Global Network Interference Detection over the RIPE Atlas Network

Collin Anderson
University of Pennsylvania

Philipp Winter
Karlstad University

Roya
Independent Researcher

Abstract

Existing censorship measurement platforms frequently suffer from poor adoption, insufficient geographic coverage, and scalability problems. In order to outline an analytical framework and data collection needs for future ubiquitous measurements initiatives, we build on top of the existent and widely-deployed RIPE Atlas platform. In particular, we propose methods for monitoring the reachability of vital services through an algorithm that balances timeliness, diversity, and cost. We then use Atlas to investigate blocking events in Turkey and Russia. Our measurements identify under-examined forms of interference and provide evidence of cooperation between a well-known blogging platform and government authorities for purposes of blocking hosted content.

1 Introduction

An important counter strategy for the proliferation of Internet filtering mandates is to measure, document, and expose interference in the free flow of information. Sunlight is said to be a disinfectant and, by shedding light on these events, the public's attention can be drawn toward information controls. Several methods exist to detect and assess Internet filtering. Ideally, the analyst has direct control over a censored source host that can perform measurements against an external destination. This is typically not the case, however, so research often has to opportunistically resort to open proxies, the help of volunteers, and existing measurement platforms. All of these methods have advantages and disadvantages. Open proxies suffer from low network coverage, are unreliable or questionably reflect typical conditions, and are often limited to TCP streams or HTTP requests. Cooperation with volunteers exposes individuals to potential harm and is time-consuming. Current, specially-designed censorship measurement platforms suffer from limited deployment and insufficient maintenance. Therefore, in order to develop representative and real-time perspectives of interference, we build a prospective mechanism on top of an existing, widely-deployed measurement platform, the RIPE Atlas network [18].

Measurement and analysis of information controls is a non-zero sum development effort. Existing platforms, such as PlanetLab [16], Herdict [11], and OONI [10] are complementary and provide unique perspectives on the diverse forms of interference. We believe that an interference analysis platform based on Atlas can provide an additional perspective to the bigger picture, one whose strengths are wide deployment, rapid results, and the foreshadowing of broader community lessons. Toward these objectives, we make the following contributions.

- We evaluate the aptitude of the RIPE Atlas platform for analysis of information controls and propose an algorithm to balance timeliness, network diversity, and cost, in order to facilitate effective analysis.
- We apply the platform and algorithm for monitoring of ongoing filtering events across different countries, and provide results based on several months of measurements.

The remainder of this paper is structured as follows. Section 2 begins by giving an overview of related work, which is then followed by our framework's structure in Section 3. After, we present two case studies in Section 4 and conclude the paper with final thoughts in Section 5.

2 Related work

It is not difficult to conduct one-off studies on filtering because administrators and governments typically do not have sufficient time to react and thwart the research. Longitudinal studies, on the other hand, are more challenging as they have to be designed in a tamper-proof and

Platform	Flexibility	Coverage	Blocking resistance	Main use
PlanetLab [16]	High	Low/Medium	Medium	Network measurements
Atlas [18]	Low	Medium/High	Medium	Network measurements
M-Lab [6]	Low	High	Medium	Network measurements
Tor [5]	Medium	Medium	Low	Low-latency anonymity
OONI [10]	High	Low	Medium	Interference analysis
Herdict [11]	Low	Low/Medium	Low	Interference analysis
OpenNet [14]	Low	Medium	High	Interference analysis

Table 1: Comparison between several popular filtering analysis platforms.

sustainable way. In 2007, Crandall et al. proposed ConceptDoppler [4]. The design enables longitudinal analysis by detecting which keywords are filtered by the Great Firewall of China (GFW) over time. More recently, CensMon was introduced by Sfakianakis et al. in 2011 [21]. CensMon is a web censorship monitor which is run on top of PlanetLab [16]. In 2012, Filastò and Appelbaum presented OONI [10]. In contrast to CensMon and ConceptDoppler, OONI is deployed and has been used successfully.[1] In parallel to these measurement tools are centrally-maintained platforms and proprietary collection agents [12, 14].

Table 1 contains a comparison between popular and deployed platforms that are or can be used for analysis of information controls. Our comparison is based on *flexibility* (i.e., how many types of measurements can be run), *coverage* (i.e., how many probes in how many countries are available), and *blocking resistance* (i.e., how easy it is for network intermediaries to disable the respective platform). We qualitatively compare all platforms and assign them the labels “Low”, “Medium”, or “High”. Note that we do not propose Atlas as *replacement* for any existing measurement platforms. Instead, we see it as a *complement* that contributes to the already existing and growing landscape of initiatives.

Additionally, in the absence of deployed platforms or other means to access machines inside countries of interest, analysts have resorted to exploiting TCP/IP side channels. In particular, Ensafi et al. demonstrated how to measure intentional packet dropping without controlling either the source or the destination machine [8].

Atlas has already been used as platform for analysis of network disruptions outside an academic setting. In 2014, Maass used Atlas to find inconsistencies in the DNS records and X.509 certificates for torproject.org [13]. In the same year, Bortzmeyer and Aben independently discussed service interference in Turkey [1, 3]. While we discuss the same topic in Section 4, we do so with significantly more data and in a more rigorous fashion.

3 Framework structure

In order to assess Atlas’s aptitude as an interference measurement platform, we continue by presenting available data collection mechanisms and our analytical framework.

3.1 RIPE Atlas background

Founded in 2010 by RIPE NCC, Atlas [18] is a globally distributed Internet measurement network consisting of physical probes hosted by volunteers. Once a user connects her probe to the network, it can be used by other participants for measurements. So-called *credits* are awarded automatically based on the uptime of contributed probes, which are expended in order to perform custom measurements. Queries to probes can be initialized centrally either over the web frontend, or over a RESTful API.

An ideal measurement platform features high geographic and topological diversity, thereby facilitating measurements in any region where filtering occurs. While Atlas probes are distributed throughout the world, there is a significant bias towards the U.S. and Europe as can be seen in Figure 1. As for Atlas’s topography, only 68 autonomous systems contain 40% of all Atlas probes with the three most common autonomous system numbers being AS7922 (4.4%, Comcast Cable Communications), AS3320 (3.2%, Deutsche Telekom), and AS6830 (2.8%, Liberty Global Operations). While not optimal, most regions of particular interest still contain at least several probes.

As of May 2014, Atlas allows four types of measurements; ping, traceroute, DNS resolution, and X.509 certificate fetching (henceforth called SSLCert). All four measurement types can further be parameterized for more fine-grained control. HTTP requests are not possible at this point due to abuse and security concerns. While Atlas clearly lacks the flexibility of comparable platforms (see Table 1), it makes up for it with high diversity, responsiveness, and continued growth. After all, we do not expect Atlas to replace existing platforms, such as OONI, but rather to *complement* them.

[1] Gathered reports are available online at: https://ooni.torproject.org/reports/.

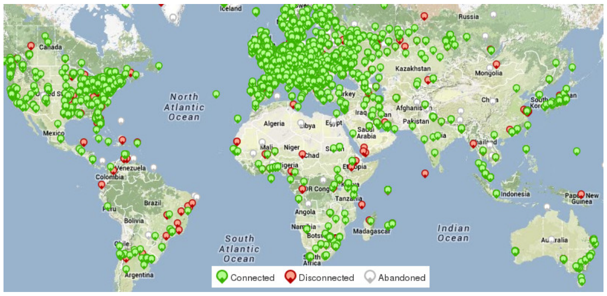

Figure 1: The geographic distribution of Atlas probes as of May 2014. Green icons represent active probes whereas red icons represent probes which are currently offline. The distribution is heavily biased towards the U.S. and Europe.

3.2 Atlas's cost model

As previously mentioned, Atlas measurements are paid with platform credits. The exact "price" of a measurement depends on the measurement type, its parameters, and the number of destinations. The credit system works based on a linear cost model. Each user has a credit balance that can be increased steadily by hosting Atlas probes[2] or by receiving credits from other users.

Table 2 lists the currently available measurement types as well as their associated costs. While DNS and SSLCert measurements have a fixed cost, ping and traceroutes vary depending on the amount and sizes of packets. Also, one-off measurements cost twice as much as repeated measurements. When scheduling a new experiment, the user first specifies the details (e.g., measurement type as well as measurement parameters). Afterwards, Atlas's web-based frontend calculates the measurement costs on the server side and shows it to the user. Finally, upon completion of the measurement, the respective cost is subtracted from the user's credit balance.

Due do the non-deterministic nature of pings and traceroutes, and measurements in general, we developed a command-line based tool to help users create new measurements and estimate their costs.[3] As input, the tool expects *1)* a country of interest, *2)* the amount of credits, the user is willing to "pay", and *3)* a measurement type. Our tool then determines the amount of available probes (if any), the expected costs, and runs the measurement if the cost is below the user's expected cost.

Measurement	Cost in credits
DNS/DNS6 (TCP)	20
DNS/DNS6 (UDP)	10
SSLCert/SSLCert6	10
Ping/Ping6	$N*(int(S/1500)+1)$
Traceroute/Traceroute6	$10*N*(int(S/1500)+1)$

Table 2: The cost for all available Atlas measurements. The variable N refers to the number of packets whereas the variable S refers to packet sizes.

3.3 Assessing measurement integrity

Despite their distribution across a diversity of countries and networks, RIPE Atlas may not fully reflect the Internet as it is experienced by the general public, as probes neither fully emulate the network position nor the configuration of an average user. As an immediate control, efforts are taken to verify the reachability of non-controversial content and identify whether probes use domestic domain name servers. These probes may be excluded from measurements in order to avoid Type 1 and Type 2 errors.

Even with additional precautions, idiosyncratic observations are an inevitable product of the high rate of placement of probes on commercial and academic networks. These institutions may have alternative connectivity that is faster and less highly regulated than consumer networks. Additionally, disrupting or degrading connections based on plain text data, traffic classification or application headers would fall outside of the measurements currently possible with Atlas probes. Lastly, Atlas, as with most measurements outlined within Section 2, is unlikely to detect content restrictions imposed by the platforms themselves, such as search manipulation or withholding of content based on a user's location.

3.4 Rough consensus validity

The international distribution of web services, such as content delivery networks, has created additional complexity in the determining whether measurement results are genuine. While SSL and DNSSEC utilize third-party trust to validate answers, Certificate Authorities have been previously compromised by state and non-state actors, and DNSSEC is not widely implemented. In order to validate answers within the Atlas network, we use a cross-country comparison of results to queries. This methodology assumes that intermediaries who interfere with connectivity do not coordinate strategies internationally. States and service providers impose different filtering approaches for purposes of localization, infrastructure or even the monetization of blocked traffic. Furthermore, interference is more effective when the

[2] As of June 2014, 21,600 credits per day of uptime.

[3] The tool is available at http://cartography.io.

public is unaware of the practices and technologies employed against them, placing a strong incentive on secrecy. Therefore, as a simple test of validity, we count the number of countries or ASNs that an answer, such as a DNS A record, final network of transit or a certificate hash, is seen. Any response with fewer than the mean number of jurisdictions, or those within private network spaces (RFC 1918 [17]), are treated as potentially aberrant and flagged for further investigation.

3.5 Ethical aspects

Atlas was not designed explicitly for analysis of information controls and accordingly, its volunteers likely may not expect that their probes will be used for such purposes. Careless measurements could attract attention and cause repercussions for probe operators. In addition, an increased used of Atlas for politically-sensitive analysis could scare away probe operators and jeopardize the usefulness of the platform. These concerns extend beyond merely complying with Atlas's acceptable usage policies and guide our selection of measurements.

Atlas's measurement types are limited in scope. As of May 2014, it is not possible to create HTTP requests or engage in actual, meaningful communication with arbitrary destinations, which limits the damage caused by reckless measurement. We are not aware of environments where low-level network queries to commonly frequented platforms or services solicit attention from authorities, even when blocked, nor where answers are falsified in order to stifle research. Requests for sites such as Facebook are generated as a part of normal web use due to script inclusion, and Google Public DNS is commonly used due to its reliability. Commonplace sites are a different class of potential monitoring targets than content promoting child abuse or violent extremism.

The balance between research interests and exposure to risk is an area of concern shared across all initiatives identified in Section 2. This has stimulated a broader discussion that will play a factor in future utilization of Atlas. Nevertheless, we stress that great care must be taken when planning measurements because the volume and types of measurements could still be suspicious. We believe that Atlas has a place in the domain of censorship analysis but it has to remain a small place, lest it endangers users or the platform itself. For a more comprehensive ethical discussion, see Wright et al. [24, § 5].

4 Case studies

After having presented our measurement platform and parameters, built on top of Atlas, we now evaluate it by discussing two cases of large-scale restrictions to online content and social media. In particular, Turkey's ban on media platforms in Section 4.1 and Russia's filtering of opposition LiveJournal content in Section 4.2. All dates and times reported follow Coordinated Universal Time.

4.1 Turkey's ban of Twitter

In late March, social media users began to report limitations on the availability of Twitter across Turkey's Internet Service Providers. YouTube and Twitter had both become the target of condemnation by Prime Minister Recep Tayyip Erdoğan in preceding months. By March 20, the Turkish government's Information and Communication Technologies Authority (BTK) mandated the filtering of Twitter across the country's service providers.

Turkey's Internet filtering has previously been characterized as DNS tampering and IP blocking [2], which both fall under the measurements possible through Atlas. Upon news of the Twitter ban, we scheduled hourly measurements of local DNS answers, SSL connectivity, and traceroute reachability for Twitter, YouTube, Google Public DNS and the Tor Project through ten probes, covering nine ASNs. The selected measurement targets sought to longitudinally document the Turkish government's disruption of controversial political content, identified based statements by authorities and potential use for circumventing controls. Seeking to address an immediate interest for real-time awareness, the measurements did not attempt to assess the whole of the country's content restrictions. As illustrated in Figure 2, we found at least six shifts in content restrictions and blocking strategies within a two week period.

While the BTK and compliant ISPs rely on DNS manipulation and IP blocking, it appears that the former is more popular. As of April 24, 2014, the Turkish-language anti-censorship site Engelliweb [7], which tracks blocked content, only lists 167 IP addresses restricted in country, compared to 40,566 domain names. In absence of address blocking or HTTP filtering, users that received valid DNS answers for Twitter's domain names could browse without further interference. As a result, foreign DNS servers quickly became both a circumvention mechanism and a political statement, with the addresses of alternative services offered by Google and OpenDNS reportedly graffitied across the the country in protest of the ban.

On the morning of March 22 (see Figure 2, **Event A**), between 01:00 and 02:00, backbone providers Tellcom İletişim Hizmetleri and Türk Telekom began disrupting Google Public DNS service through the IP blocking of its two prominent addresses (8.8.8.8 and 8.8.4.4). By 06:00 the same morning, the DNS blocking had been removed across all ISPs. Instead, to buttress the restrictions, providers shortly began to drop all outgoing traffic to IP addresses associated with the twitter.com domain,

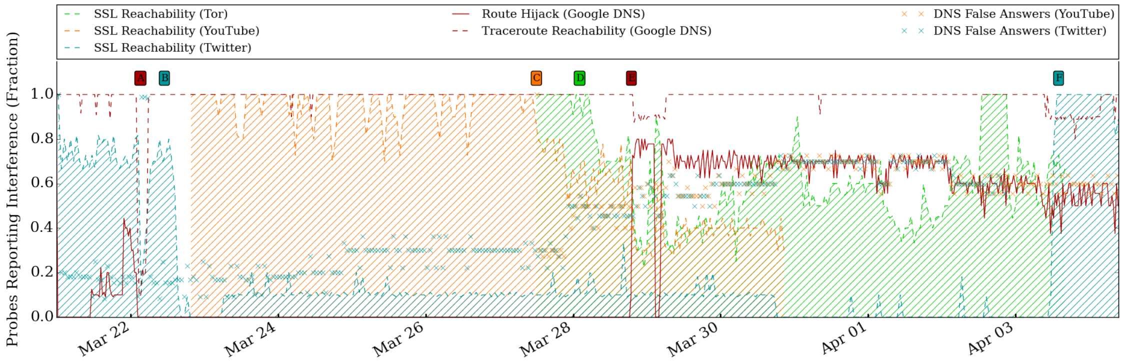

Figure 2: Disruption of Social Media Platforms in Turkey, March – April 2014

regardless of port or provider (**Event B**). By 16:00 of that day, no Atlas probe could directly negotiate an SSL connection with Twitter until the removal of the ban nearly two weeks later.

On March 27 (**Event C**), after recordings were posted of Turkish national security officials discussing possible military action against Syria, YouTube was blocked through false DNS answers for the youtube.com domain. Within the Atlas network, this restriction appears as a slow decline in the number of probes able to establish a connection to the platform. However, unlike Twitter, a significant minority of probes remained able to communicate with YouTube. Google's intertwined infrastructure presents risk of collateral damage with network prefix restrictions, which were not present with Twitter. Thus, clients that were able to receive a valid address could reliably bypass the ban.

Beginning March 28, Turkish probes began to fail to establish SSL connections to torproject.org (**Event D**). However, this restriction neither included IP blocking, nor apparent interference with the accessibility of the actual Tor network. Atlas probes could continue to negotiate valid connections to Tor's directory authories. Throughout the increased manipulation of local DNS services, nearly half of the Atlas probes remained connected due to their use of foreign DNS services.

Later in the evening, March 28, hosts querying foreign-based DNS servers began to receive the same false answers as those provided domestically, leading to a rapid drop in availability of YouTube and Tor (**Event E**). A publicly-available traceroute scheduled by third-parties on the Atlas network against Google Public DNS returned idiosyncratic and spontaneous shifts in Turkey's network topology timed with these changes. This appears within traceroutes as a shortening in the number of hops to Google, with a multifold reduction in traffic latency and the absence of international hosts in path. The core telecommunications provider Türk Telekom had begun to reroute traffic destined for Google to a local DNS server serving false answers. Only TEKNOTEL Telekom maintained consistently valid routes for Google, through Telecom Italia Sparkle. However, two days later Doruk İletişim and Net Elektronik Tasarım reestablished connectivity through Euroweb Romania, circumventing upstream interference. Türk Telekom's redirection was finally removed late on April 7.

Target	Type	Probes	Freq (s)	Credits
Twitter	SSL	10	3,600	2,400
YouTube	SSL	10	3,600	2,400
Tor	SSL	10	3,600	2,400
Twitter	DNS (U)	10	3,600	2,400
YouTube	DNS (U)	10	3,600	2,400
Twitter	Tracert	10	3,600	7,200
Total (Daily)				19,200
Probes required				0.89

Table 3: Cost of measurements for Section 4.1.

By April 3, despite continued hijacking of Google Public DNS and interference with YouTube, Twitter was unblocked for all probes (**Event F**). The total measurement credits we spent in order to conduct this experiment are shown in Table 3.

4.2 Private sector cooperation in Russian filtering of Alexei Navalny

On March 13, 2014, Russia's Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (Roskomnadzor) ordered the blacklisting of opposition figure Alexei Navalny's LiveJournal blog.

At the same time, independent media portals were filtered, including the news site grani.ru [22]. Similar to Turkey, Internet filtering in Russia is frequently conducted by IP blocking and DNS poisoning [9, 23]. However, with a random sample of 255 probes across 147

ASNs in Russia, only 38 probes on 20 ASNs received *aberrant* DNS answers for Grani. Within this subset, probes received a diverse, consistent selection of *ten unique addresses*, including two within private network address space (10.52.34.222 and 192.168.103.162). A greater selection, 40 probes across 23 ASNs, of traceroutes to port 80 for the primary address associated with Grani (as of April 30, 23.253.120.92) failed within Russia network space.

In contrast to Grani, a locally-resolved DNS query for navalny.livejournal.com over 255 probes on 146 ASNs received a consistent reply of 208.93.0.190, which matched answers internationally with only one anomalous response, a formerly valid address. The blocking of Navalny's blog must be different from Grani. While the returned DNS A record of 208.93.0.190 falls within a network prefix owned by LiveJournal Inc. (208.93.0.0/22), over the 1,462 LiveJournal subdomains in Alexa's Top 1 million list, 1,450 blogs resolved to another address, 208.93.0.150. Based on requests made independently of the Atlas network from Europe, both hosts appear to be front servers for the LiveJournal platform, as they return the same SSL Certificate and content. Requests to 208.93.0.150 with a HTTP Host header set to navalny.livejournal.com retrieves the correct content, and non-blacklisted content is retrievable through 208.93.0.190.

As of April 2014, only five subdomains on livejournal.com could be found whose DNS A records resolved to the address 208.93.0.190, Table 4, four of which are listed within Alexa's top sites. All the blogs found on this alternative host have been publicly declared by Russian authorities as in violation the country's media laws for the promotion of political activities or extremism, and two are listed within publicly-available filter site lists.

Based on timing, filtering lists, available domain names records, and Atlas network measurements, it appears that a host was specially established to facilitate Russian restrictions on content within the LiveJournal platform. Using HTTPS Ecosystem Scans as a metric of accessibility [20], the LiveJournal frontend at 208.93.0.190 came online between February 10 and February 17, with the address otherwise unused until then. Two months later, the Ukrainian LiveJournal blog 'Pauluskp' (pauluskp.livejournal.com), which had covered Russian involvement in Crimea, was filtered with the administrative order listing an IP Address of 208.93.0.190. However, as recently as six days before, the blog was recorded as pointing to the main LiveJournal host. Similarly, the movement of Navalny's blog was noticed within social media [15]. It appears that in the lead up to or at the time of filtering orders, LiveJournal coordinates with authorities to alter the DNS A record for blogs designated by Roskomnadzor, in order to segregate blacklisted content from the rest of the platform.

Subdomain	Language	Roskomnadzor
drugoi-nnover	Russian	Yes
m-athanasios	Russian	Yes
imperialcommiss	Russian	Yes
pauluskp	Russian	Yes
navalny	Russian	Yes

Table 4: LiveJournal DNS A Records of 208.93.0.190.

Segregated LiveJournal content and blacklisted addresses are subject to an additional, unknown method of network-layer interception performed within the backbone network of Rostelecom (AS12389). While blog content is not accessible over HTTPS, frontend hosts for LiveJournal offer SSL services for the purpose of securing the transmission of user credentials. On April 28, 78 of 343 Russian probes returned either irregular responses or failed to connect to the alternative LiveJournal host by address. Of this subset, 40 probes on 29 ASNs returned SSL certificates with common name or locations fields attributed to Russian ISPs. Based on HTTPS data, the four aberrant certificates captured have been seen previously on seven Russian addresses belonging to the State Institute of Information Technologies, Rostelecom and Electron Telecom Network. Three of these hosts are responsive by their alternative, public address and still match certificates. Two are generic ISP homepages and one notifies of the blocking of the site 'rutracker.ru.' Other measurements that are unresponsive could be indicative of port blocking or the redirection of traffic to a server that is not listening for SSL connections.

The invalid certificates indicate that an intermediary in transit has redirected the traffic out of its expected path to a third-party server controlled by Russian entities. This approach is different from the normal man-in-the-middle injection of responses seen in countries such as Iran and Syria, and highlights the potential for Russian ISPs to falsify content or gather user credentials. The observed behavior is not limited to protocol or port, although the end host appears to be only responsive to TCP requests, Figure 3. This holistic interference across Rostelecom's downstream peers suggests redirection at the network layer, rather than application-based classification of traffic associated with deep packet inspection. Moreover, adjacent addresses within the same network, such as the normal frontend for LiveJournal, traverse a valid international path. Instead, blacklisted traffic appears to be coerced into a path controlled by Rostelecom, indicating a narrowly-crafted interference with normal routing through false advertisements or forwarding.

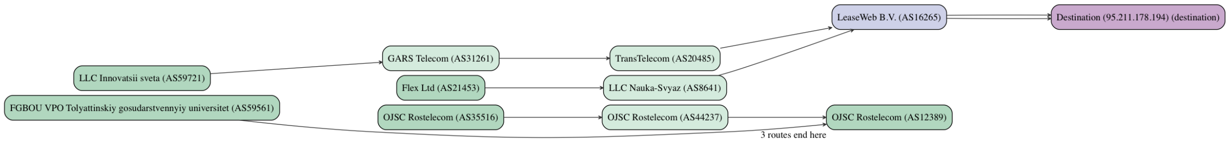
LeaseWeb B.V. (AS16265)
Destination (95.211.178.194) (destination)
GARS Telecom (AS31261)
TransTelecom (AS20485)
LLC Innovatsii sveta (AS59721)
Flex Ltd (AS21453)
LLC Nauka-Svyaz (AS8641)
FGBOU VPO Tolyattinskiy gosudarstvennyiy universitet (AS59561)
OJSC Rostelecom (AS35516)
OJSC Rostelecom (AS44237)
OJSC Rostelecom (AS12389)
3 routes end here

Figure 3: Rostelecom's (AS12389) hijack of grani.ru Traffic in April 2014.

5 Conclusion

In this paper, we have presented a model of an interference detection platform that builds on top of the RIPE Atlas platform. Previous examinations of Internet filtering have tended to analyze specific national apparatuses on a per-country unit, assuming internal consistency across providers and time. This past approach has been appropriate for describing the diversity of methods used to control access globally, as well as for when the primary research focus is on countries that impose restrictions at central points of international transit.

As Internet filtering has proliferated to countries with competition and private markets at the international frontier, researchers can no longer assume direct and consistent control by authorities. The two recent and developing cases of interference in Russia and Turkey demonstrate this shifting environment. Russia and Turkey's networks are more administratively and technically decentralized than China and Iran [19]. Through longitudinal observation, our initial research demonstrates substantive differences of methods and rates of implementation for content restrictions. In both, the Atlas network provided a unique opportunity for documenting rapidly-evolving information controls due to its nearly ubiquitous geographic presence, stability for recurrent measurements, and external queuing of targets. Reliance on alternative models outlined in Section 2 would have imposed delays on deployment, and limited the vantage points from which data could be collected.

These findings contribute to broader discussions on anti-filtering strategies. Collateral damage, urgency and level of difficulty appears to have shaped the implementation of Turkey and Russia's filtering mandates. The quick removal of restrictions on Google Public DNS, and then attempts to impersonate the service, indicate that enforcing an absolute prohibition on content is partially an economic question. Where there are high collateral costs, such as with Google infrastructure in Turkey and LiveJournal in Russia, authorities appear to have limited their restrictions or found cooperative arrangements with platform owners.

Atlas was well positioned for documentation of both blocking incidents based on telecommunications companies reliance on interfering with network reachability and domain name translation. If administrators had utilized traffic inspection, or more subtly degraded connectivity without outright blocking access, the platform would not have been capable of measuring these events.

Despite these analytical precautions, Atlas-based measurements provide an early perspective on the opportunities and methodologies possible with pervasive network observation. We document multifaceted filtering infrastructures in both countries, notably reliant on DNS manipulation and redirection of traffic by transit providers. Additionally, the latter manipulation of network routes represents an under-explored method of interference and invokes the need for tools to collect path information to complement other forms of documentation. Furthermore, differences of restrictions shed light on inconsistencies in the application of administrative orders, and could provide early warning of increased controls in the future. Our initial research demonstrates that across national networks there are substantive differences of methods, rates of implementation, and, in at least one case, even selective compliance for information controls that are measurable by Atlas and future initiatives.

Finally, our code and data sets are available online at: http://cartography.io.

Acknowledgments

We would like to thank the anonymous reviewers for their constructive feedback on earlier versions of this paper. Additionally, we would like to thank Vesna Manojlovic for providing startup credit and the RIPE Atlas community for their insight. Collin Anderson was supported by the Internet Policy Observatory program at the University of Pennsylvania's Annenberg School for Communication. Philipp Winter was supported by a research grant from Internetfonden.

References

[1] Emile Aben. *A RIPE Atlas View of Internet Meddling in Turkey*. 2014. URL: https://labs.ripe.net/Members/emileaben/a-ripe-atlas-view-of-internet-meddling-in-turkey.

[2] Yaman Akdeniz. *Report of the OSCE Representative on Freedom of the Media on Turkey and The Internet Censorship*. Tech. rep. OSCE, 2010. URL: http://www.osce.org/fom/41091?download=true.

[3] Stéphane Bortzmeyer. *Hijacking of public DNS servers in Turkey, through routing*. 2014. URL: http://www.bortzmeyer.org/dns-routing-hijack-turkey.html.

[4] Jedidiah R. Crandall et al. "ConceptDoppler: A Weather Tracker for Internet Censorship". In: *CCS*. ACM, 2007. URL: http://www.csd.uoc.gr/~hy558/papers/conceptdoppler.pdf.

[5] Roger Dingledine, Nick Mathewson, and Paul Syverson. "Tor: The Second-Generation Onion Router". In: *USENIX Security*. USENIX, 2004. URL: http://www.onion-router.net/Publications/tor-design.pdf.

[6] Constantine Dovrolis et al. "Measurement Lab: Overview and an Invitation to the Research Community". In: *Computer Communication Review* 40.3 (2010), pp. 53–56. URL: http://www.sigcomm.org/sites/default/files/ccr/papers/2010/July/1823844-1823853.pdf.

[7] Engelliweb. URL: http://engelliweb.com.

[8] Roya Ensafi et al. "Detecting Intentional Packet Drops on the Internet via TCP/IP Side Channels". In: *PAM*. Springer, 2014. URL: http://arxiv.org/pdf/1312.5739.pdf.

[9] Federal Service for Supervision of Communications, Information Technology, and Mass Media. 2013. URL: http://rkn.gov.ru/docs/Analysys_and_recommendations_comments_fin.pdf.

[10] Arturo Filastò and Jacob Appelbaum. "OONI: Open Observatory of Network Interference". In: *FOCI*. USENIX, 2012. URL: https://www.usenix.org/system/files/conference/foci12/foci12-final12.pdf.

[11] *Herdict*. URL: http://www.herdict.org.

[12] Tim Hwang. "Herdict: A distributed model for threats online". In: *Network Security* 2007.8 (2007), pp. 15–18. URL: http://cartography.io/pdf/Hwang2007.pdf.

[13] Max Jakob Maass. *torproject.org censorship detection using RIPE atlas?* 2014. URL: https://lists.torproject.org/pipermail/tor-talk/2014-February/032173.html.

[14] *OpenNet Initiative*. URL: https://opennet.net.

[15] Moscow Institute of Physics and Technology. URL: http://board.rt.mipt.ru/?read=8820778.

[16] *PlanetLab*. URL: https://www.planet-lab.org.

[17] Y. Rekhter et al. *Address Allocation for Private Internets*. RFC 1918 (Best Current Practice). Internet Engineering Task Force, Feb. 1996. URL: http://www.ietf.org/rfc/rfc1918.txt.

[18] *RIPE Atlas*. URL: https://atlas.ripe.net.

[19] Hal Roberts et al. "Mapping Local Internet Control". In: *Computer Communications Workshop*. IEEE, 2011. URL: http://cyber.law.harvard.edu/netmaps/geo_map_home.php.

[20] Mark Schloesser et al. *Project Sonar: IPv4 SSL Certificates*. URL: https://scans.io/study/sonar.ssl.

[21] Andreas Sfakianakis, Elias Athanasopoulos, and Sotiris Ioannidis. "CensMon: A Web Censorship Monitor". In: *FOCI*. USENIX, 2011. URL: http://static.usenix.org/event/foci11/tech/final_files/Sfakianakis.pdf.

[22] International Business Times. *Kremlin Blocks Four Opposition Websites As Ukraine Crisis Brews*. 2014. URL: http://www.ibtimes.com/kremlin-blocks-four-opposition-websites-ukraine-crisis-brews-1561356.

[23] John-Paul Verkamp and Minaxi Gupta. "Inferring Mechanics of Web Censorship Around the World". In: *FOCI*. USENIX, 2012. URL: https://www.usenix.org/system/files/conference/foci12/foci12-final1.pdf.

[24] Joss Wright, Tulio de Souza, and Ian Brown. "Fine-Grained Censorship Mapping Information Sources, Legality and Ethics". In: *FOCI*. USENIX, 2011. URL: http://static.usenix.org/event/foci11/tech/final_files/Wright.pdf.