Distributed Detection of Tor Directory Authorities Censorship in Mexico

Gunnar Eyal Wolf Iszaevich
Instituto de Investigaciones Económicas, UNAM
Facultad de Ingeniería, UNAM
Mexico City, Mexico
Email: gwolf@gwolf.org

***Abstract*—The Tor network relies on individuals to set up relays for it to operate. Campaigns have in the past been successfully made to invite more people to join, and the network currently consists of close to 6,500 relays, spread globally. Although the Latin American region has many characteristics that make it natural to expect a wide participation in Tor, it has lagged behind most of the world in its Tor activity — Both considering client usage and participation as relays. This study focuses on the difficulties the Mexican user community has faced in setting up Tor relays, and presents how –and why– we deployed a relatively very simple and unsophisticated network censorship reporting system, as well as the results we have received so far. While this is still considered a work in progress, it has yielded important results as an aide allowing to specify the needed characteristics for potential relays, with a clear, measurable result.**

Keywords: *ISP*; *Tor*; *Censorship*; *Detection*; *Mexico*.

I. Introduction

Anonymity loves company, says the adage. In our networked world, this means that the technical excellence of an anonymity technology is often second in importance to its usability [3], as a great program with very low usability will keep its mass adoption low — and if few people use it, de-anonymizing one of its users becomes easier. Hence, a fundamental concern for any anonymity-providing network is how to get more people to adopt it.

The best known, best studied and most popular anonymization technology is Tor [15]. It provides a low latency network, overlaid over the regular Internet, based on *onion routing* [14]. Tor is a network that relies on volunteers to provide the servers (*relays*) and their respective bandwidth for its operation.

One of the clearest ways people can help the Tor project is by running new relays; several campaigns and proposals have been launched by individuals and organizations asking committed users to set up new relays [6][10][13].

While the campaigns have been successful on a global scale, some regions' participation in the network remains quite low. In April 2017, the Tor Project started its *Global South* working group [9] to increase awareness and participation in the project for users in countries in said regions, be it as users or as participants in the network.

As the time of this writing, the Tor network consists of slightly over 6,300 relay nodes as reported by the Tor Metrics site [12]. As can be seen in Fig. 1, while there was a constrant growth in the number of relays until 2015, the number has since remained fairly stable. Tor Metrics also reports the number of daily users of the network to be close to two million; the Latin American region represents only a tiny percentage, with a combined weight of only 1.53% of the network's users [11].

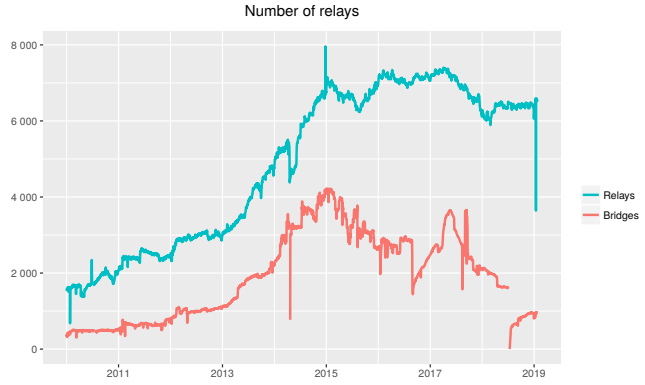

Figure 1. Number of relays and bridges over time, 2010–2019. Source: Tor Metrics [12]

The number of relays is not a core concern for the Tor network, nor is –as can be understood from Fig. 2– its available bandwidth; even though the number of relays available in the last years has not changed, advertised bandwidth has kept an overall increasing trend, and more importantly, consumed bandwidth is kept close to half of it.

However, Tor relatively lacks *diversity*, a fundamental and most desired property to be able to withstand deanonymization attacks by nation-state adversaries [8]. One of the goals of the aforementioned *Global South* group is to promote the installation of Tor relays worldwide.

Tor usage throughout the world is superbly depicted in Graham's 2014 visualization [5]; it shows that in 2014 Mexico had a similar amount of users as Sweden or Austria, a countries with a tenth of Mexico's population — and with a much better record on human rights and freedom of the press. This trend continues, as Tor Metrics reports all said countries in the 10,000 to 15,000 daily users range.

The number of relays ran from each of the aforementioned countries, however, is dozens of times larger than in Mexico. The sum of the factors so far mentioned led us to pursue convincing other sympathizers to set up Tor relays.

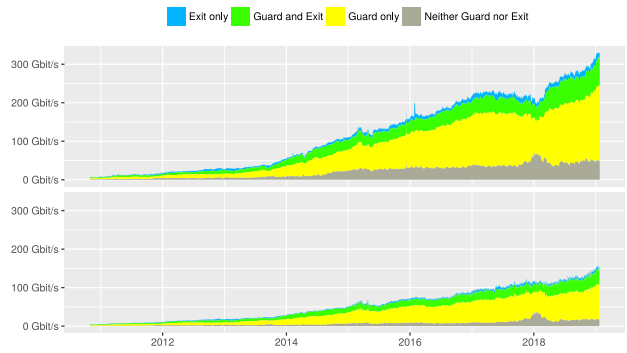

Figure 2. Advertised (above) and consumed (below) bandwidth in the Tor network, 2010–2019. Source: Tor Metrics [12]

While large numbers of relays have never been observed in Mexico, Fig. 3 shows a clear symptom of network censorship: while there was only one stable relay before 2013, in the lapse of a year the number grew (partly due to the aforementioned campaigns) to stabilize between seven and eight relays. However, in late 2015 there is a sharp drop, and while there are some spikes, Mexico's presence was clearly limited.

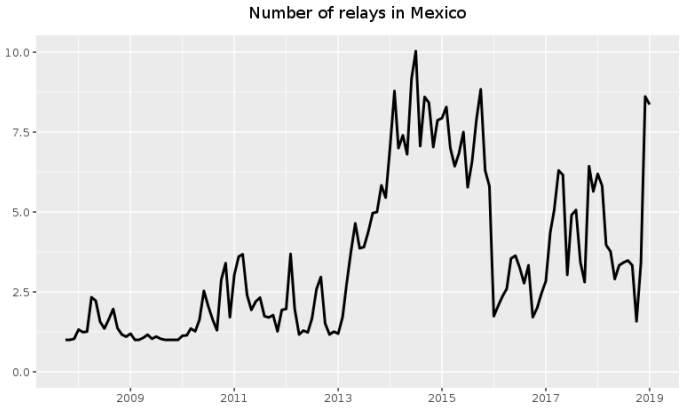

Figure 3. Number of relays in Mexico, 2008–2019. Source: Tor Metrics [12]

Besides anecdotal evidence by several former relay operators, we have found online reports from around the time this censorship was instated [1].

During 2017 and 2018, our project (UNAM/DGAPA/PAPIME PE102718) in consonance with Derechos Digitales' (see Section VI) engaged in promoting further relays, but it wasn't until after we had partial results of the project this article presents that the spike at the end of the graph.

The results we can report so far are, sadly, not that we managed to stop the censorship — but by finding an ISP (Internet Service Provider) amenable to running relay nodes, managed to successfully improve participation. However, with this information documented, we are starting to contact relevant ISPs in order to work legally and socially against their blocking.

II. Network censorship, architecture or policy?

When attempting to set up Tor relays in residential (Digital Subscriber Line, or DSL) connections in Mexico, we found they repeatedly failed to be recognized by the Metrics site. Although we did have some anecdotal evidence pointing towards the ISP blocking connectivity to the Tor directory authorities (DirAuths) [1], we needed further validation to ensure whether this was effectively due to network censorship (and not misconfiguration).

Also, as we were embarking on a project to distribute Raspberry Pi computers donated by the *Derechos Digitales* NGO for volunteers interested in setting up a relay, we felt necessary to do a more thorough review to check the status of the different providers.

We identified the three following points as in need of an answer:

1) Does the ISP *actively interfere* with connections?
   We need to know if there are technical measures *purposefully* set up by the ISP to block connections to Tor.
2) Does the ISP perform *deep NAT* (Network Address Translation) to its customer's networks?
   Due to the scarcity of IPv4 addresses, many ISPs (specially the local ones, or the latecomers to the market) don't provide a network-visible IP address to each user. Instead, several layers of NAT can be traversed on the way to the real network (we have detected up to seven *hops* inside NATted networks). If users cannot be reached from the outside network, there is no way they can set up relays.
3) Does the ISP allow end users to reconfigure their routers and receive incoming connections?
   Even having the necessary network capabilities to reach the user's connection, and allowing unfettered access to the Tor DirAuths, residential-grade routers are usually configured –in good measure for security– to reject any connections not started within the client's network.
   All modern routers have the capability to set up network forwarding for specific ports. Not all ISPs, though, allow the user to configure in this fashion their routers.

Given that item number 3 needs actually reconfiguring network equipment, we decided not to pursue it at this stage.

From the three points mentioned above, although they are all important for the project's goals, only item 1. qualifies as network censorship.

III. Interface description

As we needed to survey different networks countrywide, we decided to make a public call for participation: asking individuals to run some tests for us. We needed to design a simple task which any interested person could easily follow.

We considered adopting preexisting tools for this task, mainly OONI (Open Observatory of Network Interference, see Section IV for further details). However, given that our interest was specific and limited to getting information as to give to

potential relay operators (which ISPs would be feasible to set up relays with), we did not consider necessary to design a full application; we decided to request information based on tools readily available in default installations of any general-purpose operating system.

With this in mind, we set up a simple form, reproduced in Fig. 4, collecting only some data about the connection of each probe, and giving instructions to run *traceroute*, either on Unix or Windows-based systems. We request our users to provide the results of running *traceroute* to all of the Tor DirAuths. *Traceroute*, being an ICMP probing tool, has many known shortcomings and many readily available tools would probably do a better job. The criteria for choosing *traceroute* is, again, that it is available and preinstalled in every major operating system.

Figure 4. Interface at http://rutas.priv-anon.unam.mx with the form shown to users when submitting a trace

We acknowledge the main blocker for this form is the means we requested participants to submit their information from: They have to open an interactive terminal, paste into it a long command, wait for a couple of minutes (we have observed run times between one and two minutes from non-censored networks, and between four and six minutes from censored ones) for it to finish, and paste back their results in the browser. We reproduce here the command for Unix systems:

```
for i in 171.25.193.9 86.59.21.38
199.58.81.140 194.109.206.212
204.13.164.118 131.188.40.189
128.31.0.34 193.23.244.244
154.35.175.225 128.31.0.39
199.254.238.52; do traceroute $i;
done
```

It is far from user friendly. This design was chosen due to the limited time and resources we had.

IV. Related work

There are many projects with different scopes aimed at detecting network censorship in the context of Tor participation. Even with this stated level of specificity, this section is far from comprehensive.

OONI [4] is a global project aimed at finding and reporting several instances of network censorship worldwide. OONI operates in a fashion comparable to what Tor does, based on a large amount of *probes* run continuously on hosts provided by volunteers, performing network connections and looking for censorship or filtering evidence in many ways, including tests for Tor connectivity. OONI also has user-friendly applications that can be installed in mobile devices. A major output of OONI's work is the interpretation of the gathered data in a global fashion, often correlating censorship events with news items.

The OONI application, however, includes the probes only for web connectivity, instant messaging, network performance, and middleboxes detection. But even in the server-based probe, Tor connectivity is measured by trying to connect as a client to network. Our tests verify the reachability of the DirAuths, needed for setting up relays, but not for client connections.

Quite probably, we will be able to work with the OONI developers to add DirAuth reachability to their probes. This is a clear next step for our project.

The *traceroute.org* site, set up by Thomas Kernen [7], provides a directory of servers offering a Web form from which they run *traceroute* on behalf of the users. This site has sadly not been updated since 2011, and thus contains many broken links. While the linked sites do provide a valuable resource to network administrators, it does not provide any severs in Mexico, and is thus not suitable for our needs.

The model presented by Danezis [2] has many items compatible with what we try to achieve, but goes to greater lengths to assure a given address is blocked. It also presents a series of user connections to directory servers to detect censorship. While Danezis' model contemplates repeating measurements at time intervals of one week, given the nature of participation and our goal of not installing any software in the participating clients, ours is based on one-shot measurements. Besides, this work is presented as a model, not as a comparable implementation.

Another country-specific project worth mentioning is research on Internet censorship in China [16]. This works has a very different focus than our project's. China is probably, together with Iran the foremost country-level censorship example, and the researchers' approaches are applied in a much bigger scale. Of course, citing said article in no way means that Mexico's censorship is in any way comparable to China's.

The article starts by describing the *Great Firewall of China* at a BGP (Border Gateway Protocol) level, analyzing the conformation of the Autonomous Systems (AS), and explaining where they discovered the different filtering devices and presenting the filtering not as a *Great Firewall*, but as an

Internet Panopticon, with local and peripherial filtering points performing different tasks.

V. Results and discussion

Throughout five months, we received 79 reports from 12 states (out of 32 in the country). Table I shows the distribution of reported ISPs.

TABLE I. Number of reports received from each of the different available ISPs

ISP	Reports
Telmex	32
Axtel	1o
Izzi	7
Total Play	7
AT&T	6
Megacable	4
Alestra	2
UNAM	2
Avantel	1
Bestel	1
Cablevisión	1
Express VPN	1
Maxcom	1
Movistar	1
Nextel	1
Telcel	1

The distribution is close to what we expected, with Telmex (which spans its constituents, Uninet and Infinitum) clearly dominating the scene.

The results are aggregated and presented, one report per row, in a table as the one (partially) shown in Fig. 5; row colors represent the percentage of DirAuths each IP could reach: Red (0-25%), orange (25-50%), yellow (50–75%) and green (75-100%).

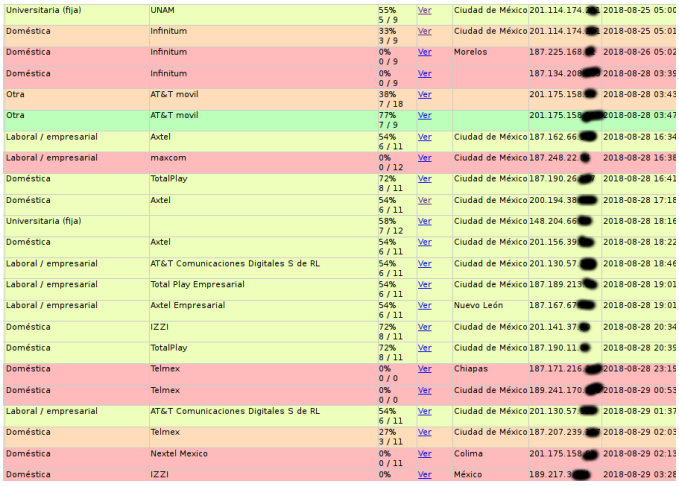

Figure 5. Results table. Last octet of all IP addresses has been manually obscured.

By the time this project was started, we knew for a fact that Telmex censored connections to DirAuths; this was confirmed, as most connections report 3 out of 11 successful connections. There are several records showing 0/11 — Given the similarities in them, we believe this to be caused by old modems not properly implementing NAT forwarding support for *traceroute*.

A second interesting finding was the high amount of connections providing sufficient but still not perfect returns — this means, connections where Tor relays could be installed, as they can exceed the 50% mark, but not by much — Most strikingly, the two tested connections at UNAM, Mexico's largest and most important university, can barely withstand being a relay, as they can reach only 55% of the DirAuths. This is another item to verify, both technically (what kind of communications exactly are being censored) and politically (why are they being censored).

Since we managed to systematize the results, we have been inviting prospective relay operators to connect via Axtel, the ISP that has the highest success rate. This has led to the spike at the right of Fig. 3.

VI. Further work

As for the reasons of the censorship, we have contacted Customer Support for the ISP with the largest market share, Telmex. As it was expected, they denied instrumenting this blocking. We have started contacting the Federal Institute for Telecommunications (IFT) so we can push for a real reply.

As it was said in the Abstract, this article presents a Work in Progress. We still have not analyzed the records to find evidence of *deep NAT*. ISPs, particularly smaller or newer ones, do not do this because of censorship, but because of their limited network resources; nevertheless, their connections are being censored.

Acknowledgement

The author wishes to acknowledge the UNAM/DGAPA/PAPIME PE102718 project for the needed facilities to carry out the activities here presented, as well as Derechos Digitales for its logistical and economical support.

Personal thanks to Vasilis Ververis, for preparing Fig. 3 which, in short, supports the writing of this article.

References

[1] F. Bustillos, *Is tor being blocked by isp? (mexico)*, 2016. [Online]. Available: https://lists.torproject.org/pipermail/tor-relays/2016-January/008491.html (visited on 03/15/2019).

[2] G. Danezis, "An anomaly-based censorship detection system for tor," *The Tor Project*, 2011. [Online]. Available: https://research.torproject.org/techreports/detector-2011-09-09.pdf.

[3] R. Dingledine and N. Mathewson, "Anonymity loves company: Usability and the network effect.," in *WEIS*, 2006. [Online]. Available: https://www.freehaven.net/anonbib/cache/usability:weis2006.pdf (visited on 03/15/2019).

[4] A. Filasto and J. Appelbaum, "Ooni: Open observatory of network interference.," in *FOCI*, 2012.

[5] M. Graham and S. D. Sabbata, "The anonymous internet," Internet Geographies, Oxford Internet Institute, Tech. Rep., 2014. [Online]. Available: http://geography.oii.ox.ac.uk/the-anonymous-internet/ (visited on 03/15/2019).

[6] R. Jansen, N. Hopper, and Y. Kim, "Recruiting new tor relays with braids," in *Proceedings of the 17th ACM conference on Computer and communications security*, ACM, 2010, pp. 319–328. [Online]. Available: http://www.robgjansen.com/publications/braids-ccs2010.pdf (visited on 03/15/2019).

[7] T. Kernen, *Traceroute.org*, 1998–2011. [Online]. Available: http://www.traceroute.org/ (visited on 03/15/2019).

[8] I. R. Learmonth, "Strength in numbers: Measuring diversity in the tor network," Tor Project, Tech. Rep., Dec. 11, 2018. [Online]. Available: https://blog.torproject.org/strength-numbers-measuring-diversity-tor-network (visited on 03/15/2019).

[9] A. Macrina, *Next steps from tor meeting*, Apr. 10, 2017. [Online]. Available: https://lists.torproject.org/pipermail/global-south/2017-April/000000.html (visited on 03/15/2019).

[10] D. McDevitt, "Tor exit relays to be run in libraries: Library freedom project," Open Technology Fund, Tech. Rep., Jul. 9, 2015. [Online]. Available: https://www.opentech.fund/news/tor-exit-relays-to-be-run-in-libraries-library-freedom-project/ (visited on 03/15/2019).

[11] J. Nájera, A. Argüelles, and S. Alcántar, "La internet anónima: Tor en méxico," Enjambre Digital, Tech. Rep., 2018. [Online]. Available: https://tor.enjambre.net/ (visited on 03/15/2019).

[12] T. Project, *About tor metrics*, 2009–2018. [Online]. Available: https://metrics.torproject.org/about.html (visited on 03/15/2019).

[13] R. Reitman, "Tor challenge inspires 1,635 tor relays," Electronic Frontier Foundation, Tech. Rep., Sep. 19, 2014. [Online]. Available: https://www.eff.org/deeplinks/2014/09/tor-challenge-inspires-1635-tor-relays (visited on 03/15/2019).

[14] P. Syverson, R. Dingledine, and N. Mathewson, "Tor: The secondgeneration onion router," in *Usenix Security*, 2004. [Online]. Available: https://www.onion-router.net/Publications/tor-design.pdf (visited on 03/15/2019).

[15] *The tor project*. [Online]. Available: https://www.torproject.org/ (visited on 01/25/2019).

[16] X. Xu, Z. M. Mao, and J. A. Halderman, "Internet censorship in china: Where does the filtering occur?" In *International Conference on Passive and Active Network Measurement*, Springer, 2011, pp. 133–142. [Online]. Available: https://censorbib.nymity.ch/pdf/Xu2011a.pdf (visited on 03/15/2019).