Detecting Network Interference Without Endpoint Participation

Extended Abstract

Sadia Nourin* Kevin Bock* Nguyen Phong Hoang† Dave Levin*
*University of Maryland †University of Chicago

1 Introduction

Authoritarian regimes around the world deploy pervasive censorship apparatuses to restrict free and open communications on the Internet. Extensive censorship measurement efforts have, over the years, shed impressive light on how censors operate. Typically, prior efforts have included recruiting volunteers to run network probes [8, 15], relying on public servers [16, 17, 20, 21], or using remote vantage points (e.g., VPNs) [10, 12, 13]. While each powerful in their own right, these approaches tend to concentrate on large countries that have high Internet penetration, such as China [1, 9, 11], Iran [2, 4], and Russia [18].

Several challenges limit the use of these methods in smaller countries: low Internet penetration rate, small populations, and scarcity of public servers or remote vantage points make it difficult to recruit volunteers or identify safe endpoints to test. Even when one is able to overcome these challenges with a rare vantage point, there may be large periods of no or low measurement.

In this extended abstract, we sketch the design of a system that can longitudinally detect Internet censorship *without relying on local endpoints or volunteers inside a censoring nation*. Our system includes a set of measurement techniques that can be run from vantage points outside a censoring regime with no participating endpoints within. We emphasize that our goal is not to replace existing measurement efforts, but to complement them by providing insight into networks they cannot reach.

The central idea behind our approach is to take advantage of the fact that many censoring middleboxes are not TCP compliant [3, 5]. All middleboxes have to assume that they may miss some network packets within a connection. Bock et al. [3] showed that this assumption could be exploited to trick HTTP middleboxes into injecting block pages and other traffic without any communication with a destination, let alone first establishing a TCP connection.

In our earlier work, we demonstrated how to apply Bock et al.'s approach of triggering censors with specific packet sequences to measure the censorship infrastructure of Turkmenistan [3, 14]. However, our earlier approach relied on manual effort to discover a packet sequence that could trigger the censor. Such manual efforts will not feasibly scale to other countries, or even to multiple ISPs within a single country with decentralized censorship.

The system we present in this paper seeks to operate on a *global* scale by automating the discovery of censorship-triggering packet sequences. Our system uses Geneva [6], a genetic algorithm-based tool that trains directly against censors to discover packet sequences that can achieve a given goal. While originally used to discover evasion strategies [5, 6, 9], we have modified Geneva to discover sequences of network packets that can trick censors to take blocking actions. Critically, our application of Geneva does not rely on any participation from an endpoint inside censoring countries. However, it does require certain properties of the censor itself, which we describe. While we mainly focus on discovering censorship triggering packet sequences for HTTP censorship in this paper, our approach is extensible and applicable for other network protocols as well (e.g., HTTPS).

In the remainder of this paper, we discuss the high-level design of this measurement system, and we introduce some preliminary results from detecting censorship in two previously under-explored nations: Brunei and Tajikistan.

2 High-level Design

In order for a country to be eligible for our censorship detection approach, we need the following:

Censored domain First, we need at least one confirmed censored domain for each AS within the country in order to test whether a packet sequence triggers censorship. These domains can be obtained from available sources such as OONI [8] or CensoredPlanet [17], from anecdotal information given by residents of the country, or from other reputable sources.

Bidirectional censorship Second, the country must have bidirectional censorship. This is to ensure that we can suc-

cessfully trigger censorship from outside the censored country using our own machines, as it may be infeasible to obtain machines or vantage points within the country.

We detect bidirectional censorship in an automated fashion by running a ZMap scan on the country and finding all live HTTP servers. We send a request containing an uncensored domain to the servers followed by a request for a censored domain. Assuming that none of the servers host the actual domains that we have sent requests for, if there is a difference in the responses we receive, we can confirm bidirectional censorship.

Little to no residual censorship Finally, we require that the country does not have excessive *residual censorship*. Residual censorship occurs when a censor filters *all* requests from a client for an extended period of time after it has determined that the client sent a censored request. Residual censorship may impact our measurements and add false positives to our results, as an innocuous domain may seem to be censored, when it is really just subjected to residual censorship. We determine whether there is residual censorship by sending a censored query followed by an innocuous query to the same server and port, and determining whether both queries were censored. In the event that there is residual censorship, we must either circumvent it (e.g., altering port numbers works in some countries) or simply wait for it to expire.

When a censorship infrastructure meets the above conditions, it permits an approach similar to what was described in our recent study of Turkmenistan [14]. Specifically, we can use an extension to the automated tool Geneva that determines packet sequences to send that trigger censorship. Originally, Geneva was created to *circumvent* censorship, but a change to its fitness function allows us to discover how to *get* censored, not how to avoid it.

Most importantly, we train Geneva by sending packets to *non-responsive IP addresses*. This has two important benefits: First, it simplifies training, as we can assume that any responses we receive are from middleboxes. Second, when Geneva finds a successful packet sequence to trigger censorship without requiring responses from a destination, then it means that Geneva can be applied to *all* IP addresses in that network (e.g., we could TTL-limit the packets so they cross the censoring middleboxes, but do not reach the destination).

Once Geneva discovers a packet sequence that triggers censorship, we can conduct a widespread measurement of the country by sending requests to non-responsive IP addresses in the country and observing responses we receive, if any.

3 Preliminary Results

Here, we present preliminary results evaluating how applicable our methodology is. We report on two countries that meet our requirements, and discuss several countries that appear not to.

Brunei Brunei fits all of the requirements for our censorship detection approach. It is relatively small with 132,000 IP addresses, it has bidirectional censorship in the form of RST packet injections within one of its seven active ASes (AS10094), and this AS covers approximately 70% of the country's Internet users [19]. Brunei's HTTP censorship occurs on all ports and there is no residual censorship in the country. In addition, Brunei's censor has a clear fingerprint—it injects a RST packet with the censored query's IP ID field.

Geneva discovered that a SYN, followed immediately by a PSH+ACK with a payload is able to trigger censorship within the country. As described in Bock et al. [3], this packet sequence succeeds in triggering the censor as the censor presumes the TCP three way handshake has already completed and injects a RST packet based on the censored HTTP Host Header in the payload. We can use this packet sequence to conduct wide-scale measurement in this AS.

Tajikistan Tajikistan also fits our requirements. Tajikistan has 79,000 IP addresses and has a centralized censorship system due to a national decree announced in 2016 in which all national egress and ingress traffic is to pass through AS51346, a state-run telecommunication company, The Opened Joint Stock Company Tojiktelecom [7]. As such, virtually every traceroute in the country passes through AS51436.

We were able to find bidirectional censorship within one AS so far, AS24722. Interestingly, the national gateway that resides in AS51346 *is* the censor, as they are located at the exact same hop. Naturally then, one can ask why the other ASes do not trigger bidirectional censorship if all traffic is routed through the censor. We do not know why this is the case but believe that there may be certain policies in place.

The censor injects a RST+ACK packet that has a unique fingerprint containing a 22 byte payload of zeroes, does not employ residual censorship, and censors on all ports. When we train Geneva against this specific AS in Tajikistan, we discover a censorship triggering strategy in which we send a PSH+ACK packet with a censored payload *twice* before we receive the censor's injected RST+ACK response.

Previous work has discovered that a single PSH+ACK packet can trigger censorship in other countries [3]. Sending a PSH+ACK packet twice seems redundant, but it may be reflective of a more typical client browsing behavior in which many PSH+ACK packets are sent to fetch multiple resources from the same origin to render a full webpage. Regardless, since there is no residual censorship, we can use this strategy for wide-scale measurement in this AS.

Negative Results We also have attempted to apply our methodology in countries that do not appear to satisfy the required criteria. In particular, our experiments with Burundi, Equatorial Guinea, Myanmar, and Kyrgyzstan were not successful as we were unable to detect bidirectional censorship.

4 Conclusion

In this paper we outline the beginnings of a newfound censorship detection approach that can fill a gap in the censorship measurement community by providing a longitudinal measurement system that does not depend on endpoint participation. This approach exploits a design choice made by censoring middleboxes to not fully adhere to the TCP protocol and to censor bidirectionally. With Geneva, we automatically discover censorship triggering packet sequences for two countries using this approach and leave conducting the actual measurements of these countries to future work.

Acknowledgments

We thank the anonymous reviewers and our shepherd, Paul Pearce, for their helpful feedback. This work was supported in part by NSF award CNS-1943240 and an Internet Society Pulse Research Fellowship.

References

[1] Anonymous, Arian Akhavan Niaki, Nguyen Phong Hoang, Phillipa Gill, and Amir Houmansadr. Triplet Censors: Demystifying Great Firewall's DNS Censorship Behavior. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, 2020.

[2] Simurgh Aryan, Homa Aryan, and J. Alex Halderman. Internet Censorship in Iran: A First Look. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, 2013.

[3] Kevin Bock, Abdulrahman Alaraj, Yair Fax, Kyle Hurley, Eric Wustrow, and Dave Levin. Weaponizing Middleboxes for TCP Reflected Amplification. In *USENIX Security Symposium*, 2021.

[4] Kevin Bock, Yair Fax, Jasraj Singh, Kyle Reese, and Dave Levin. Iran: A New Model for Censorship. https://geneva.cs.umd.edu/posts/iran-whitelister.

[5] Kevin Bock, George Hughey, Louis-Henri Merino, Tania Arya, Daniel Liscinsky, Regina Pogosian, and Dave Levin. Come as You Are: Helping Unmodified Clients Bypass Censorship with Server-Side Evasion. In *ACM SIGCOMM*, 2020.

[6] Kevin Bock, George Hughey, Xiao Qiang, and Dave Levin. Geneva: Evolving Censorship Evasion Strategies. In *ACM Conference on Computer and Communications Security (CCS)*, 2019.

[7] eurasianet. Tajikistan: Data Gateway Deals Blow to Internet Freedom. https://eurasianet.org/tajikistan-data-gateway-deals-blow-to-internet-freedom, 2016-02-16.

[8] Arturo Filasto and Jacob Appelbaum. OONI: Open Observatory of Network Interference. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, 2012.

[9] Michael Harrity, Kevin Bock, Frederick Sell, and Dave Levin. GET /out: Automated Discovery of Application-Layer Censorship Evasion Strategies. In *USENIX Security Symposium*, 2022.

[10] Nguyen Phong Hoang, Sadie Doreen, and Michalis Polychronakis. Measuring I2P Censorship at a Global Scale. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*. 2019.

[11] Nguyen Phong Hoang, Arian Akhavan Niaki, Jakub Dalek, Jeffrey Knockel, Pellaeon Lin, Bill Marczak, Masashi Crete-Nishihata, Phillipa Gill, and Michalis Polychronakis. How Great is the Great Firewall? Measuring China's DNS Censorship. In *USENIX Security Symposium*, 2021.

[12] Nguyen Phong Hoang, Michalis Polychronakis, and Phillipa Gill. Measuring the Accessibility of Domain Name Encryption and its Impact on Internet Filtering. In *Passive and Active Network Measurement Conference (PAM)*, 2022.

[13] Arian Akhavan Niaki, Shinyoung Cho, Zachary Weinberg, Nguyen Phong Hoang, Abbas Razaghpanah, Nicolas Christin, and Phillipa Gill. Iclab: A global, longitudinal internet censorship measurement platform. *CoRR*, abs/1907.04245, 2019.

[14] Sadia Nourin, Van Tran, Xi Jiang, Kevin Bock, Nick Feamster, Nguyen Phong Hoang, and Dave Levin. Measuring and Evading Turkmenistan's Internet Censorship. In *International World Wide Web Conference (WWW)*, 2023.

[15] OpenNet Initiative. https://opennet.net, 2014.

[16] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick. Feamster, Nick Weaver, and Vern Paxson. Global Measurement of DNS Manipulation. In *USENIX Security Symposium*, 2017.

[17] Ram Sundara Raman, Prerana Shenoy, Katharina Kohls, and Roya Ensafi. Censored Planet: An Internet-wide, Longitudinal Censorship Observatory. In *ACM Conference on Computer and Communications Security (CCS)*, 2020.

[18] Reethika Ramesh, Ram Sundara Raman, Matthew Bernhard, Victor Ongkowijaya, Leonid Evdokimov, Anne

Edmundson, Steven J. Sprecher, Muhammad Ikram, and Roya Ensafi. Decentralized control: A case study of russia. In *Network and Distributed System Security Symposium (NDSS)*, 2020.

[19] RIPE. RIPE Atlas Population Coverage - Brunei. https://sg-pub.ripe.net/petros/population_coverage/country.html?name=BN.

[20] Will Scott, Thomas Anderson, Tadayoshi Kohno, and Arvind Krishnamurthy. Satellite: Joint Analysis of CDNs and Network-Level Interference. In *USENIX Annual Technical Conference*, 2016.

[21] Benjamin VanderSloot, Allison McDonald, Will Scott, J. Alex Halderman, and Roya Ensafi. Quack: Scalable Remote Measurement of Application-Layer Censorship. In *USENIX Security Symposium*, 2018.