produce network graphs for a geographically diverse set of destinations for source parameters that showed variation, on a per country/protocol basis. Choosing a diverse set of destinations allows us to study potentially different effects at play even within a single country. Using these network graphs we find three type of district effects contributing to ECMP routing induced observed censorship variation.

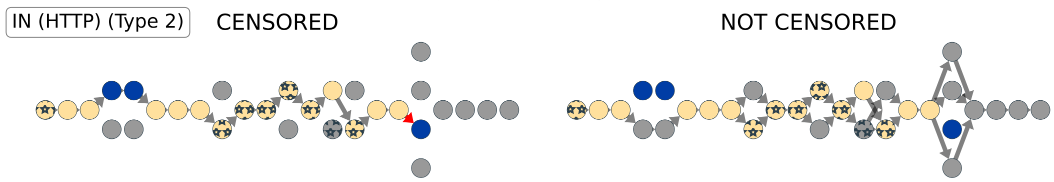
Figure 5 shows network graphs for a representative set of countries and protocols that demonstrate the different effects. For each destination we group all experiments that yield censorship into one graph (left) and all experiments that did not yield censorship into another graph (right). We now discuss these effects in greater detail in the context of how they manifest in each of the different countries and protocols. Appendix C in the extended version [76] contains additional graphs.

5.4.1 ***Type 1****: ECMP routing (Inter-AS or Intra-AS) exercising possibly failed/misconfigured censoring nodes.* Here, the packet is routed from a single node at a particular hop to a series of different nodes at the subsequent hop, that typically belong to an AS known to perform censorship. These different nodes often exist in the same /24 subnet and geographic region. *We observe different censorship behavior because some of these devices/paths perform censorship while others do not*, pointing to potentially failed/misconfigured censoring devices. This was one of the most commonly observed effects.

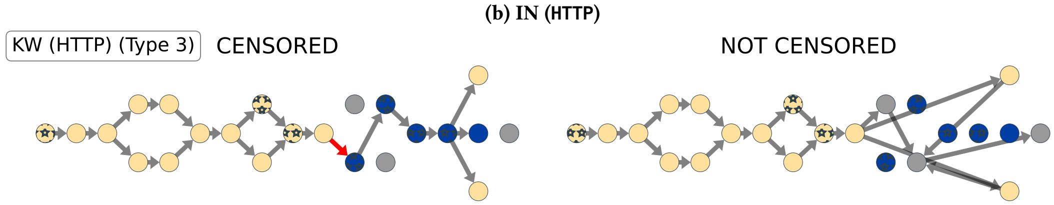
Identifying exact failed/misconfigured censoring nodes is challenging without detailed knowledge of the underlying censoring infrastructure, but we can infer this from comparing network graphs that have different observed censorship behavior side-by-side. We do this by identifying particular nodes in the censoring AS that are



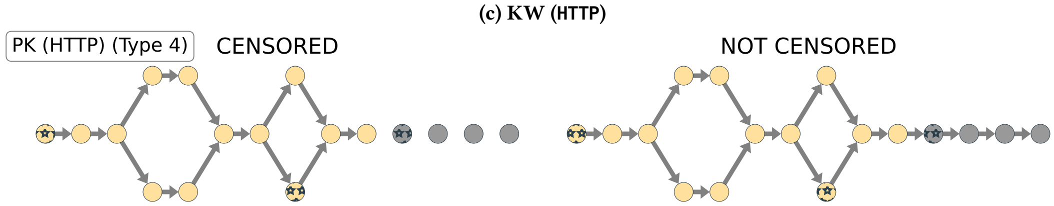
(a) ID (HTTPS)



(b) IN (HTTP)



(c) KW (HTTP)



(d) PK (HTTP)

Figure 5: Network graphs of censored routes (left) and non-censored routes (right). All nodes are shown in both, only routes change. Blue nodes were found only in paths that caused censorship, black found only in paths that had no censorship, yellow found in both, and censoring edges (if found) are red. We observe notable structural differences between cases.

only present in the paths (for source parameters) that *do not* exhibit censorship, suggesting failure or misconfiguration.

Inter versus Intra AS. When routing to different nodes in the censoring AS, ECMP routing can take place either: 1) completely within the censoring AS (Intra AS), *or* 2) from a completely different AS to the censoring AS (Inter AS). This highlights that variation in censorship measurement based on ECMP is not solely due to routing differences inside censoring ASes, but also stems from routing changes yielding different ASes before the censoring AS.

We observe this effect in Indonesia (HTTPS), shown in Figure 5a, as the sensitive packet transits from AS3491 (PCCW Global IP in Singapore) to AS7632 (PT Link Net, ISP in Indonesia known to perform censorship [48]), it is routed to several different IPs within several different /24 subnets in AS7632, some of which experience (RST) censorship while others do not. We also note that these nodes are mutually exclusive for the two behaviors.

We also observe this behavior in: 1) Belarus (HTTP and HTTPS) as the packets traverse through a set of mutually exclusive paths in AS6697 (Beltelecom, one of the ISPs known to perform blockpage based censorship [19, 49]), where one transits nodes that always produce blockpage censorship and the other passes through nodes with no censorship, 2) in Bangladesh (HTTP) as the packet undergoes ECMP routing from AS6939 (Hurricane Electric IP in Singapore) or AS2914 (NTT America IP in Singapore), to several different IPs in AS58717 (Summit Communications, a major ISP in Bangladesh known to perform censorship [50]), some of which perform blockpage/drop censorship, while others do not, 3) in Bangladesh (HTTPS) where ECMP routing entirely within AS17494 (BTCL, a large ISP in Bangladesh) causes the sensitive *client-hello* to pass through certain nodes that never experience censorship while others that always produce RST censorship, 4) in Turkey (HTTPS), as the sensitive packet transits from AS1299 (Arelion Sweden IP in Germany) to AS9121 (Turk Telekom, ISP in Turkey known to perform censorship [34]), it is routed to different IPs is the same /24 subnet in AS9121. On some of these paths, the sensitive packet experiences censorship while on others it does not, and 5) in South

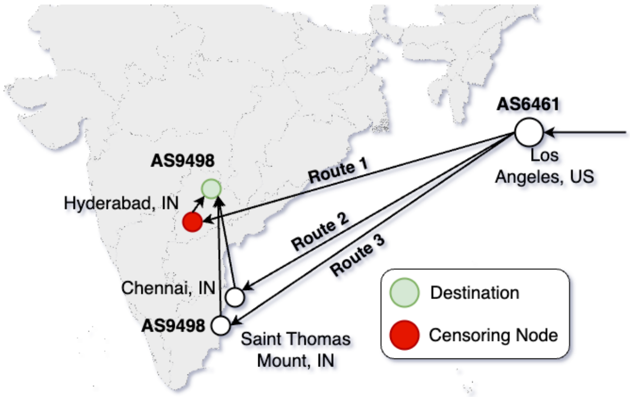


Figure 6: ECMP routing yielding *geographically* diverse routes. We observe that the sensitive packet is routed to different cities based on source IP or port, producing different observed censorship behavior. Censorship is only observed on Route 1. We note this exemplar is persistent.

Korea (HTTP) where the packet transits from AS9848 (Sejong Telecom) to AS4670 (Shinbiro), the packet takes two distinct paths going into AS4670; on one path the sensitive packet produces a blockpage while on the other it passes through to the destination and we get the response from the server. Although AS4670 is not known to perform censorship, it has customers (e.g., AS9848) that censor [30].

Such per-destination variation (presence or absence) of measured censorship due to source parameters causing the packet to pass through potentially failed/misconfigured censoring nodes could possibly manifest as non-uniform censorship across different destinations within a country/ISP, which could be a contributing factor in variation observation in prior work [17, 23, 28, 53, 55, 66, 68, 70].

5.4.2 ***Type 2****: ECMP routing through geographically different regions with different censorship behavior.* Here packets are routed from a particular AS to a fixed endpoint, but along the way transit IPs that are in different /24 subnets and are geographically diverse (but in the same AS). As a result of such persistent substantial change in the network path we observe different censorship behaviors. Accounting for such effects is critical for outside-in measurement so as to not conflate route/path specific censorship behavior as censorship behavior at the measurement destination.

We observe this in India (HTTP), shown in Figure 5b, where the deviation in path producing differing censorship behavior occurs when the packet transits ***from*** one of: 1) AS6461 (Zayo Bandwidth IP in London, UK, or in Los Angeles, USA), or 2) AS7473 (Singapore Telecom in Singapore)- ***to*** AS9498 (BHARTI Airtel, known to perform censorship [29, 37, 60]). The IP in AS6461 (or AS7473) routes the packet on IPs in noticeably different /24s in AS9498, sometimes in completely different cities simply based on the packet construction. Some of these paths produce a blockpage while others do not. Only a handful of routers in AS6461 and AS7473 control the path taken into AS9498 and consequently if we observe censorship.

Figure 6 demonstrates this geographically diverse *censorship-along-the-way* effect, with packets between two fixed endpoints taking different paths based on source parameters, those paths varying geographically (per Ipinfo [33]), and censorship correlating to geographic regions within India. Only a fraction of these paths pass through devices that perform censorship, owing to larger impact of ECMP on differences in observed censorship in India (Table 2). Prior work found AS-level variation in measured censorship for India [37, 73] which highlights the need to account for such ECMP routing induced censorship changes, and also the possibility that router regions, rather than end-host regions, are being measured.

5.4.3 ***Type 3****: ECMP routing* around *censorship.* In this mode we observe that at a particular hop, the packet undergoes ECMP routing, causing it to take completely different AS paths before reaching converging at the destination. Some of these paths never go through the censoring AS, experiencing no censorship, while others do.

We observed this effect in Kuwait (HTTP, Figure 5c), where depending on source parameters, the packet either transits directly from AS3356 (LEVEL3 in London, UK) to AS59605 (Zain Group), reaching the destination without censorship, *or* it had unresponsive hops between AS3356 and AS59605, where we observed censorship.

We also observed the effect in: 1) Kuwait (HTTPS) as the packet transits from a hop in AS3356 (LEVEL3, IP in UK), it either passes through several unresponsive hops leading to an IP in AS47589 (Kuwait Telecom, known to perform censorship [30]), where it terminates producing censorship, *or* transits directly to another IP in AS47589 eventually leading to the destination and no censorship occurs, and 2) Rwanda (HTTP) as the packet transits from AS16637 (MTN SA IP in Kenya) we observe two mutually exclusive paths, one that transits to another AS16637 IP in SA then AS36890 (MTN Rwandacell, large ISP in Rwanda) before reaching the destination, another that passes through a set of non-responsive hops and reaches the destination. The former path that passed through AS36890 always experiences censorship, while the latter does not.

Such per-destination variation in censorship caused by simply changing source parameters resulting in completely different AS paths not experiencing censorship *can also* manifest as non-uniform country/ISP level, which could be a contributing factor in variation observation in prior work [17, 23, 28, 53, 55, 66, 68, 70].

5.4.4 ***Type 4****: Behavior that is unknown or cannot be directly attributed to variation in path.* In this scenario the changes in censorship results cannot be directly attributed to an observable change in path. This could occur due to different forms censorship infrastructure in the path that do not answer with any ICMP messages *or* ones that are completely off-path. We observed this behavior with Pakistan (HTTP), shown in Figure 5d, as the packet transits from AS3356 (LEVEL3 IP in France) to AS17557 (Pakistan Telecom, known to perform censorship [41, 74]), there exists an unresponsive node where for some source parameters we observe censorship (packet drops) and for others we do not.

We also observed this behavior in: 1) Iran (HTTP), where both sets of source parameters have the same network path until they reach an IP in AS58224 (Iran Telecom) at which point a blockpage is always issued for certain set of source parameters while there is no censorship for others, and 2) Algeria (HTTP and HTTPS) where ECMP routing in AS174 (Cogent) produces mutually exclusive sub-paths (each for a combination of source parameters) that exhibit different observed censorship behavior. In this case as the packet leaves the

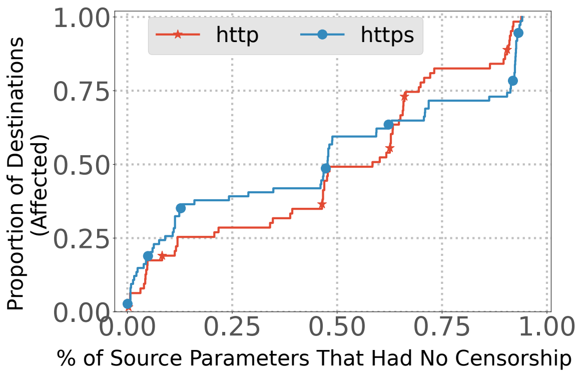


Figure 7: Simulation of impact of ECMP on a replication of prior work's [55] selection method. CDF of percent of (source IP, source Port) combinations per destination for which *we* observed no censorship for the sample set of destinations.

border router in AS174 towards AS36947 (Telecom Algeria), there are several unresponsive hops before the destination.

5.4.5 Result and Route Stability. Our conservative selection criteria specifically controlled for stable routes (by controlling for Flow-ID, Section 3.1). Additionally, our experiments took place over a two week window, requiring that the different routes and corresponding variation results were stable over that time horizon. We then conducted a spot-check manual exploration approximately 2 months after the initial experiments of the specific packet parameters and IPs that caused variation, and found much of the variation remained. This is consistent with prior work [51], that while dated, demonstrated that routes tend to be stable over time.

5.5 RQ5: How do these results contextualize with specific prior works?

We have qualitatively and quantitatively established the impact of ECMP routing on global outside-in censorship measurement while also investigating the underlying causes of such variation. Finally we want to contextualize prior results in the presence of ECMP. We first look at the applicability of ECMP router induced censorship changes on the different forms of variance observed in prior work (enumerated in Section 2). We then focus on specific findings/results from prior work that are potentially impacted by ECMP routing [53, 55].

5.5.1 Applicability in Prior Work. Section 2 detailed the different forms of country and ISP level inconsistencies in measured censorship in prior work. Not all of this variation is ECMP routing induced, with the root causes ranging from clearly demonstrated different censorship implementations across a country [37, 45, 71–73], to network load on censors manifesting as sporadic absences in censorship [17, 23]. We posit however that the per-destination changes in censorship observed in our results by changing the source parameters of the sensitive packet, can manifest as non-uniform censorship at a country/ISP level as observed in prior work [17, 23, 28, 53, 55, 66, 68, 70]. While we cannot conclude that ECMP routing is the (sole or contributing) cause for such variation ECMP routing is applicable to these methods and studies, and should be accounted for to identify variation root-causes and stabilize measurements. ECMP routing is also applicable for external measurements attempting to disambiguate ISP-granularity censorship differences, since prior work has found concrete evidence of differing censorship implementation at the ISP/AS-level [37, 45, 73] and our findings (Section 5.4.2) suggest that for the same destination, ECMP routing causes the probe to pass through completely different ASes with different censorship behavior.

5.5.2 Potential Impact of ECMP routing on 2 Prior Works. We now look at two prior outside-in studies focusing on: how their observations can potentially be explained by ECMP routing [53], and how their results could be impacted by ECMP routing [55], based on our results.

Global DNS censorship measurement. Our results can shed light on the causes of heterogeneity in observed censorship seen by prior DNS censorship measurement. e.g., Figure 7 from Pearce et al. [53] shows banding effects at ~10% for Iran and ~20% in China, which they attribute to non-determinism in censors. When we compare their results to our own, we find that ~7% of destinations for Iran, and ~25% of destinations for China are impacted by ECMP routing. While the age of the study and lack of specific source parameters precludes definitive conclusion, the similarity between results is consistent with ECMP routing induced censorship differences causing these banding effects. We note that Pearce et al. measures all forms of DNS manipulation, whereas we measure only "injected" DNS manipulation, which limits our comparison.

Global HTTP and HTTPS censorship measurement. In this case study we consider the example of a global outside-in censorship study conducted by Raman et al. [55]. We see that they choose either 11 (for HTTP) or 13 (for HTTPS) destinations per country for their study (from their Table 1 [55]). To understand the potential impact of ECMP routing on such a study we randomly sampled 11 (for HTTP) or 13 (for HTTPS) destinations per country that observed censorship in our experiments for RQ2. We found that across the countries out of 362 destinations that we sampled, *138 (~35%)* destinations showed some form of variations in observed censorship results. For this sample, Figure 7 shows the CDF of percent of source parameters that no censorship was observed on per destination. We find that *for a median affected destination, ~47% of source parameters produced no censorship.* As such, depending on the specifics of multiple experiments, repeated trials, and specific source parameters used, prior work may have observed what appeared to be failures or lack of censorship, when in reality they were observing routing-induced censorship changes.

6 RECOMMENDATIONS AND CONCLUDING DISCUSSION

Our work demonstrates the choice of intra-subnet source IP and ephemeral port influences censorship measurement routes, and those routing changes in turn impact observed censorship across DNS, HTTP, and HTTPS, globally. We show these variations are significant in terms of the number of affected countries, IPs, and source parameters, and explore *why* such changes exist. We note that censorship measurement is a critical tool for policymakers and evasion tool designers, and thus accurate understanding of methods, results, and confounding factors such as those discussed in this work are important for *correct* assessment of information control.

ECMP Routing Impacts Measurement, Globally. Prior work observed the Chinese GFW's DNS injection changing based on routing [7], but it was unclear if such behaviors were an idiosyncrasy of that system or a global phenomenon rooted in routing. We generalize our understanding of the intersection between routing and censorship measurement, discovering that ECMP routing's impact on censorship measurement is both pervasive and significant. We also observed differences in how particular source parameters influence censorship results and found not just source IPs but source ports can have an impact on censorship results. These differences in changes point to the complexity of routing and distributed systems, and call for further work in understanding localized observed differences in censorship measurement globally. We further contextualized results with prior work and found instances where non-determinism in prior global censorship measurement is consistent with this phenomenon. We also note that this phenomenon may impact other forms of measurement, such as fast Internet scanning.

Causes of Variation. We also explored *why* variation exists and found several causes: 1) ECMP routing exercising likely failed or misconfigured censoring devices, 2) ECMP routing causing a difference in AS path to cause "routing around" censorship behavior, and 3) ECMP routing causing the packets to traverse completely different geographical regions, producing different censorship behavior *en-route*. From this exploration we learn that in addition to diversity in source parameters, censorship studies must account for potentially observing different censorship behaviors *along* the path rather than *at* the destination, when performing measurement.

Extending to IPv6 Censorship Measurement. IPv6 censorship is broadly under-studied, irrespective of ECMP routing. Given that prior work [1] found that ECMP is *more* prevalent (75% of their measured routes) in IPv6 than IPv4, we believe that its impact on censorship measurement could be commiserate. Thus ECMP routing should be taken into account as a first order concern as nascent IPv6 censorship measurement studies are conducted. We note that the the presence of the *Flow Label* field in IPv6 may aid in methods to produce route-stable censorship measurements.

Recommendations for Future Studies. We find that diversity in both source IP and source port is critical when performing censorship measurement to avoid potentially incorrect results or what appears to be transient failures, globally, across protocols. We also note repetition is necessary to differentiate between *network effects* and *actual censorship variation*. Measurements must be repeated with significant parameter diversity, and localized geographic effects must be validated with route diversity. Finally, we find significant differences in measured censorship in some countries by protocol. These differences indicate a need for cross-protocol measurement for a complete picture of Internet censorship.

Future Work and Evasion. Our work focuses on outside-in measurement that is, by design, external. Future work aimed at extending these observations to volunteer vantage points within a country is needed. We however note such work is potentially challenging as volunteer systems may not have access to numerous IPs, or repeated experiments may be ethically challenging. Further, the prevalence across both countries and protocols, as well as the overlap between protocols for given IPs, suggests the viability of future work leveraging these differences to construct packets to *route-around* censorship, thus effectively evading. We note that Tor Bridges [40] are a particularly apt potential use.

7 ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers and the shepherd for all their thoughtful feedback provided during the review process. This work was supported by funding from the National Science Foundation (NSF) CAREER award 2239183.

REFERENCES

[1] Rafael Almeida, Osvaldo Fonseca, Elverton Fazzion, Dorgival Guedes, Wagner Meira, and Ítalo Cunha. 2017. A characterization of load balancing on the IPv6 Internet. In *Passive and Active Measurement (PAM 17)*.

[2] Anonymous. 2014. Towards a Comprehensive Picture of the Great Firewall's DNS Censorship. In *USENIX FOCI*.

[3] Simurgh Aryan, Homa Aryan, and J Alex Halderman. 2013. Internet Censorship in Iran: A First Look. In *USENIX FOCI*.

[4] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira. 2006. Avoiding traceroute anomalies with Paris traceroute. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. 153–158.

[5] Brice Augustin, Timur Friedman, and Renata Teixeira. 2010. Measuring multipath routing in the internet. *IEEE/ACM Transactions on Networking* (2010).

[6] belmont 1978. *The Belmont Report - Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html.

[7] Abhishek Bhaskar and Paul Pearce. 2022. Many Roads Lead To Rome: How Packet Headers Influence DNS Censorship Measurement. In *USENIX Security Symposium (USENIX)*.

[8] Dave Levin Kevin Bock. 2020. Exposing and Circumventing China's Censorship of ESNI.

[9] Kevin Bock, Pranav Bharadwaj, Jasraj Singh, and Dave Levin. 2021. Your censor is my censor: Weaponizing censorship infrastructure for availability attacks. In *2021 IEEE Security and Privacy Workshops (SPW)*. IEEE, 398–409.

[10] Kevin Bock, Gabriel Naval, Kyle Reese, and Dave Levin. 2021. Even Censors Have a Backup: Examining China's Double HTTPS Censorship Middleboxes. In *Proceedings of the ACM SIGCOMM 2021 FOCI*.

[11] CAIDA. [n. d.]. Routeviews Prefix to AS mappings Dataset (pfx2as) for IPv4. https://www.caida.org/catalog/datasets/routeviews-prefix2as/.

[12] Zimo Chai, Amirhossein Ghafari, and Amir Houmansadr. 2019. On the Importance of Encrypted-SNI(ESNI) to Censorship Circumvention. In *USENIX FOCI*.

[13] Cisco. 2007. Understanding etherchannel load balancing and redundancy on catalyst switches. https://www.cisco.com/c/en/us/support/docs/lan-switching/etherchannel/12023-4.html.

[14] Cisco. 2021. LAG Load Balancing on Cisco 350 and 550 Series Switches. https://www.cisco.com/c/en/us/support/docs/smb/switches/cisco-550/x-series-stackable-managed-switches/1388-LAG-Load-Balancing-on-Cisco-350-and-550-series-Switches.pdf .

[15] Citizen Lab. [n. d.]. Block Test List. https://github.com/citizenlab/test-lists.

[16] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson. 2006. Ignoring the Great Firewall of China. In *Privacy Enhancing Technologies Symposium (PETS)*.

[17] Jedidiah R. Crandall, Daniel Zinn, Michael Byrd, Earl Barr, and Rich East. 2007. ConceptDoppler: A Weather Tracker for Internet Censorship. In *Computer and Communications Security*. ACM, Alexandria, VA, USA, 352–365.

[18] D. Dittrich and E. Kenneally. 2012. *The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research*. Technical Report. U.S. Department of Homeland Security.

[19] Shawn P Duncan and Hui Chen. 2023. Detecting network-based internet censorship via latent feature representation learning. *Computers & Security* (2023).

[20] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, and J. Alex Halderman. 2015. A Search Engine Backed by Internet-Wide Scanning. In *ACM Conference on Computer and Communications Security (CCS)*.

[21] Kathrin Elmenhorst, Bertram Schütz, Nils Aschenbruck, and Simone Basso. 2021. Web censorship measurements of HTTP/3 over QUIC. In *21st ACM IMC*.

[22] Roya Ensafi, Jeffrey Knockel, Geoffrey Alexander, and Jedidiah R. Crandall. 2014. Detecting Intentional Packet Drops on the Internet via TCP/IP Side Channels. In *Workshop on Passive and Active Network Measurement (PAM)*.
[23] Roya Ensafi, Philipp Winter, Abdullah Mueen, and Jedidiah R. Crandall. 2015. Analyzing the Great Firewall of China Over Space and Time. *Privacy Enhancing Technologies Symposium (PETS)* (2015).
[24] Oliver Farnan, Alexander Darer, and Joss Wright. 2016. Poisoning the Well – Exploring the Great Firewall's Poisoned DNS Responses. In *ACM Workshop on Privacy in the Electronic Society (WPES)*. http://dl.acm.org/authorize?N25517
[25] Sergiu Gatlan. 2019. South Korea is censoring the Internet by snooping on SNI traffic.
[26] gfw report. 2022. Large scale blocking of TLS-based censorship circumvention tools in China. https://github.com/net4people/bbs/issues/129.
[27] Phillipa Gill, Masashi Crete-Nishihata, Jakub Dalek, Sharon Goldberg, Adam Senft, and Greg Wiseman. 2008. Characterizing Censorship of Web Content Worldwide.
[28] Phillipa Gill, Masashi Crete-Nishihata, Jakub Dalek, Sharon Goldberg, Adam Senft, and Greg Wiseman. 2015. Characterizing Web Censorship Worldwide: Another Look at the OpenNet Initiative Data. *Transactions on the Web* 9, 1 (2015).
[29] Devashish Gosain, Anshika Agarwal, Sahil Shekhawat, H. B. Acharya, and Sambuddho Chakravarty. 2018. Mending Wall: On the Implementation of Censorship in India. In *Security and Privacy in Communication Networks*.
[30] Nguyen Phong Hoang, Sadie Doreen, and Michalis Polychronakis. 2019. Measuring I2P Censorship at a Global Scale. In *USENIX FOCI*.
[31] 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 *Proceedings of the 30th USENIX Security*.
[32] Freedom House. 2016. Silencing the Messenger: Communication Apps Under Pressure. https://freedomhouse.org/report/freedom-net/2016/silencing-messenger-communication-apps-under-pressure
[33] ipinfo [n. d.]. IPInfo.IO. https://ipinfo.io/.
[34] Lin Jin, Shuai Hao, Haining Wang, and Chase Cotton. 2021. Understanding the Practices of Global Censorship through Accurate, End-to-End Measurements. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*.
[35] Ben Jones, Tzu-Wen Lee, Nick Feamster, and Phillipa Gill. 2014. Automated Detection and Fingerprinting of Censorship Block Pages. In *ACM Internet Measurement Conference (IMC)*.
[36] Ben Jones Nick Feamster Joseph Lorenzo Hall, Michael D. Aaron. 2014. A Survey of Worldwide Censorship Techniques. (2014). https://tools.ietf.org/id/draft-hall-censorship-tech-05.html
[37] Divyank Katira, Gurshabad Grover, Kushagra Singh, and Varun Bansal. 2023. CensorWatch: On the Implementation of Online Censorship in India. (2023).
[38] Sheharbano Khattak, Mobin Javed, Syed Ali Khayam, Zartash Afzal Uzmi, and Vern Paxson. 2014. A Look at the Consequences of Internet Censorship Through an ISP Lens. In *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, 271–284. http://dl.acm.org/citation.cfm?id=2663750
[39] Graham Lowe, Patrick Winters, and Michael L. Marcus. 2007. *The Great DNS Wall of China*. Technical Report. New York University. http://cs.nyu.edu/~pcw216/work/nds/final.pdf
[40] Srdjan Matic, Carmela Troncoso, and Juan Caballero. 2017. Dissecting Tor Bridges.
[41] Zubair Nabi. 2013. The Anatomy of Web Censorship in Pakistan. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*.
[42] 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) (Oakland 2020)*. IEEE Computer Society.
[43] Arian Akhavan Niaki, Nguyen Phong Hoang, Phillipa Gill, Amir Houmansadr, et al. 2020. Triplet Censors: Demystifying Great Firewall's DNS Censorship Behavior. In *10th USENIX FOCI*.
[44] Arian Akhavan Niaki, William R. Marczak, Sahand Farhoodi, Andrew McGregor, Phillipa Gill, and Nicholas Weaver. 2021. Cache Me Outside: A New Look at DNS Cache Probing. In *Passive and Active Measurement - 22nd International Conference, PAM 2021, Virtual Event, March 29 - April 1, 2021, Proceedings*.
[45] Aqib Nisar, Aqsa Kashaf, Ihsan Ayyub Qazi, and Zartash Afzal Uzmi. 2018. Incentivizing censorship measurements via circumvention. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*. 533–546.
[46] Sadia Nourin, Van Tran, Xi Jiang, Kevin Bock, Nick Feamster, Nguyen Phong Hoang, and Dave Levin. 2023. Measuring and Evading Turkmenistan's Internet Censorship: A Case Study in Large-Scale Measurements of a Low-Penetration Country. In *Proceedings of the ACM Web Conference 2023*. 1969–1979.
[47] ooni [n. d.]. Ooniprobe: A Network Interference Detection Tool. https://github.com/thetorproject/ooni-probe
[48] OONI. 2017. The State of Internet Censorship in Indonesia. https://ooni.org/post/indonesia-internet-censorship/
[49] OONI. 2020. Belarus protests: From internet outages to pervasive website censorship. https://ooni.org/post/2020-belarus-internet-outages-website-censorship/
[50] OONI. 2022. OONI Explorer. https://explorer.ooni.org/search?probe_cc=BD&since=2022-11-01&until=2023-11-01&failure=true&only=confirmed&probe_asn=58717
[51] Vern Paxson. 1997. End-to-end routing behavior in the Internet. *IEEE/ACM transactions on Networking* 5, 5 (1997), 601–615.
[52] Paul Pearce, Roya Ensafi, Frank Li, Nick Feamster, and Vern Paxson. 2017. Augur: Internet-Wide Detection of Connectivity Disruptions. In *IEEE Symposium on Security and Privacy (S&P)*.
[53] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nick Weaver, and Vern Paxson. 2017. Global Measurement of DNS Manipulation. In *USENIX Security Symposium (USENIX)*.
[54] Ram Sundara Raman, Leonid Evdokimov, Eric Wurstrow, J Alex Halderman, and Roya Ensafi. 2020. Investigating large scale HTTPS interception in Kazakhstan. In *Proceedings of the ACM Internet Measurement Conference*. 125–132.
[55] Ram Sundara Raman, Adrian Stoll, Jakub Dalek, Reethika Ramesh, Will Scott, and Roya Ensafi. 2020. Measuring the Deployment of Network Censorship Filters at Global Scale.. In *NDSS*.
[56] Ram Sundara Raman, Mona Wang, Jakub Dalek, Jonathan 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*. 18–34.
[57] Reethika Ramesh, Ram Sundara Raman, Matthew Bernhard, Victor Ongkowijaya, Leonid Evdokimov, Anne Edmundson, Steven Sprecher, Muhammad Ikram, and Roya Ensafi. 2020. Decentralized control: A case study of russia. In *NDSS 20*.
[58] Reethika Ramesh, Ram Sundara Raman, Apurva Virkud, Alexandra Dirksen, Armin Huremagic, David Fifield, Dirk Rodenburg, Rod Hynes, Doug Madory, and Roya Ensafi. 2023. Network responses to Russia's invasion of Ukraine in 2022: a cautionary tale for internet freedom. In *32nd USENIX Security Symposium*.
[59] Will Scott, Thomas Anderson, Tadayoshi Kohno, and Arvind Krishnamurthy. 2016. Satellite: Joint Analysis of CDNs and Network-Level Interference. USENIX Association.
[60] Kushagra Singh, Gurshabad Grover, and Varun Bansal. 2020. How India censors the web. In *Proceedings of the 12th ACM Conference on Web Science*. 21–28.
[61] Ram Sundara Raman, Prerana Shenoy, Katharina Kohls, and Roya Ensafi. 2020. Censored Planet: An Internet-wide, Longitudinal Censorship Observatory. In *Proceedings of the 2020 ACM SIGSAC CCS*.
[62] transmission control 1981. Transmission Control Protocol. RFC 793.
[63] Benjamin VanderSloot, Allison McDonald, Will Scott, J. Alex Halderman, and Roya Ensafi. 2018. Quack: Scalable Remote Measurement of Application-Layer Censorship. In *27th USENIX Security Symposium (USENIX Security 18)*.
[64] Kevin Vermeulen, Justin P Rohrer, Robert Beverly, Olivier Fourmaux, and Timur Friedman. 2020. Diamond-Miner: Comprehensive Discovery of the Internet's Topology Diamonds. In *17th USENIX NSDI*. 479–493.
[65] VMWARE. 2020. Understanding IP Hash load balancing (2006129). .
[66] Zhongjie Wang, Yue Cao, Zhiyun Qian, Chengyu Song, and Srikanth V Krishnamurthy. 2017. Your state is not mine: a closer look at evading stateful internet censorship. In *Proceedings of the 2017 Internet Measurement Conference*. 114–127.
[67] Nicholas Weaver, Robin Sommer, and Vern Paxson. 2009. Detecting Forged TCP Reset Packets.. In *NDSS*.
[68] Zachary Weinberg, Diogo Barradas, and Nicolas Christin. 2021. Chinese Wall or Swiss Cheese? Keyword filtering in the Great Firewall of China. In *Proceedings of the Web Conference 2021*.
[69] Florian Westphal. 2016. bridge filtering with nftables. (2016).
[70] Philipp Winter and Stefan Lindskog. 2012. How the Great Firewall of China is Blocking Tor. In *USENIX FOCI*.
[71] Joss Wright. 2014. Regional variation in Chinese internet filtering. *Information, Communication & Society* (2014).
[72] Xueyang Xu, Z. Morley Mao, and J. Alex Halderman. 2011. Internet Censorship in China: Where Does the Filtering Occur?. In *Workshop on Passive and Active Network Measurement (PAM)*. http://www.eecs.umich.edu/~zmao/Papers/china-censorship-pam11.pdf
[73] Tarun Kumar Yadav, Akshat Sinha, Devashish Gosain, Piyush Kumar Sharma, and Sambuddho Chakravarty. 2018. Where The Light Gets In: Analyzing Web Censorship Mechanisms in India. In *Proceedings of the Internet Measurement Conference 2018* (Boston, MA, USA) *(IMC '18)*. Association for Computing Machinery, New York, NY, USA, 252–264. https://doi.org/10.1145/3278532.3278555
[74] Ethan Zuckerman. 2008. How a Pakistani ISP briefly shut down YouTube. https://ethanzuckerman.com/2008/02/25/how-a-pakistani-isp-briefly-shut-down-youtube/
[75] Ethan Zuckerman. 2010. Intermediary Censorship. In *Access Controlled - The Shaping of Power, Rights, and Rule in Cyberspace*. MIT Press, 71–86.
[76] Abhishek Bhaskar and Paul Pearce 2024. *Understanding Routing-Induced Censorship Changes Globally*. arXiv preprint. https://arxiv.org/abs/2406.19304