Measuring Internet Censorship in Iran from Multiple Vantage Points

Neeki Hushyar
UMass Amherst
nhushyar@cs.umass.edu

Angela Upreti
UMass Amherst
aupreti@cs.umass.edu

ABSTRACT

We analyze Internet censorship in Iran through data collected using ICLab and RIPE Atlas. We use results from HTTP GET requests, TCP connect requests and DNS queries to confirm previous findings reported by Halderman et al. in 2013. Halderman et al. used a single VP in Iran whereas we use n different VPs to determine the consistency of the results from different vantage points. Through DNS queries conducted using network measurement platform, RIPE Atlas, we measure the scope of the censorship by probing 60 different domains. Through the HTTP GET requests, tcp connect requests and DNS queries conducted using ICLab, we measure the scope of the censorship by probing more than 500 domains. These domains include a test list for Iran compiled by citizen lab and top websites for Iran reported by the Alexa website. We compare the results obtained through Iranian VPs with the results from a VP located with the US. We conclude with the results confirming the blocked webpages either by way of HTTP filtering or DNS hijacking as previously reported by Halderman. In addition, we find that the censorship encountered at varying vantage points is consistent. In the process, we also discover some weaknesses of the ICLab platform and offer some recommendation.

CCS Concepts

•**Computer systems organization** → **Embedded systems;** *Redundancy;* Robotics; •**Networks** → Network reliability;

Keywords

ACM proceedings; LaTeX; text tagging

1. INTRODUCTION

Freedom House, the U.S. based, government-funded, nonpartisan organization which researches the levels of democracy, political freedom and human rights in countries all over the world, ranked Iran as one of the worst countries in terms of free internet, second only to China. The government of Iran blocks social media and political and social content from a wide range of categories. Over the past few years users of blocked domains in Iran have faced persecution and circumvention tools, which assisted civilians in accessing blocked domains, have been blocked through a variety of methods. The affects of internet censorship and lack of net freedom have been widely publicized. Notable cases include a complete shutdown in access to social media beginning during the 2012 presidential election campaigns in Iran, the 2014 arrests of Facebook activists who had used the social media website to motivate government reform, and unrelated arrests that same year of individuals who insulted the Supreme Leader, Imam Khomeni, over a text message application.

2. BACKGROUND

3. RELATED WORKS

The previous experiments, pinpointing the methods used for and sources of censorship in Iran was conducted in 2013. These experiments revealed many methods used by ISP and DNS resolvers in Iran in the effort to censor Internet users. Those methods include:

DNS Redirection: The return of false, redirected pages in response to DNS requests for censored domain names. The previous experiments reported this redirect IP as 10.10.34.34.

HTTP Filtering: Access to specific websites blocked, based either on the content of the GET request itself, or the content in the response packet[s].

Connection Throttling: Iran reportedly throttles the connection of specific connections, namely those running over SSL or virtual private networks implemented using IPSec, among others methods of encryption. Furthermore, attempts to access specific websites, independent of protocol were reported to cause connection throttling.

Broadband Speed Limitations: The maximum bandwidth afforded to home userâĂŹs is limited. This was previously believed to be enforced as a method of censorship, however more recent reports reveal this may be forced by limited network capabilities.

Our experiment simulates the HTTP filtering and DNS redirection described above in order to determine the consistencies of the filters and redirections. The previous experiment, conducted in 2012, was the first to pursue the technical aspects of the Internet censorship in Iran. Most of the research was done from a single vantage point and our experiment seeks to expand upon them.

WOODSTOCK '97 El Paso, Texas USA

DOI: 10.475/123_4

4. EXPERIMENTS AND RESULTS

We conducted experiments using ICLab and RIPE Atlas. ICLab is a research platform used to measure global Internet censorship. ICLab consists Raspberry Pis and VPN vantage points (VPs) located all around the word. This platform allows researchers to run sef-written experiments on chosen VPs. ICLab was used to conduct tests targeted at over 500 domains. The domains tested fall in a variety of categories, including the top Alexa domains as accessed by people within Iran as well as several hundred additional domains, which were previously reported to be blocked.
RIPE Atlas is a global network of probes that measure Internet reachability and connectivity. RIPE was used determine several vantage points from which to send DNS queries.

4.1 ICLab

ICLab platform was used to conduct DNS, traceroute, TCP connect, TLS handshake and HTTP GET measurements. All except the traceroute experiments were successful.

4.1.1 DNS:

DNS resolution of each domain in our list was conducted from the Iranian VP using two resolvers: A local resolver of VP's choice and the google resolver, 8.8.8.8 were used. A python script was then used to compare the DNS responses from a US-based VP to the results obtained from Iran.
All queries using the google resolver yielded a null response. Some queries using the local resolver also yielded a null response even though a valid IP was obtained for the query using the US-based VP. Figure x shows a comparison of DNS responses from a US VP and a Iranian VP.

4.1.2 HTTP:

We executed HTTP GET requests from sources in both the United States and Iran to use as control and experiment results, respectively. We analyzed the differences and compiled a list of inaccessible addresses by comparing the results from Iran to the results we received in the United States. Our original tests from the U.S. resulted in a negligible number of failures to resolve which were of the form: HTTP 400 Bad Request or HTTP 404 Page Not Found. This occured in the case the request timed out or the domain name was intentional invalid. Invalid domains were used to test for HTTP request filtering by internet service providers (ISPs) in Iran. An invalid domain should return a 404 Page Not Found Error, however, the invlaid domains which were blocked resulted in 400 Bad Request or 403 Forbidden.

4.1.3 Traceroute:

We setup our experiment to conduct traceroute measurements before we handed it over to the person responsible for ICLab. Our intention was to pinpoint the location of the censor within IranâĂŹs network by using the traceroute emasurements. However, we got "traceroute not found or not installed" error for all of our measurements. We see this as one of the limitations of the ICLab platform compared to RIPE atlas. Researchers do not control the remote machines in the test countries so they lack a clear expectations of what kinds of tests will be successful.

4.2 RIPE Atlas

4.2.1 DNS:

As reported in the previous study, the method used to conduct DNS hijacking is false resolutions. We encountered several domains for which the DNS resolutions returned a false redirect address, 10.10.34.36 instead of the domainâĂŹs correct IP address.
RIPE Atlas allowed us to set specific parameters for the DNS requests. First, we used a total of three different source probes from which our DNS requests were sent. Parties affiliated with RIPE Atlas operate the probes. Second, we specified the IP addresses of specific resolvers in a variety of locations within Iran. The constraints, from which we narrowed down our choice in resolvers by, include reliability and location. In total, the experiments used a series of three different probes and five DNS resolvers.
Our aim was to determine if the false resolution results vary depending on probe and/or DNS resolver. We used the top 40 most often accessed domains in Iran from Alexa as well as 20 additional domains previously reported to be victims of DNS hijacking.
Our results revealed that there probe IP from which the DNS query was sent from had no effect on the accuracy of the result from any one resolver. That is to say, given two separate probes, resolving the same domain name at the same resolver, always received the same response. This suggests that all of the DNS requests were sent as is, and never tampered with on the way to the intended resolver. On the other hand, we determined that the resolver themselves were not centralized in terms of where they retrieved their rules from. In other words, the different resolvers could be inconsistent as we saw several cases of some resolvers returning false resolutions for the same domains that other resolvers returned accurate IP addresses for. As a result of this inconsistency, we compiled three categories in which our tested domains fall in.

Always Resolves: The domain names in this category were all accurately resolved. The resolutions were verified by through the use of scripts that ran the whois command to confirm the organization name matched the intended domainâĂŹs organization name. The initial tests executed 60 domain name resolution requests from three probes to three targeted resolvers. Thus each domain was tested from three probes, targeted at three solvers, for a total of 9 resolutions per domain. A domain was assigned to this category if all of the resolutions were accurate.

Resolves: Our initial test of 60 domain names over three probes and three target resolvers revealed inconsistencies in resolutions. Specifically, of our first three resolvers, located in Tehran (2) and Isfahan, the results revealed that one resolver in Tehran, would consistently return accurate resolutions for domains which returned the 10.10.34.36 redirect pages by the other two resolvers. To determine whether or not this anomalous resolver was one of a kind, or if there exists true diversity in replies, depending on the chosen DNS resolver, we extended further testing. Our additional tests utilized the same three probes, with two additional resolvers in Tabriz and Zehadan. We tested every domain that returned the false redirect address in the previous round of resolutions, against these additional resolvers. We found further inconsistencies. The domain names in this category were accurately resolved by some resolvers and falsely resolved by others. This category contained a wide variety of domains in topics including: human rights, local blogs, national news and censorship circumvention.

Resolves: The domain names in this category were never accurately resolved. This applies to all five resolvers. The category contained domains from topics including: political reform, gay+lesbian, international news, and western social media. The result returned from each DNS request in this category, was either no result at all, as a result of a timeout, or the redirect IP, 10.10.34.36.

Figure 1. Shows the accurate DNS resolution rate of a few of the queried domains

Figure 2. Shows the percentage accuracy per DNS resolver, of the domains that were blocked by at least one resolver (Percentage blocking per resolver location of âĂIJBadâĂİ Domain Names).

5. CONCLUSION

6. FUTURE WORKS

7. ACKNOWLEDGMENTS

This section is optional; it is a location for you to acknowledge grants, funding, editing assistance and what have you. In the present case, for example, the authors would like to thank Gerald Murray of ACM for his help in codifying this *Author's Guide* and the **.cls** and **.tex** files that it describes.

APPENDIX

.

.1 References

Generated by bibtex from your .bib file. Run latex, then bibtex, then latex twice (to resolve references) to create the .bbl file. Insert that .bbl file into the .tex source file and comment out the command `\thebibliography`.