Monitoring Internet Censorship with UBICA

Giuseppe Aceto, Alessio Botta, Antonio Pescapè, Nick Feamster, M. Faheem Awan, Tahir Ahmad, Saad Qaisar

Monitoring Internet Censorship with UBICA

Giuseppe Aceto[1] ⋆, Alessio Botta[1], Antonio Pescapè[1] ⋆ , Nick Feamster[2], M. Faheem Awan[3], Tahir Ahmad[3], and Saad Qaisar[3]

[1] University of Napoli Federico II (Italy),
{giuseppe.aceto, a.botta, pescape}@unina.it
[2] Georgia Institute of Technology (GA, USA), feamster@cc.gatech.edu
[3] NUST SEECS (Pakistan),
{10mscsemawan,11msccstahmad,saad.qaisar}@seecs.edu.pk

Abstract. Censorship is becoming increasingly pervasive on the Internet, with the Open Net Initiative reporting nearly 50 countries practicing some form of censorship. Previous work has reported the existence of many forms of Internet censorship (e.g., DNS tampering, packet filtering, connection reset, content filtering), each of which may be composed to build a more comprehensive censorship system. Automated monitoring of censorship represents an important and challenging research problem, due to the continually evolving nature of the content that is censored and the means by which censorship is implemented. UBICA, User-based Internet Censorship Analysis, is a platform we implemented to solve this task leveraging crowdsourced data collection. By adopting an integrated and multi-step analysis, UBICA provides simple but effective means of revealing censorship events over time. UBICA has revealed the effect of several censorship techniques including DNS tampering and content filtering. Using UBICA, we demonstrate evidence of censorship in several selected countries (Italy, Pakistan, and South Korea), for which we obtained help from local users and manually validated the automated analysis.

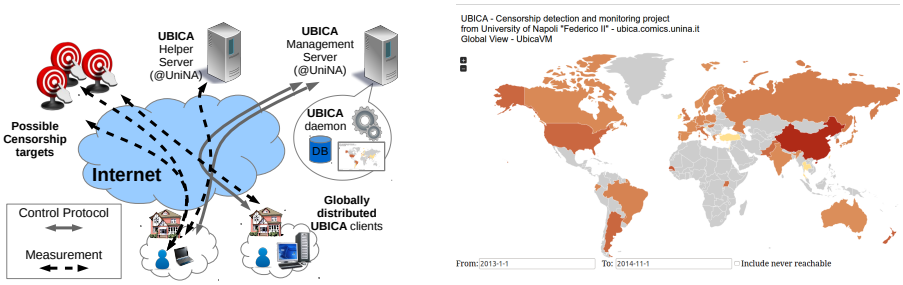
1 Introduction

Akin to network monitoring for faults, attacks, and performance variations, Internet censorship monitoring is a relatively new field of research with methodologies, tools and practices still in course of definition. We consider Internet censorship *detection* as "*the process that, analyzing network data, reveals impairments in the access to content and services caused by a third party (neither the client system nor the server hosting the resource or service) and not justifiable as an outage*". In turn, Internet censorship monitoring is the automated and continuous process of detecting Internet censorship over time, with the aim of revealing status changes in terms of the affected targets or the adopted censoring techniques. Regardless of the ethical and political positions regarding censorship, the interference with Internet protocols standard and intended behavior

⋆ This work has been carried out thanks to a *Google Faculty Research Award* for the project UBICA (User-Based Internet Censorship Analysis).

has practical implications. Moreover significant aspects of censorship, such as its enforceability, its transparency, and the accountability of the censors to the affected population, strongly depend on the technical details of the censorship technique adopted and thus evolve with both the technology and its application in practice. Collection of the appropriate network measurements for monitoring censorship is thus a fundamental part of understanding the existence, prevalence and evolution of censorship [3,12], and to tell it from unintentional network outages or performance issues. Although tools for monitoring censorship abound, most of them do not base their analysis or conclusions on widespread, scalable, continuous network measurement. One well-known censorship monitoring tool is Herdict [1], a crowd-sourced platform. Its main interface is a website allowing users to report about "accessibility"of URLs from within their browser; this way the platform leverages crowdsourcing both for the collection of *targets* of interest for the users, and by having the users to perform an application-level censorship test. A browser plugin also allows users to submit reports without accessing the web interface. Another tool called *CensMon* is specifically designed for censorship monitoring [16]. It is designed for continuous and automatic functioning, and addresses the "needle in a haystack" problem of selecting *targets* worth checking by feeding the system with URLs automatically harvested from a variety of online sources. The most complete and wide-ranging tool for censorship detection is provided by the OONI project [10]. It is a Free Software project, part of the wider *Tor project* with which it is tightly integrated. The main component is a Python script offering a list of censorship detection tests to be performed using Tor. In addition to platforms or tools for censorship detection and monitoring, previous work has performed many studies of various censoring systems and techniques [8,18], often focused on the Great Chinese Firewall [5,15,19], or investigate outage-like censorship events [7,2].

In this paper we discuss results obtained by means of a platform for censorship monitoring called UBICA, standing for User-based Internet Censorship Analysis. Due to space constraints, we focus on the results and on the analyses allowed by the platform, and describe the platform at a functional level, referring to future works for a more in-depth discussion. UBICA adopts an integrated and multi-step analysis and provides a simple but effective dashboard thanks to which censorship events are easily spotted and described also in their temporal evolution. UBICA integrates an algorithm for detecting censorship based on Internet measurements: if the test finds evidence of blocking, additional tests attempt to identify possible mechanisms, including DNS blocking, IP blocking, No HTTP Reply, RST (TCP-level tampering), Infinite HTTP Redirect, and Block page. Using UBICA for several months on selected targets, we found evidence of several censorship techniques, such as DNS tampering and content filtering. We validated the accuracy of UBICA with the help of users in selected countries and also show evidence of censorship in several countries (Italy, Pakistan, and South Korea).



(a) Architecture diagram. (b) Report interface (detail: global map).

Fig. 1: The UBICA platform.

2 UBICA

The main objective of UBICA is to provide users with a *censorship monitoring* system that presents both a report on world-wide Internet censorship status and a quick view of censorship from users' perspectives. To gather data, the platform leverages a distributed deployment of probes belonging to different kinds (router-based, headless client, GUI-client) that are orchestrated by a central management server. The platform provides: (i) dynamically updated censorship tests; (ii) dynamically updated targets to be verified; (iii) support for different types of probing clients; (iv) automatic censorship detection and censorship technique identification. Fig. 1a shows the UBICA architecture. An example of monitoring report is shown in fig. 1b.

Monitoring control flow. The collection of evidences of censorship is performed through *active measurements* from the probes, that periodically retrieve from the Management Server a list of test requirements (eventually updating the necessary targets lists and code). The target lists are build from up-to-date reports from Herdict [1], a list of worldwide top accessed websites, and lists suggested from in-country volunteers; they are distributed to probes based on the country the probe is located at. After the evidence collection each probe packs all the results in a report file and uploads it back to the Management Server. Such server asynchronously parses the reports and inserts the relevant information into an SQL Database. The *Analysis Engine* periodically processes data in the database, performing the censorship detection analyses described in the *Experimental results* section. The different types of measurements performed are described in the following.

DNS resolution. To collect clues about this phase, a name resolution is elicited: given a fully qualified domain name, a DNS request of `type A` is issued from the probe towards its default resolver. The tool used to issue the request is `nslookup`. To distinguish among different DNS tampering techniques, the same

request is issued also towards a list of open resolvers, used as *control* resolvers from inside the censored network. The list of open resolvers is the same as the one used in [14].

TCP reachability. To check for filtering triggered by *IP:port*, this test tries to establish a TCP connection, starting a three-way handshake with a given timeout. The tests takes as parameters *targetIP:port* and a timeout value in seconds, that has been set by default to $15s$.

HTTP reachability. This test issues an `HTTP GET` request: the response (or lack of it) from the server is collected, along with application level values. The HTTP header field `User-Agent` is chosen randomly from a list of the most common user agent strings, according to [9]. The tool used to issue the request and collect application level information is `curl`.The report from this test includes several values, such as content type, HTTP response code, number of redirects, etc., not reported for the sake of brevity.

3 Experimental Results

With the help of professional and personal contacts, a number of software probes have been deployed in different countries worldwide, plus more than a dozen BISmark routers [17] from an experimental deployment in Pakistan, one in Italy and another one in USA. The distributed platform PlanetLab [4] has also been employed, deploying UBICA probes in the most diverse set of countries available at the time of the experiments. The measurement campaigns have been conducted using more than 200 probes, constituted by: 47 clients with GUI (run by volunteers both in Italy and abroad); 188 headless clients (of which 19 run by volunteers worldwide and 169 in PlanetLab nodes); and 16 BISmark home routers run by volunteers (mostly in Pakistan). The target lists for each country included Herdict reports for the country, a list of worldwide top accessed websites, and URLs suggested by local volunteers. Measurements have been made from 31 different countries, testing more than $16K$ different *targets* (about $15K$ different hostnames) on a timespan of 4 months.

The application of the UBICA detection algorithm to data collected in this experimental campaign and the time analysis of the related outcomes have tested the functionalities of the platform. In the following we report an extract of the most interesting results, concentrating on those for which we had a ground truth.

3.1 Censorship in Pakistan

In *ONI* country profiles, Pakistan (PK) is classified as applying "selective filtering", showing a consistent level of censorship and tight control on Internet communications across the national border. The government body Pakistan Telecommunication Authority (PTA) is in charge of the management of the Pakistan Internet Exchange, the exchange point connecting the country to the rest of the Internet, and maintains a blacklist of URLs to be censored [14]. According to the last report from *The OpenNet Initiative*, blocked resources belong to the classes: religion, sex, and politics.

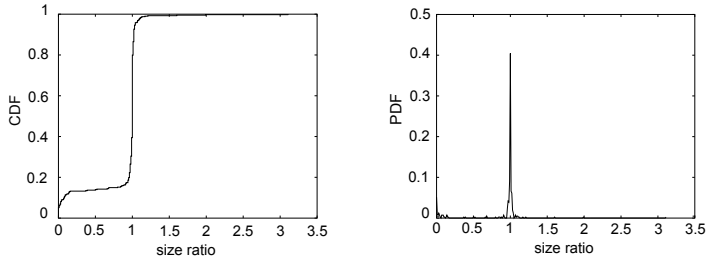
A general view. Our experimental campaigns performed through UBICA probes in Pakistan evidenced that many resources were actually censored in this country. The censorship detection algorithm reported that the techniques used were mainly two: DNS injection and HTTP tampering. To understand what happened and to confirm these results we analyzed the intermediate data, comprising the results of the different tests performed by UBICA. We describe the overall results and the details about the intermediate ones in the following.

As for DNS, 68% of the resources are identically resolved from inside PK and USA (USA has been used for comparison purposes). Thus the algorithm for censorship detection excluded the occurrence of DNS-based censorship for the related resources. Therefore, for the remaining resources, the analysis has exploited information about the size of the resource (the *content size* tests). Similar analysis based on content size has been recently published in [11], but it leverages the availability of a ground truth, i.e., a copy of the content known to be uncensored, to compare with. Our algorithm, described hereafter, does not need such knowledge.

Considering the size of the resource (webpage) that has been retrieved, and averaging on all measurements from within a country, we expect to find a significant difference between different countries if one of the two is censoring the content by means of a "blocking page". For each URL u, the average resource size per country $s_{u,PK} = \frac{\sum_{u in PK} size(u)}{|PK|}$ is calculated and divided by the corresponding size averaged on all the other countries; as an example, we show the ratio with USA in this case, but in the following reports the more general setup is adopted. Considering the empirical CDF of such ratio (Fig. 2a), we can see that while most of URLs show a comparable average size, there is an interesting fraction of them whose size is much smaller in Pakistan than in USA. The empirical probability mass function distribution reported in Fig. 2b clearly shows two modes: one centered in 1 and a smaller one close to 0. The variability around 1 can be considered as due to differences in parts of the HTML code that are updated in the dynamic generation of the resource. The relatively big variations that lead to the mode close to zero hint to a different phenomenon, on which we will focus to find evidence of censorship. To differentiate between the two modes, we choose a threshold of 0.3, which is halfway between the two modes minus a guard interval of 0.2 to account for variability across multiple countries and coherently with the design principles of the detection algorithm. An excerpt of some URLs whose size ratio falls below this threshold (in total 56, of which 28 are *youtube* videos) are reported in Tab. 1. We took one of the URLs selected through the *average content size ratio* test, namely *ninjaproxy.com* (accounting for $343Bytes$ in Pakistan and $14753Bytes$ from USA) and looked at the HTML code received by the client in Pakistan. The inspection confirmed that the page is completely different from the one retrieved from outside Pakistan (not shown for space constraints). Indeed censorship has been enacted providing a webpage with iframe redirection to a blocking page. These results are consistent with [14], and the analysis in the report by *The Citizen Lab* on this country. More details

URL	size PK	size USA	Ratio
barenakedislam.wordpress.com	453.0	49095.63	0.01
ninjaproxy.com	342.45	14085.42	0.02
NinjaProxy.com	342.39	13154.06	0.03
www.similarsites.com	375.33	13701.44	0.03
www.youtube.com	4183.91	144177.2	0.03
www.freefacebookproxies.com	9041.17	241485.33	0.04
friendlyatheist.com	7881.34	205294.23	0.04
www.loonwatch.com	2661.73	65075.19	0.04
www.sodahead.com	3575.67	73969.7	0.05
www.hotspotshield.com	731.8	10789.91	0.07
face-of-muhammed.blogspot.com	6208.7	85342.93	0.07
www.foxnews.com	4705.53	63425.26	0.07
www.buzzfeed.com	22097.93	287001.77	0.08
www.freefacebookproxies.com	18245.93	233254.73	0.08
www.hotspotshield.com	870.1	10632.97	0.08
www.cagle.com/news/muhammad	3594.5	40974.12	0.09
www.smugbox.com/facebook/...	1883.93	21455.95	0.09
www.faithfreedom.org/Gallery/...	1438.93	15423.32	0.09
www.turbohide.com/	896.91	8744.12	0.1
www.unblockbook.net	812.48	6348.47	0.13
www.thesecretninjaproxy.info	469.79	3416.17	0.14
www.kproxy.com.	647.47	4694.55	0.14
www.kproxy.com	666.39	4618.71	0.14
www.unblock-facebook.net	840.26	5783.3	0.15
www.blockedsiteaccess.com	1271.46	7780.19	0.16

Table 1: Selection of URLs whose content size ratio (size PK divided by size USA) is smaller than 0.3; URL path is truncated for presentation constraints.

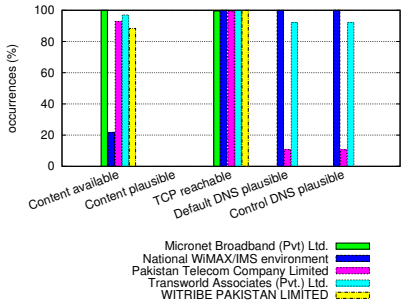


(a) Empirical CDF (b) Empirical PDF

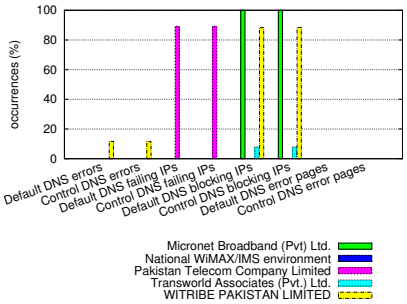
Fig. 2: Distribution of content size ratios (size PK divided by size USA) for each URL, tested URLs are from [14] (468 URLs).

on reports generated by UBICA are described in the following for specific *targets* that better expose the detection algorithm inner working.

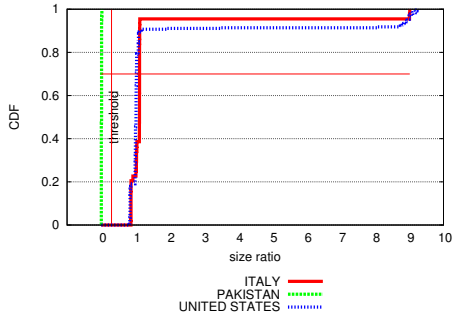
The case of YouTube. One of the final results of the UBICA detection algorithm is the summary of the censorship techniques detected for a given *target* as accessed from different ISPs. This report shows an evaluation of censorship conditions and technologies in the considered country for the specified resource. An example of blocked URL showing interesting differences among ISPs is the streaming video platform - with content and comment sharing from users - *YouTube* (`www.youtube.com`), integrated with the social network *google plus* and the search engine google). The report that UBICA generated for the URL of a resource on *YouTube*, as tested from different ISPs in Pakistan, is shown in form of a bar chart in Fig. 3a. The lack of bars in the second aggregate (with label "Content plausible") means that this resource is never reachable, even though for all but one ISP, a resource is returned when performing an HTTP request (first aggregate of bars, labeled "Content available"). We recall that "Content plausible" is the percentage of URLs that passed the *size ratio* censorship test, and thus present a content size comparable to the average on all countries. The outcome of this test is represented in Fig. 3c as a CDF of the ratio of the size of the downloaded content in one sample over the global average of such size. The CDF generated for Pakistan is shown (in green) along with other countries



(a) Comparison of results for different censorship techniques.



(b) Detail of DNS analysis.



(c) Empirical CDF of content size ratio

Fig. 3: Censorship in Pakistan: the case of YouTube.

for comparison: Italy (cyan) and U.S.A. (in dark blue); the aggregation level is *country*, thus considering samples for the whole nations regardless of the ISP. The graph shows clearly that the size ratios in Pakistan are close to 0 (i.e. the content size is very small compared with the global average) with relative frequency 1 (always), while for both the other countries the occurrences fall close to 1 (thus same content size as the global average) with relative frequency greater than 0.9 (for U.S.A. 0.91 for a size 1.23 times the average, for Italy 0.95 for a size 1.11 times the average). Comparing with the size ratio threshold (set to 0.3) we notice that the test has correctly separated results in Pakistan from the ones in the other countries. Moreover as the detected condition is above the coherence threshold, the reported results are consistent over each country dataset.

TCP-level tests (Fig.3a, third aggregate, label "TCP reachable") show almost 100% reachability for all the ISPs, thus either no censorship is enacted at this layer, or DNS tampering precedes it. By considering the *default* DNS results for two ISPs "Micronet Broadband (Pvt) Ltd." and "Witribe Pakistan Ltd." no result yields a plausible IP address (i.e. neither a known block page or a failing IP, nor a DNS error), similarly for "Pakistan Telecom Company Ltd." only 11.7% is plausible. These ISPs clearly block the resource with *DNS tampering*. The DNS overall results show equal values for the *default* and the *control* resolvers, thus the inferred technique is DNS *injection*. The ISPs "Transworld Associates" (cyan in Fig. 3b) and "National Wi-Max/IMS" (dark blue) do not perform *DNS tampering* on the resource under analysis; yet for both the *content*

size ratio analysis has detected censorship: an *HTTP tampering* technique has been applied. To gather information regarding the *symptom* the user gets in the censored networks, we leverage the detailed DNS analysis, shown in Fig. 3b. It can be noted that, while two ISPs (namely, "Micronet Broadband (Pvt) Ltd." and "Witribe Pakistan Ltd.") both use DNS tampering to provide the user with an explicit *block page*, the ISP "Pakistan Telecom Company Ltd." provides an address that will likely cause an error (either at TCP-level or an HTTP-404), thus confounding the customer without providing explicit notification of censorship.

From the comparison between the summarized view (Fig. 3a) and the DNS analysis details (Fig. 3b) the behavior of one ISP ("Pakistan Telecom Company Ltd.", in magenta) seems inconsistent with the expected *symptom*, as the detected technique ("DNS injection - failing IP") should have elicited an error, and not the high percentages found both in *TCP reachable* and *Content available* bars (3a). By inspecting the collected evidence data it resulted that the IP address returned by the ISP under analysis is `127.0.0.1`, corresponding to `localhost`, i.e. for each machine is the address of the machine itself (network level loopback). While other "specialized" network address ranges [6] are unlikely to be assigned to active hosts in the same LAN of the probe, `localhost` for sure is, and the outcome of a TCP connection to the port 80 and possibly an HTTP request depend on the presence of a service listening on that port, and the response the service will return, if present. The inspection confirms the verdict of the platform, that detected censorship and the actual technique *DNS injection* regardless of the misleading *symptoms* (no errors at any level of the stack - DNS, TCP, HTTP).

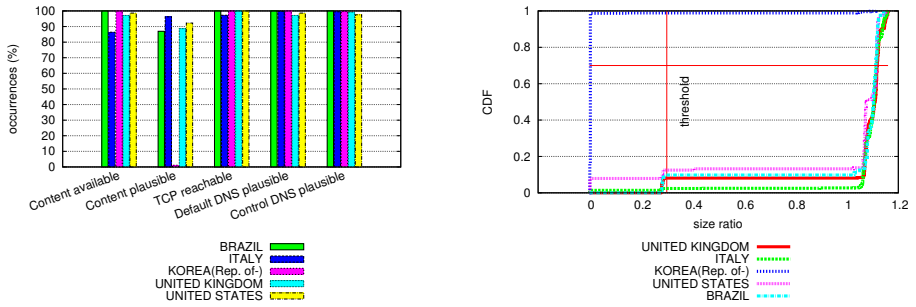
3.2 Censorship in Korea

The access to online content in South Korea is regulated by a government body, Korea Communications Standards Commission (KCSC) nominated by the president and in charge of the Ethics of Internet communications. The nation is reported by *ONI* as applying "selective filtering" for *Social* topics and "pervasive filtering" for the *Conflict/Security* category.

Adult websites. A category of websites that is forbidden per order of the Ethical authority is the one showing adult content (classified among "obscenity and perversion"). The detection algorithm has signaled censorship for URLs such as `hardsextube.com`, `pornhub.com` and `redtube.com`, coherently with the expectations. We will consider the case of `hardsextube.com` in detail, as the other presents analogous results.

Considering the summarized view for the different tested techniques aggregated by country (Fig. 4a), it becomes evident the peculiar response in Korea with respect to the other tested countries. More specifically, the "content plausible" percentage of tests, result of the analysis based on the size ratio of the downloaded resource, is near 0% while other countries show near 100%, thus limiting to Korea only the issue in accessing the original content. Also no other

censorship detection technique has been matched, thus excluding DNS Tampering and TCP-level filtering.



(a) Comparison of censorship evidences results across countries.

(b) Empirical CDF of the content size ratio.

Fig. 4: Censorship in Korea: porn websites.

To inspect in more detail the test that has detected censorship we refer to Fig. 4b, where the Empirical Cumulative Distribution Function is drawn of the ratio of each sample content size over the global average. It can be seen that only results for Korea (in dark blue, close to the top border of the graph) are almost completely (0.98%) below the detection threshold (empirically set to 0.3 as for the preceding analyses). All other countries have the almost totality of samples beyond 1.1, with the exception for U.K., U.S.A., and Brazil, with small fraction (less than 0.16) falling just short of the threshold.

These results have not raised a censorship verdict due to the small relative occurrence (pre-filtering data cleansing ignores cases that represent less than 70% of the results). We have manually checked the content and found that corresponds to mobile versions of the requested website. The detection algorithm based on the size ratio has proved robust to content adaptation [13] in this scenario, but further research should be pursued in order to generalize this result.

To validate the censorship verdict, we have manually inspected the returned resource.We have seen that the returned webpage, result of the *HTTP tampering* technique, consists of a single JavaScript section whose effect when interpreted by the browser is to redirect to the address `http://warning.or.kr`, the official block page of the Korean authority for Internet censorship.

3.3 Censorship in Italy

Internet censorship in Italy is enforced mainly against websites proposing online gaming, betting and copyright infringement. Another significant motivation for censorship is the block of child pornography, but due to ethical issues in potentially involving volunteers in police investigations the latter has not been tested. Thanks to UBICA we could see that no centralized censoring infrastructure is present, as censoring is detected for different ISPs starting and ending at differ-

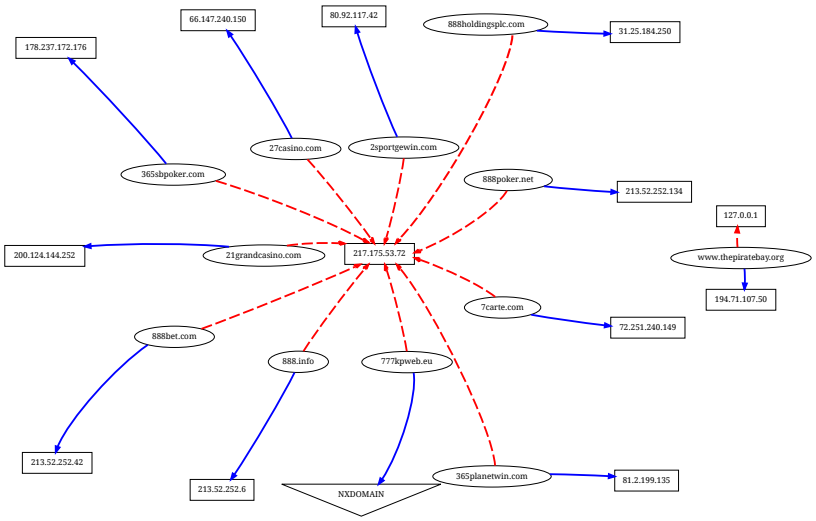


Fig. 5: DNS hijacking in Italy: DNS resolution graph for betting websites. Ellipses contain host names, rectangles contain IP addresses, arrow lines are resolutions requested by probes to their default resolver: for dashed lines probes are inside Italy, while for solid lines the probe is in USA.

ent times, and censoring techniques are sometimes different (in the vast majority DNS hijacking, and case-specific TCP blocking).

The Italian Agency for State Monopolies (AAMS) [4] provides an official list of domains[5] that have been blocked because of infringement of the Italian laws on online gaming and betting (both require a state license). Another -but non official - source is provided by an independent researcher in his "observatory on censorship" website[6] where a list of censored domains together with the authority that issued the censoring order and the date it was issued are reported.

In the case of blocks of websites proposing online gaming and betting the block is explicit (by means of a blockpage), while for websites related to *file sharing* the block is not motivated, resulting in a network error or a website describing a generic error. The censoring technique used most across all the tested ISPs is DNS hijacking, whose effect is graphically shown in Fig. 5 and in which DNS resolution requested to the probe default resolver is compared between probes from inside Italy (red lines) and USA (blue lines).

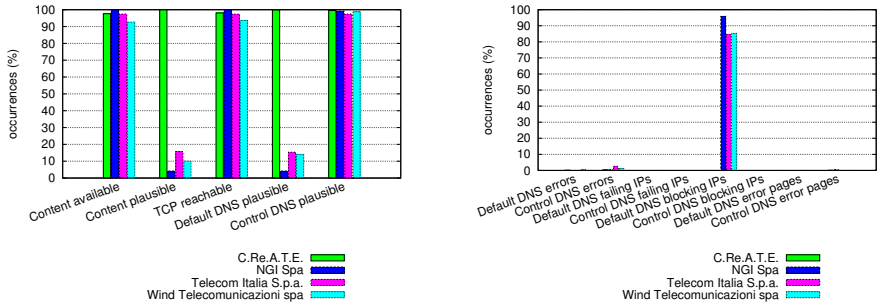
A few specific examples are described in the following.

Betting and gaming. The website `http://bet365.com` will be used as a representative of the *betting and gaming* website class. The results of censorship analysis algorithms for the resource `bet365.com` is reported in Fig. 6a. We can

[4] Amministrazione Autonoma dei Monopoli di Stato, `http://eee.aams.gov.it`

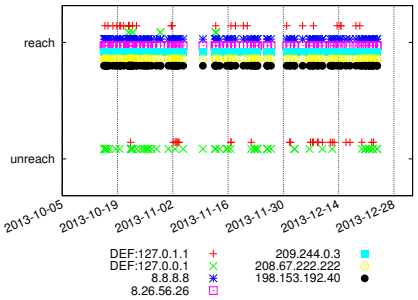
[5] `http://www.aams.gov.it/sites/aams2008/files/documenti_old/private/downloads/documentazione/scommesse/Elenco_siti_inibiti/elenco_siti_inibiti.rtf`

[6] `http://censura.bofh.it/elenchi.html`



(a) Comparison of censorship techniques per different ISPs.

(b) Detail of DNS analysis.



(c) DNS temporal analysis.

Fig. 6: Censorship in Italy: gaming and betting websites, the case of `bet365.com`.

see that for the ISP "NGI" the percentage of DNS resolutions performed by the probe default resolver is as little as 4.5%. This is reflected by an analogous percentage of content of plausible size. From the same graph it can be seen that also for "Wind Telecomunicazioni" and "Telecom Italia" providers there are low percentages of plausible DNS resolution (31.2% and 46.1% respectively) and similar percentages of plausible content size (23.8% and 46.1% respectively). Only for the "Center for REsearch And Telecommunication Experimentation" ISP, serviced by the *GARR* [7] , both the DNS resolutions and the downloaded content size are always plausible, showing no censorship on this network for the considered resource.

The verdict for the other ISPs is of *censorship by means of DNS hijacking towards an explicit blockpage*, in fact by comparing the result between the default DNS resolver and the control ones it can be noted that no control DNS is affected.

The reason for the specific kind of DNS hijacking (*blockpage*) is evident when inspecting the results of the DNS analysis, reported as a bar chart in Fig. 6b. Here we can see that for all the three ISPs implementing censorship, the resulting DNS response belongs to the list of known *blockpages*. Thus the adopted censoring technique has the effect of presenting the user with a block webpage explicitly telling him/her of the censorship. From the Fig. 6b it can also be noted that

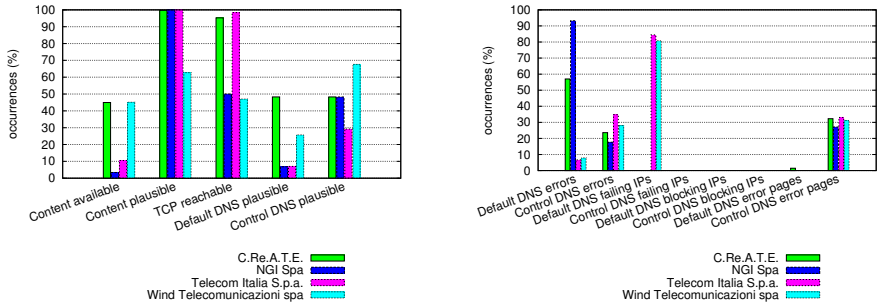
[7] The "GARR" is the Italian Academic and Research telecommunication network.

with the exception of "NGI Spa", with 95.4%, no ISP gives percentages close to the totality. The possible causes of this behavior can be: (i) a variability of the censor behavior in the analysis time interval (beginning or ending of censorship); (ii) heterogeneity of the probe environment at a granularity smaller than the *ISP* level. The temporal evolution of the case under description is shown in Fig. 6c. It can be seen that the oscillating results between reachability (upper line) and unreachability is limited to the *default* resolvers (the first two entries in the key, prepended with "DEF:"), while the *control* resolvers always report the domain as uncensored. It can be noted that the default DNS server address - as reported in the DNS reply - corresponds to `localhost`: a local caching application such as `dnsmasq`[8] is in function on the probe system, preventing the collection of the local default resolver.

Streaming and File Sharing. The second class of websites censored in Italy is constituted by repositories and index directories for file sharing and multimedia streaming. For this class of websites UBICA has reported a much more diverse scenario across the different ISPs; we will describe it in the following taking as an example the index directory `http://thepiratebay.sx`. The overall behavior of censorship techniques used by different Italian ISPs is summarized in Fig. 7a. Besides the low percentages of *plausible DNS* responses for the default resolver, low percentages are present also for *control* DNS servers. Moreover, differently from the case of betting websites, also the ISP connected through the Academic and Research network GARR presents low percentages (less than 50% for both default and control resolvers, and close to 40% of content availability). Another notable difference is in the result for TCP reachability: while for the online betting website this measure scored close to 100% reachability for 3 out of 4 ISP (and more than 75% for the remaining one), in the case of the file sharing website 2 ISPs show less than 50% reachability at the TCP level. A more in-depth inspection of the results of DNS tests, reported in Fig. 7b, shows a more diverse condition with respect to the case of betting websites (Fig. 6b).

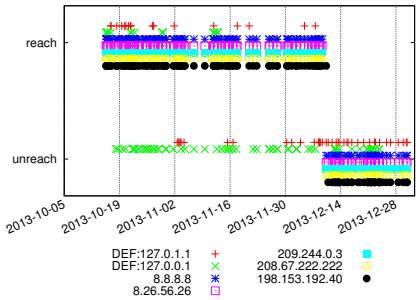
All the ISPs show different DNS errors, both for default and control DNS servers. One ISP ("Wind Telecomunicazioni") shows a 65.5% responses returning a *failing IP* (`127.0.0.1`) for the default resolver, and 7.7% of `NXDOMAIN` or `TIMEOUT` DNS errors. Different percentages of errors are shown by the other ISPs, each characterized by the presence of multiple symptoms of DNS unreachability in strong discordance with the case of betting websites (each ISP concentrated in one kind of DNS unreachability symptom). The temporal analysis of the DNS measures, represented in the time series of Fig. 7c, helps explaining such combination of results for the "Wind" ISP. In fact, similarly to the case of betting websites (Fig. 6c), there is an oscillation between reachability and unreachability for the default resolvers, spanned over the first half of the timeline, again explainable with the lack of control over the default DNS set for the probe.

[8] `dnsmasq` is an open source DNS cache and forwarder, installed by default on several distribution of Linux, including OpenWrt and Ubuntu, main OSes for the UBICA probes. Website: `http://www.thekelleys.org.uk/dnsmasq/doc.html`



(a) Comparison of results of different techniques.

(b) Detail of DNS analysis.



(c) DNS temporal analysis for the "Wind" ISP.

Fig. 7: Censorship in Italy: file sharing websites, the case of `thepiratebay.sx`.

In this case, however, all the resolvers, no matter if default or control, report unreachability. The unreachability of `thepiratebay.sx` starting from December 10th 2013 is verified by the probes in *all* the countries, signaling that a server-side event has occurred. From manual check of external information (the news section of the same website, freshly moved to another Top Level Domain: `http://thepiratebay.se/blog/234`) we can validate the finding of the UBICA platform: the old hostname has been dismissed on December 10th.

4 Conclusions

In this paper we have presented results obtained by means of UBICA (User-based Internet Censorship Analysis), a crowdsourced platform for Internet censorship monitoring. We have ran UBICA for several months on selected targets and we have found evidences of several censorship techniques, such as DNS tampering and content filtering. In this paper we have shown practical results from the following countries: Italy, Pakistan, and South Korea. In these countries we obtained help from local users (and we really thank them) and we validated our analysis using a ground truth built by manual inspection of evidences. We have shown how the UBICA architecture and its main features are able to run an integrated and multi-step analysis to provide a simple but effective dashboard

thanks to which censorship events are easily spotted and described also in their temporal evolution. Being based on crowdsourced data and on repeated measurements, the completeness and accuracy of the monitoring depend on user participation: to foster community participation we have provided a lightweight UBICA client for linux platforms and the online access to client reports, both available at `http://ubica.comics.unina.it`.

References

1. Herdict Project. `http://www.herdict.org`
2. Anderson, C.: Dimming the internet: Detecting throttling as a mechanism of censorship in iran. arXiv preprint arXiv:1306.4361 (2013)
3. Chaabane, A., Chen, T., Cunche, M., Decristofaro, E., Friedman, A., Kaafar, M.A., et al.: Censorship in the wild: Analyzing internet filtering in syria. In: ACM SIGCOMM IMC (2014)
4. Chun, B., Culler, D., Roscoe, T., Bavier, A., Peterson, L., Wawrzoniak, M., Bowman, M.: Planetlab: an overlay testbed for broad-coverage services. ACM SIGCOMM Computer Communication Review 33(3), 3–12 (2003)
5. Clayton, R., Murdoch, S., Watson, R.: Ignoring the great firewall of china. In: Privacy Enhancing Technologies, 2006
6. Cotton, M., Vegoda, L., Bonica, R., Haberman, B.: Special-Purpose IP Address Registries. RFC 6890 (Best Current Practice) (Apr 2013)
7. Dainotti, A., Squarcella, C., Aben, E., Claffy, K.C., Chiesa, M., Russo, M., Pescapé, A.: Analysis of country-wide internet outages caused by censorship. In: SIGCOMM. pp. 1–18. ACM (2011)
8. Dornseif, M.: Government mandated blocking of foreign web content (2003), `http://arxiv.org/abs/cs/0404005`
9. Eckersley, P.: How unique is your web browser? In: Privacy Enhancing Technologies, 2010
10. Filastò, A., Appelbaum, J.: Ooni: Open observatory of network interference. In: USENIX FOCI 2012
11. Jones, B., Lee, T.W., Feamster, N., Gill, P.: Automated detection and fingerprinting of censorship block pages. In: ACM SIGCOMM IMC (2014)
12. Khattak, S., Javed, M., Khayam, S.A., Uzmi, Z.A., Paxson, V.: A look at the consequences of internet censorship through an ISP lens. In: ACM SIGCOMM IMC (2014)
13. Md Fudzee, M.F., Abawajy, J.: A classification for content adaptation system. In: iiWAS. pp. 426–429. ACM (2008)
14. Nabi, Z.: The anatomy of web censorship in pakistan. In: USENIX FOCI 2013
15. Park, J.C., Crandall, J.R.: Empirical study of a national-scale distributed intrusion detection system: Backbone-level filtering of HTML responses in China. In: ICDCS. pp. 315–326. IEEE (2010)
16. Sfakianakis, A., Athanasopoulos, E., Ioannidis, S.: Censmon: A web censorship monitor. In: USENIX FOCI 2011
17. Sundaresan, S., De Donato, W., Feamster, N., Teixeira, R., Crawford, S., Pescapè, A.: Measuring home broadband performance. CACM 55(11), 100–109 (2012)
18. Verkamp, J.P., Gupta, M.: Inferring Mechanics of Web Censorship Around the World. In: FOCI. USENIX (2012)
19. Xu, X., Mao, Z.M., Halderman, J.A.: Internet censorship in china: Where does the filtering occur? In: Passive and Active Measurement. pp. 133–142. Springer (2011)