The Impact of Geoblocking and the U.S. Embargo on Internet Freedom in Cuba

ABSTRACT

Today, service operators increasingly play a crucial role in implementing geolocation-based discrimination and restricting Internet freedom. A majority of such geoblocking is applied to abide by regulations and legal restrictions such as the U.S. embargoes. In this paper, we conduct an in-depth investigation into the impact of the U.S. embargo and geoblocking on the Internet freedom in Cuba. We help empirically-inform our analysis by conducting a qualitative study of tech-savvy Internet users in Cuba, and through these efforts, we curate a test list of 12,384 domains and conduct large-scale measurements from a vantage point in Cuba's largest ISP. We document the various ways that geoblocking is implemented by popular domains in a variety of categories. We highlight that the lack of standardization and accountability surrounding overcompliance to U.S. embargoes results in access disparity and often ends up burdening the users. We hope that our work raises awareness into the effect of large-scale geoblocking due to embargoes.

1 INTRODUCTION

It is becoming increasingly evident that service operators play an active role in violating Internet freedom and discriminating against users for a variety of reasons. As a prime exemplar of this phenomenon, we witnessed the damaging, compounding effects of both censorship and geoblocking due to hasty sanctions imposed on Russia after their invasion of Ukraine in 2022 [33, 38, 39]. As a result, civil society groups condemned these sanctions in a letter to the U.S. president [1] and emphasized that isolating users would only exacerbate disinformation and result in the formation of a "splinternet." In Cuba, the sanctions imposed by the U.S. government were proposed in 1962 and continue to impede innovation and development of infrastructure [9, 14, 19]. With limited previous work studying the Cuban internet, and the closest works dating back over six to eight years ago [6, 29], current information on what web services are accessible for users in the country remains understudied.

Characterizing Cuba's connectivity in a detailed manner, especially due to the presence of the embargo, is challenging. Not only is it difficult to determine what to test and how to gain a vantage point, but also, because some services implement geoblocking using timeouts, connection resets, or non-informative errors, differentiating geoblocking from in-path filtering or censorship is complex.

In this paper, we aim to investigate the impact of geoblocking especially due to the effect of U.S. embargo on Internet freedom in Cuba. Geoblocking is a practice implemented by server-side administrators to intentionally block users in specific geographical regions often due to compliance with agreements and legal regulations. To help inform our analysis and to gain a better understanding of the human cost of geoblocking, we reach out to 10 technically-savvy users in Cuba to learn about the pain points of being a target of geoblocking, and collect an empirically-informed list of domains and category of web services to test. Thereafter, we work with our contact in the country who has extensive experience in the Internet freedom community in Cuba and, with their explicit consent, conduct measurements from their vantage point in the country's largest ISP, ETECSA.

We curate a test list of 12,384 domains and perform measurements on multiple-layers of the network stack, and use modified traceroute methods to identify cases of geoblocking. In total, we confirm geoblocking by 717 domains (5.79%), which is an order of magnitude larger than the 0.66% reported by previous work on CDN geoblocking in Cuba [22]. Interestingly, we find that more popular domains implement significantly more geoblocking than others; almost 30% of the geoblocking seen in Tranco 10K (142 of 481) is implemented by just the top 2K domains. We manually investigate the terms of services of these 142 domains and find that 66.2% mention restriction of service due to the U.S. embargo.

The lack of standardization leads to service providers implementing geoblocking in inconsistent ways. We find that 70.3% (504/717) of the geoblocking is implemented using HTTP response status codes (403 Forbidden, 451 Unavailable For Legal Reasons, 404 Not Found) and less than 10% of these responses have any direct mention of legal reasons, sanctions, or country-specific geoblocking. These differences reinforce the sentiments of confusion expressed by the participants in our qualitative study, and also researchers seeking to correlate geoblocking with the embargo. Investigating the categories, we unsurprisingly see that economy, finance, and e-commerce domains implement significant geoblocking, but we also find 8.58% of technology (*e.g.* `zoom.us`), 3.41% of education (*e.g.* `udemy.com`), and 11.69% of job search and career-related domains (*e.g.* `monster.com`) geoblock Cuban users. In our qualitative study, we see participants express that the unavailability of these categories of domains affects their daily life.

Without a study such as ours, the pernicious nature of overblocking under the guise of U.S. embargo and sanctions goes completely unnoticed. In an effort to encourage future work in this area, we will open-source our code, data, and the 14 fingerprints of geoblocking that we extracted from our blockpages. We hope that our work highlights the impact and human cost of such geoblocking and fosters efforts from the community to hold server-side operators accountable to issues of over-compliance. We need to advocate for more standardized, informative user-facing blockpages, and justifications for why web services that do not fit the definition of the embargo still implement geoblocking.

2 BACKGROUND AND RELATED WORK

The U.S. Embargo and Its Complexity: The Cuban embargo was initially proposed by President Kennedy in 1962 and expanded by Presidents Johnson and Reagan, finally being formalized with the Helms-Burton Act of 1996 [9, 14]. The goal of the embargo was to pressure Cuba to return to a democratic form of government by cutting off imports from and exports to the country [25]. In recent times, these regulations have seen significant change [5, 9, 25].

Businesses looking to provide products or services in Cuba must either apply for and receive specific authorization from the U.S. Department of Treasury or be a member of a list of allowed products and service categories promulgated by the Bureau of Industry and Security (BIS) [19]. These two bodies work together with the Office of Foreign Assets Control (OFAC) to ensure compliance with these regulations; any businesses found in violation can be fined up to U.S. $350,000 per transaction conducted depending on the scale and severity of the transgressions [19]. These restrictive regulations and the complexities of seeking legal exemptions have led to companies not operating in Cuba.

Local Regulatory and Connectivity Barriers: Despite being one of the first countries in the Caribbean region to gain access to the Internet in the late 1990s, Cuba's connectivity and infrastructure development rapidly stalled due to the communist government's passage of laws aimed at limiting the free flow of information [10]. To further regulate internet access, Cuba merged its existing telecommunications providers to make the only legal state ISP, Empresa de Telecomunicaciones de Cuba S.A. (ETECSA) [15]. The popularity of the internet grew throughout the 2010s, leading to ETECSA offering 3G/4G access plans by late 2019 [13, 20]. Though, access remains both slow and expensive [11, 26, 27].

Bischof *et al.* conducted a network analysis to explore the state of Cuban internet in 2015, finding that most traffic traveled through ALBA-1 but was rather slow, with pings sometimes exceeding 300 ms, which they attributed this to incoming traffic passing through high-latency satellites [6]. They also ran traceroute measurements to identify the underlying network structure and Autonomous System (AS) links in Cuba. Pujol *et al.* worked with network operators to explore the SNET (which was subsequently absorbed by the state), mapping its topology [29]. While the network was unreliable at times and had security concerns, they found it was widely used especially for gaming and social networking.

Geoblocking Studies: Tschantz *et al.* [36] were among the first to introduce analysis of geoblocking, retrieving domains hosted by Cloudflare present in the Alexa Top 1M websites list from vantage points in various countries. McDonald *et al.* [22] conducted a wide-scale set of measurements focused on CDN geoblocking, querying the Alexa Top 10K from thousands of vantage points in 177 countries. They found that Cuba experienced geoblocking on 66 out of 10K (0.66%) and 165 out of the Alexa Top 1M (0.0165%) websites, lagging only behind Syria, Iran, and Sudan in both categories. Crucially, they were able to validate their data by comparing it to Cloudflare data to ensure their results were indeed accurately identified. Afroz *et al.* [2] verified these findings, using traceroute measurements and response comparisons to distinguish between geoblocking and censorship from vantage points in Africa, Pakistan, and Ukraine. While our measurements are inspired by these global studies, our work is focused on the effect of U.S. embargo and geoblocking on Cuban users and the human cost of such geoblocking.

3 UNDERSTANDING THE HUMAN COST OF GEOBLOCKING

To empirically characterize the current impact of geoblocking and embargo on Internet users in Cuba, besides consulting with Internet routing infrastructure experts, we conducted a qualitative study by reaching out to 10 tech-savvy Cuban citizens leveraging our personal and NGO contacts. Through our qualitative study, first, we collect demographic and Internet usage information, and poll participants on notable web services and categories of websites that are blocked and the resulting impact. Next, we ask the participants to submit indicators of such blocking, including uploading blockpage screenshots. Further, we aim to capture the impact of geoblocking by asking participants to explain how the inaccessibility of various categories of web services affects their daily life. Responses were either collected in English or in Spanish and translated to English. Our participants all reside in Cuba, and the majority are within the ages of 26-35, and range from daily Internet users to software professionals; five of the responses came from software developers or network administrators, one from a government official, and the remaining participants had different occupations.

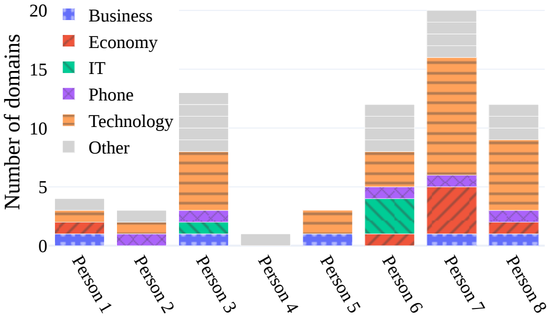


Figure 1: User-indicated categories of websites that impose geoblocking towards Internet users in Cuba.

From our qualitative study, we collected 44 user-reported blocked websites. Figure 1 shows the categories that participants most commonly indicate as being blocked. Most participants mentioned websites belonging to the categories of Technology (*e.g.* `Zoom`) and Economy & Finance (*e.g.* `Paypal`). The majority of these blocked websites match an existing, curated blocklist maintained on Github [28]. Notably, all participants mention that they have experienced a `403 Forbidden` message when trying to access an Internet service, and nine of ten participants say they have encountered a "not available for users in your region" message. Although we sought examples of geoblocking rather than ISP or government censorship, participants still listed domains *e.g.* `www.dimwcuba.com` and categories like political criticism and human-rights issues, which are more likely censored. *This underscores the complexity of distinguishing and attribution of geoblocking vs. local censorship.*

We identify the most discord between participants in how they perceive the effect of the inaccessibility of particular services on their day-to-day lives. Figure 2 shows that participants have a variety of contradictory responses throughout different aspects of their lives such as career opportunities, financial well being and education. This further highlights the importance of our work and the need to quantify what is being blocked. Considering the technical background of the users we study, such confusion and ambiguity may be *even more prevalent among regular Internet users in Cuba.* Finally, nine out of ten participants indicate that they rely on VPNs to circumvent content restrictions, with several also mentioning utilizing proxy services.

4 MEASUREMENTS

We obtain a vantage point in the largest ISP, ETECSA which is also state-owned, through our contact in Cuba. We also rented three vantage points in datacenters, two in the U.S. and one in the U.K. to serve as controls. There were no differences found across these control vantage points.

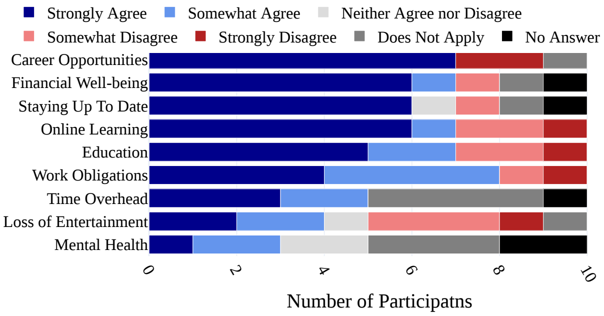


Figure 2: User-indicated aspects of day-to-day life most impacted by inaccessible services.

For our measurement test lists, we use 43 websites specifically mentioned by participants in our qualitative study (§3). To this list, we add the set of 126 domains reported by Cuban developers as being blocked in the country [28]. We also use the top 10,000 popular domains from the Tranco list [18], and the next 500 domains from the Tranco list [18] for each of the five categories that were most frequently reported to be blocked by our participants in the qualitative study: *Technology, Economy & Finance, Internet Phone & VOIP, Business, and Information Technology*. The union of these sub-lists has 12,384 domains, as shown in Table 1. We find the category of domains using Cloudflare's domain categorization [34].

Our measurement were conducted over the period of May 11–22 2023, with corresponding control and test data measurements conducted within 24 hours of each other. We send up to 100 measurements in parallel, waiting one second between successive requests to the same IP address. We also retry measurements that timed out with a 30 second gap to confirm consistent behavior. Geoblocking is often based on a user's IP address, and its AS and geolocation. As shown by previous work [22, 33, 36], geoblocking implementation varies across web services, hence we perform rigorous measurement on multiple-layers combined with TCP and TLS traceroute-like probes to identify and verify geoblocking responses and differentiate them from in-path censorship [33].

DNS Name Server Geoblocking: We perform 3 types of DNS measurements to test presence of geoblocking during DNS resolution. To capture common user behaviour, we perform a DNS request for the test domain using the local ISP DNS resolver. Also, considering the popularity of public DNS resolvers, we perform a DNS request through Cloudflare DNS (1.1.1.1). Cloudflare's nearest point of presence to Cuba is in the United States (Miami, Florida), and Cloudflare does not forward Client IP information by default [7]. Finally, we perform the recursive DNS queries ourselves, acting as a recursive DNS resolver. This is not only resistant to ISP DNS censorship, but also allows us to extract the exact name server that presented an error.

Error Stage	Error Type	# domains
Test List Size		12,384
Passed Control DNS		10,740
DNS Fails (50/10,740, 0.46%)	Failed iteration	47 (0.44%)
	Failed AuthNS connect	3 (0.03%)
Passed Control TCP		8,051
TCP Fails (125/8,051, 1.55%)	Timeout	123 (1.53%)
	Network Unreachable	1 (0.01%)
	No route to host	1 (0.01%)
TLS Fails (27/8,051, 0.33%)	Timeout	19 (0.23%)
	Reset	8 (0.09%)
HTTP Fails (34 of 8,051, (0.42%)	Timeout*	11 (0.14%)
	Reset*	4 (0.05%)
	Truncated Response	8 (0.09%)
	DNS fail in redirect*	13 (0.16%)
HTTP Geoblocking Responses (504 of 8,051, 6.25%%)	403 Forbidden	489 (6.07%)
	404 Not Found (Legal Reasons)	2 (0.02%)
	451 Unavailable due to Legal Reasons	13 (0.16%)

Table 1: Results–We find explicit geoblocking signals for 717 domains. Most present a HTTP response or fail to establish a TCP connection, some domains implement multiple methods, denoted by asterisks.

TCP Geoblocking: Next, we attempt to establish a TCP connection with the web server, using the IP address obtained during the DNS resolution. In case the DNS resolution was unsuccessful, we use the IP address obtained from our U.S. control measurements. In TCP geoblocking, web servers could silently drop the TCP SYN packet from our Cuba vantage point forcing a timeout [33]. We retry failed TCP handshakes up to three times to account for transient network failures. Considering the same errors could also occur in the presence of in-path filtering middleboxes, we run TCP SYN traceroutes to identify the location of failure [31].

Application-layer Geoblocking: Once the TCP connection is established, we attempt to perform a TLS handshake, sending a TLS Client Hello with the Server Name Indication (SNI) field set to the test domain, replicating browser behavior. Since the SNI field is a common target for Internet censorship, we conduct TLS traceroutes to identify whether messages exchanged during the TLS handshake period reach the web server [31, 33]. Over the established TLS connection, we send a HTTP GET request with the Hostname set to the test domain. We store the response from the web server, and follow up to 3 redirects. The HTTP response either presents the valid webpage of the tested domain or an error code or a page indicating geoblocking. We identify geoblocking by manual fingerprinting (as shown in §5.2).

5 MEASUREMENT RESULTS

Table 1 shows a summary of our results. Of the 12,384 domains in our test list, we find that 10,740 domains succeeded in the DNS stage and 8,051 succeeded in the TCP stage in our control measurements. We consider these as the baseline for our name server and web server geoblocking results. *Overall, we are able to detect geoblocking by 717 domains.*

5.1 DNS Name Server Geoblocking

We find that name servers of 50 domains (0.46%) perform geoblocking of Cuban requests (refer Table 1). Of these 50, 47 domains fail to obtain the authoritative name server of the domain during the trace process, while 3 obtain the name of the authoritative name server but fail to connect to it. Through our iterative trace, we locate the name server that failed to respond to our query. For instance, we find the name server `192.102.198.240` preventing access for 3 Intel-related domains (`01.org`, `intel.cn`, and `intel.com`).

18 of the 50 domains performing DNS geoblocking belong to the Economy & Finance category (*e.g.* `capitalone.com`, `edwardjones.com`) and 16 belong to the Technology category (*e.g.* `webex.com`). As observed by previous studies, we found instances of five U.S. state government domains implementing DNS geoblocking ((mt | texas | virginia | arkansas | delaware).gov) [33]. While the U.S.embargo restricts financial transactions, we find that even *services with free options such as* `01.org` *(Technology),* `monster.com` *(Job Search), and* `kugou.com` *(Music) geoblock Cuban users through DNS, showing the detrimental nature of sanctions on Internet access.*

5.2 Web Server Geoblocking

504 domains (6.25%) send HTTP geoblocking pages, while another 125 domains (1.55%) fail in the TCP handshake stage.

During TCP connections, the most common failure we observe is due to connections timing out because of web servers not responding to TCP SYN packets. We also observe timeout and reset errors during the establishment of TLS handshakes with 27 domains (0.33%). *These errors do not provide any useful information to the user about geoblocking.*

Are TCP and TLS errors due to local censorship or geoblocking? TCP and TLS failures could either indicate the presence of local censorship or geoblocking by web servers, and hence we use traceroute measurements to distinguish between the two. Of the 156 unique IPs (corresponding to 123 domains) that fail during the TCP connections, we were able to successfully run traceroutes for 154 IPs. From this analysis, we find that none of the traceroutes end within Cuba and 141 of the 154 traceroutes reach the same country as the target IP. This indicates that the blocking is applied very close to the web server and is most likely geoblocking. Clustering the AS of the last successful hop in these traceroutes, we see that 18 of them reach an Akamai AS (AS 32787), 16 reach Level3 (AS 3356), and 14 reach Hurricane Electric (AS 6939). We find similar results from our TLS traceroutes to the 27 domains that failed during the TLS handshake stage—all traceroutes exit Cuba and reach the country of the target

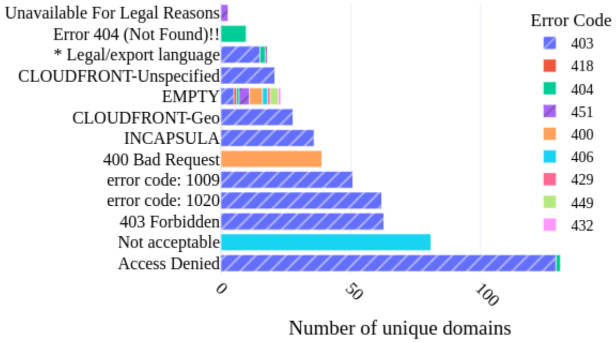


Figure 3: Parsed response pages grouped by fingerprints. Fingerprints representing the 403 and 451 status codes are explicit signals of geoblocking.

IP address. *Our traceroute experiments show failures occur close to the web server and are likely caused by geoblocking.*

Which HTTP responses indicate geoblocking and do web servers provide clear geoblocking signals? *Responses with status codes 403 and 451 indicate geoblocking but most web servers do not provide clear geoblocking messages.* 733 test domains always return a non-`200 OK` response from our Cuban vantage point. From these, we isolated the 677 responses with status codes in the `4XXs`, which indicates a server-side error. To better characterize the HTTP response bodies with `4XX` codes and examine whether they signal geoblocking, we build HTML fingerprints through manual inspection. We use an HTML parser to extract the text, and filter out content referencing JavaScript and enabling cookies. We then iteratively identify 14 matching fingerprints, the most popular of which are shown in Figure 3. Manually inspecting these fingerprints, we find that pages with a `403` or `451` response are most likely geoblocking pages, and are returned by 502 domains. We also find two domains returning a `404` status code that showed an explicit indicator of geoblocking. We conservatively do not count the responses with other status codes as geoblocking, but even with such conservative estimates, geoblocking through HTTP pages still accounts for 70.3% (504/717) of all of the geoblocking we find. We will open-source these fingerprints to enable further research on geoblocking. Notably, only 49 geoblocking domains (9.78 %) return blockpages indicating any sort of legal reasoning for the blocking.

What type of domains perform geoblocking? *Websites belonging to the Technology, Economy & Finance, Business, and Ecommerce category perform the most TCP, TLS, and HTTP blocking.* Figure 4 shows the number of domains performing different types of web server geoblocking in the top 10 categories with geoblocking. Economy & Finance, Business, and Ecommerce domains that geoblock Cuba deal primarily with financial transactions (*e.g.* `paypal.com`, `target.com`), which comes under the direct regulation of the U.S. embargo. However, many of these domains simply timeout during TCP connections, providing users with no information about why they are blocked. Moreover, Technology, Information Technology, and Education domains (*e.g.* `zoom.us`, `udemy.com`, `www.gitkraken.com`) provide free or partly free services that do not require financial transactions, but still restrict access to Cuban users. These are also the aspects that our qualitative study participants indicate as significantly impacting their everyday life, as shown in Figure 2. *The large number of domains performing TCP and HTTP geoblocking sheds light on the severe impact of the U.S. embargo on the unavailability of different types of web services.*

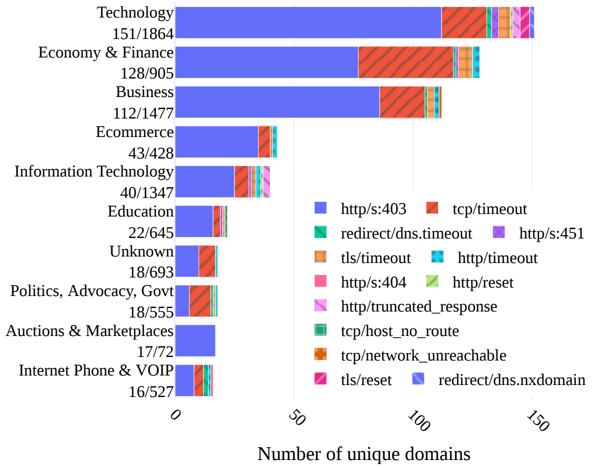


Figure 4: Web server Geoblocking in Top 10 Categories– Domains in the Technology and Economy & Finance categories perform the most geoblocking.

5.3 Website Popularity

More popular domains implement significant levels of geoblocking. Figure 5 shows a histogram of geoblocking of websites in bins of decreasing Tranco Rank for the Top 10K domains. The highest amounts of geoblocking are done by the domains in the Top 1K, and there is a negative trend of geoblocking with Tranco rank. We observe more popular domains in the Technology (`zoom.us,#41`), Information Technology (`microsoft.com,#6`) and Education (`scribd.com,#370`) categories performing geoblocking.

Do popular websites mention U.S.embargo sanctions? *Most US-based websites provide some description of the effect of U.S. embargo sanctions on their service.* We manually investigate the privacy policies and terms of service documents for all geoblocking domains within the Top 2K Tranco rank (142 domains) to identify whether these services publicly mention embargo sanctions and why they geoblock Cuban users. 103 out of the 134 services are headquartered in the

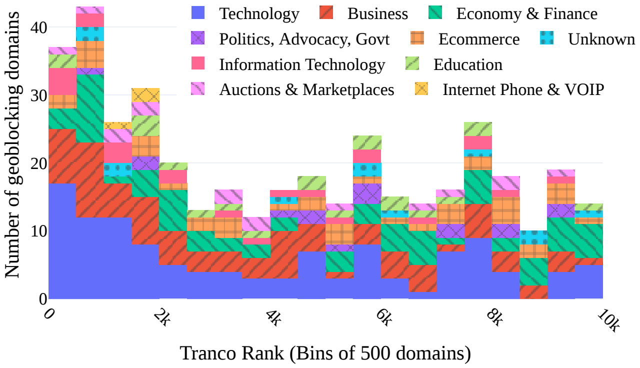


Figure 5: Geoblocking by Tranco Rank and Category– Domains within the first thousand Tranco rank perform more geoblocking, with many popular Technology and Business domains geoblocking Cuban users.

United States. 48 of the 142 services (33.8%) do not make any mention of sanctions or embargo in their terms of service, another 64 (45.07%) only make a short mention about sanctions such as "You and your Users will not use the Services from an embargoed country (currently Cuba, Iran, North Korea, Sudan and Syria)," and the remaining 30 (21.13%) provide detailed descriptions about the impact of U.S.embargo sanctions on their services. *Our investigation shows that web services commonly indicate the U.S. embargo as the reason behind the unavailability of their services in Cuba.*

6 DISCUSSION AND CONCLUSION

We record the extent and variety of methods by which global service providers implement embargo-based geoblocking in Cuba. Our analysis shows the damaging nature of ambiguous blocking and over-compliance.

Do we find all cases of geoblocking? *No!* We provide a conservative estimate of geoblocking in this paper, since we only report clear signs of consistent failures that signal geoblocking. It is possible for error codes other than 403 and 451 to be used for geoblocking. Anecdotally, we found instances of Cuban internet users attributing certain 404, 403, and 406 responses to geoblocking [3, 12, 21]. However, we do not include these response codes in our counts as they do not clearly signal geoblocking. Moreover, it is possible that certain web services only limit certain functionality for Cuban users, such as denying payment or refusing downloads. Future work can utilize dynamic analysis of websites to discover such types of discrimination.

How can web services do better? *Develop and provide standardized indication of sanction compliance!* Our results highlight the difficulty for users and researchers to distinguish between local censorship and geoblocking due to sanctions. This confusion is exacerbated by the fractured, non-standard, and at times incomprehensible ways in which services apply blocking. In addition to the ambiguity in the error codes discussed above, there is no clear indication for the user in the cases where the blocking is implemented as a DNS, TCP, or TLS error. Our qualitative study further illustrates even tech-savvy internet users are confused about whether particular web services are blocked due to the local government or due to the U.S. embargo, suggesting such confusion may be even more prevalent among average Internet users in Cuba and is corroborated by the opaque nature of blocking shown in our measurement findings. We advocate for service providers to adopt and use standardized client error response codes. For instance, content providers such as Cloudflare have made efforts to tailor their platform-specific errors to indicate deliberate blocking of certain users [8, 22].

Additionally, even within service providers, ambiguity often remains in spite of ease of indication. A stark example of this is the two identified CloudFront fingerprints, CLOUDFRONT-Geo and CLOUDFRONT-Unspecified, varying only by the inclusion of: "The Amazon CloudFront distribution is configured to block access from your country." in the Geo case and "Request blocked" in the unspecified case.

The lack of accountability surrounding over-compliance enables companies to operate fully at their own discretion, resulting in disparity within sectors that only serves to breed more end-user confusion. For example, Microsoft Azure and IBM Softlayer are two of the leading cloud service providers in the world. While Softlayer geoblocks all traffic from restricted nations including in our Cuba measurements [16], Azure passes the responsibility of ensuring the compliance of their services with regulations to their users [23]. Additionally, even when lifted by the U.S. government, removing internet sanctions can be non-trivial. We found five blockpages citing embargo that explicitly mention Iran as a sanctioned country, despite the issuance of Iran General Licence D-2 in late 2022, which authorized tech companies the capacity to provide a greater range of services to the people of Iran [37].

This culminates in an inequitable state of Internet accessibility extending far beyond the Cuba embargo and its ramifications, where the most harmed by over-compliance are the citizens of sanctioned countries themselves. Projects like sanctions.net [35] exist to aid in legal compliance for sanctions worldwide, but web service providers still must take care to accurately comply and especially to engage in corresponding disclosure to end-users. By taking a data-driven approach and quantifying the impact of the Cuban embargo and geoblocking, we seek to encourage future work and community attention to geoblocking. Only then can the overarching obscurity of Internet sanctions and associated negative human consequences be meaningfully reduced.

REFERENCES

[1] Access Now. Letter to U.S. government: Do not disrupt internet access in Russia or Belarus, Mar. 2022. https://www.accessnow.org/letter-us-government-internet-access-russia-belarus-ukraine/.

[2] Afroz, S., Tschantz, M. C., Sajid, S., Qazi, S. A., Javed, M., and Paxson, V. Exploring server-side blocking of regions. *arXiv preprint arXiv:1805.11606* (2018).

[3] Akhtar), M. M., and natalee9 (Andres Gonzalez). How i can solve 406 not acceptable webflow issue, 2022.

[4] Bailey, M., Dittrich, D., Kenneally, E., and Maughan, D. The menlo report. *IEEE Security & Privacy* (2012).

[5] Bellinger, J. B., Shannon, T. A., Barker, J. P., Weiss, B., and Mirski, S. A. Two years of title iii: Helms-burton lawsuits continue to face legal obstacles: Advisories, 2021.

[6] Bischof, Z. S., Rula, J. P., and Bustamante, F. E. In and out of cuba: Characterizing cuba's connectivity. In *Proceedings of the 2015 Internet Measurement Conference* (2015), Association for Computing Machinery.

[7] Cloudflare. Does 1.1.1.1 send edns client subnet header? https://developers.cloudflare.com/1.1.1.1/faq/#does-1.1.1.1-send-edns-client-subnet-header.

[8] Cloudflare. Troubleshooting cloudflare 1xxx errors, 2023. https://developers.cloudflare.com/support/troubleshooting/cloudflare-errors/troubleshooting-cloudflare-1xxx-errors/.

[9] Cuban liberty and democratic solidarity (libertad) act of 1996 (h.r.927), 1996.

[10] del Poder Popular, A. N. Decreto no. 209/96, Sobre el acceso de la República de Cuba a Redes de Alcance Global, 1996.

[11] ETECSA. @etecsa_cuba photo captioned "navegación en internet por datos móviles [internet browsing by mobile data]", 2020.

[12] Fischer, P. @airbnbhelp is there a known issue that causes a „404" error when trying to book an accommodation in cuba?i get 404 instead of payment page, 2016.

[13] Cuba: Freedom on the net 2022 country report, 2022.

[14] Gerhard Peters, J. T. W. Proclamation 3447—embargo on all trade with cuba, 1962. https://www.presidency.ucsb.edu/documents/proclamation-3447-embargo-all-trade-with-cuba/.

[15] Harris, J. Castro hates the internet, so cubans created their own, 2015.

[16] Ibm cloud notices. https://cloud.ibm.com/docs/overview?topic=overview-notices, 2023.

[17] Jones, B., Ensafi, R., Feamster, N., Paxson, V., and Weaver, N. Ethical concerns for censorship measurement. In *NS Ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research* (2015). https://www.icir.org/vern/papers/censorship-meas.nsethics15.pdf.

[18] Le Pochat, V., Joosen, W., Korczyński, M., Tajalizadehkhoob, S., and Van Goetham, T. Tranco: A research-oriented top sites ranking hardened against manipulation, 2023.

[19] Electronic code of federal regulations (e-cfr), 2023.

[20] Marsh, S. Cuba launching internet on cellphones, 2018.

[21] mas, C. . u. g. Puedo poner más, porque no puedo ser usuario de paypal, de airtm, cuando quiero bajar alguna apk en google play me dice no disponible para mi país y muchísimas más web que cuando ven ip de cuba ponen error 403 o 404. https://twitter.com/62_cuba/status/1420569443420225543?s=46&amp;t=v8srG0HVvWCzq7GSRfbQMw, 2021.

[22] McDonald, A., Bernhard, M., Valenta, L., VanderSloot, B., Scott, W., Sullivan, N., Halderman, J. A., and Ensafi, R. 403 forbidden: A global view of cdn geoblocking. In *Proceedings of the Internet Measurement Conference 2018* (2018), Association for Computing Machinery.

[23] Azure support for export controls. https://learn.microsoft.com/en-us/azure/azure-government/documentation-government-overview-itar#ofac-sanctions-laws, 2022.

[24] Narayanan, A., and Zevenbergen, B. No encore for Encore? Ethical questions for web-based censorship measurement, 2015. Available at SSRN: https://ssrn.com/abstract=2665148 or http://dx.doi.org/10.2139/ssrn.2665148.

[25] Oliver, I., and Venancio, M. N. Understanding the failure of the u.s. embargo on cuba, 2022.

[26] Salario medio en cifras. cuba enero-diciembre 2021 [average salary in figures. cuba january-december 2021], 2021.

[27] Speedtest global index, 2023.

[28] Opensourcers, C. Cuban-opensourcers/cuban-restricted: Awesome list about tech sites/services restricted for cuba, 2023.

[29] P. Pujol, E. E., Scott, W., Wustrow, E., and Halderman, J. A. Initial measurements of the cuban street network. In *Proceedings of the 2017 Internet Measurement Conference* (2017), Association for Computing Machinery.

[30] Partridge, C., and Allman, M. Addressing ethical considerations in network measurement papers. In *NS Ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research* (2015).

[31] Raman, R. S., Wang, M., Dalek, J., Mayer, J., and Ensafi, R. Network measurement methods for locating and examining censorship devices. In *Proceedings of the 18th International Conference on Emerging Networking EXperiments and Technologies* (2022), Association for Computing Machinery.

[32] Ramesh, R., Raman, R. S., Bernhard, M., Ongkowijaya, V., Evdokimov, L., Edmundson, A., Sprecher, S., Ikram, M., and Ensafi, R. Decentralized Control: A Case Study of Russia. In *Network and Distributed System Security* (2020), The Internet Society.

[33] Ramesh, R., Sundara Raman, R., Virkud, A., Dirksen, A., Huremagic, A., Fifield, D., Rodenburg, D., Hynes, R., Madory, D., and Ensafi, R. Network responses to russia's invasion of ukraine in 2022: A cautionary tale for internet freedom. In *32nd USENIX Security Symposium (USENIX Security 23)* (2023), USENIX Association.

[34] Ruth, K., Kumar, D., Wang, B., Valenta, L., and Durumeric, Z. Toppling top lists: Evaluating the accuracy of popular website lists. In *Proceedings of the 22nd ACM Internet Measurement Conference* (2022), Association for Computing Machinery.

[35] The Internet Sanctions Project, 2023. https://wiki.sanctions.net/.

[36] Tschantz, M. C., Afroz, S., Sajid, S., Qazi, S. A., Javed, M., and Paxson, V. A bestiary of blocking: The motivations and modes behind website unavailability. In *8th USENIX Workshop on Free and Open Communications on the Internet (FOCI 18)* (2018), USENIX Association.

[37] U.s. treasury issues iran general license d-2 to increase support for internet freedom, 2022. https://home.treasury.gov/news/press-releases/jy0974.

[38] Xynou, M., and Filastò, A. New blocks emerge in Russia amid war in Ukraine: An OONI network measurement analysis, Mar. 2022. https://ooni.org/post/2022-russia-blocks-amid-ru-ua-conflict/.

[39] Yeung, C., Iqbal, U., O'Neil, Y. T., Kohno, T., and Roesner, F. Online advertising in ukraine and russia during the 2022 russian invasion. In *Proceedings of the ACM Web Conference 2023* (2023), pp. 2787–2796.

[40] Zevenbergen, B., Mittelstadt, B., Véliz, C., Detweiler, C., Cath, C., Savulescu, J., and Whittaker, M. Philosophy meets Internet engineering: Ethics in networked systems research. GTC Workshop Outcomes Paper, 2015. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2666934.

A ETHICS

For both our qualitative study and our measurements, we consulted with and received approval from the Institutional

Review Board (IRB). We also ensure that our qualitative study participants, Cuban individuals, have a clear understanding of and provide explicit consent for the study. Additionally, during the course of the study, we collect no personally identifiable information (PIIs) in line with best practices. Moreover, we assess the potential risks associated with our measurements through consultations with our collaborators with decades of experience in the field of censorship measurement. A longtime Internet freedom community member with many years of experience especially related to network traffic analysis of Cuba explicitly consented to provide the Cuban vantage point to facilitate the study. We also modelled our measurements to imitate browser behavior and rate limit connections to the same IP address. We perform measurements in small batches with the purpose of not overloading the Cuban vantage point or affecting intermediaries and servers, in line with measurement ethics and previous works [4, 17, 24, 30, 32, 40].

B MEASUREMENT TEST LIST

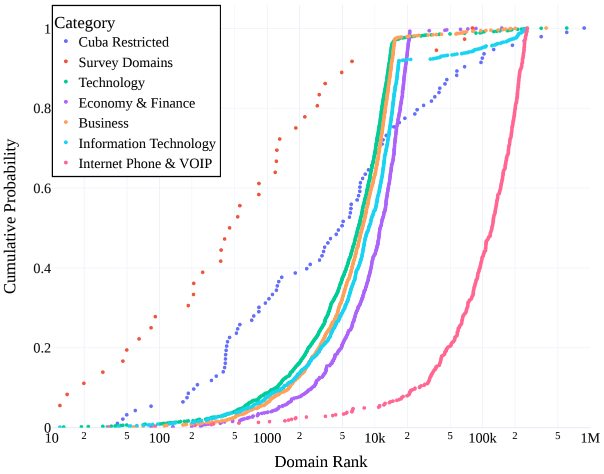


Figure 6: CDF of our Test List vs. Tranco Domain Rank

As mentioned in §4, we created our measurement test list from the community-curated list [28], our qualitative user study (§3), and the Tranco Top 10K. In Figure 6, we illustrate the ranking of domains in the community-curated list, qualitative user study, and the five categories most commonly mentioned in our qualitative survey. We note that domains could be classified under multiple categories, causing the plateaus after rank 10K.

C GEOBLOCKING TYPES IN TRANCO TOP 10K

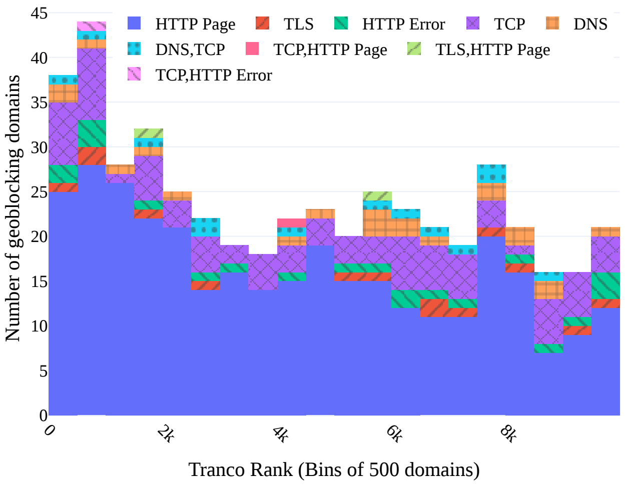
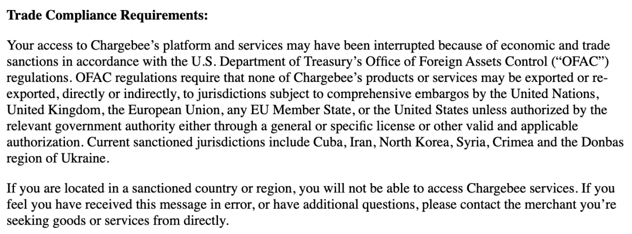


Figure 7: Geoblocking by Tranco Rank

As shown in Figure 7, we find that domains within the first thousand Tranco rank ostensibly perform more geoblocking. As illustrated in the figure, here are cases where webservers for some domains perform multiple types of geoblocking, such as blocking on both the DNS and TCP levels.

D OUTDATED BLOCKPAGE



Trade Compliance Requirements:

Your access to Chargebee's platform and services may have been interrupted because of economic and trade sanctions in accordance with the U.S. Department of Treasury's Office of Foreign Assets Control ("OFAC") regulations. OFAC regulations require that none of Chargebee's products or services may be exported or re-exported, directly or indirectly, to jurisdictions subject to comprehensive embargos by the United Nations, United Kingdom, the European Union, any EU Member State, or the United States unless authorized by the relevant government authority either through a general or specific license or other valid and applicable authorization. Current sanctioned jurisdictions include Cuba, Iran, North Korea, Syria, Crimea and the Donbas region of Ukraine.

If you are located in a sanctioned country or region, you will not be able to access Chargebee services. If you feel you have received this message in error, or have additional questions, please contact the merchant you're seeking goods or services from directly.

Figure 8: Blockpage referencing outdated Iran sanctions.

As shown in Figure 8, we observe that some geoblocking pages still incorrectly cite Iran as being subject to OFAC sanctions governing technology products despite the issuance of Iran General Licence D-2 in late 2022 [37].