A Method for Identifying and Confirming the Use of URL Filtering Products for Censorship*

Jakub Dalek⋆, Bennett Haselton⋆, Helmi Noman⋆, Adam Senft⋆, Masashi Crete-Nishihata⋆, Phillipa Gill†⋆, Ronald J. Deibert⋆
⋆ The Citizen Lab, University of Toronto
† Dept. of Computer Science, Stony Brook University

Abstract

Products used for managing network traffic and restricting access to Web content represent a dual-use technology. While they were designed to improve performance and protect users from inappropriate content, these products are also used to censor the Web by authoritarian regimes around the globe. This dual use has not gone unnoticed, with Western governments placing restrictions on their export.

Our contribution is to present methods for identifying installations of URL filtering products and confirming their use for censorship. We first present a methodology for identifying externally visible installations of URL filtering products in ISPs around the globe. Further, we leverage the fact that many of these products accept user-submitted sites for blocking to confirm that a specific URL filtering product is being used for censorship. Using this method, we are able to confirm the use of McAfee SmartFilter in Saudi Arabia and the United Arab Emirates (UAE) and Netsweeper in Qatar, the UAE, and Yemen. Our results show that these products are being used to block a range of content, including oppositional political speech, religious discussion and gay and lesbian material, speech generally protected by international human rights norms.

Categories and Subject Descriptors: C.2.2 [Computer-Communication Networks]: Network Protocols

General Terms: Measurement **Keywords:** Censorship; Network Measurement; URL filtering

1. INTRODUCTION

URL filtering products, used for managing Web traffic and restricting access to content, are extremely common in corporate, educational and ISP networks around the globe. However, these technologies, which were designed to improve performance and filter inappropriate content in the enterprise setting, represent a dual-use technology. Indeed, there have been numerous reports of URL filtering products produced by Western companies being sold in countries with poor human rights records, where they are used for censorship and surveillance [2–5, 28, 30]. The use of technology developed in North America and Europe against citizens by authoritarian regimes presents many legal and ethical issues. In the past two years, the United States [26], Europe [29], and Israel [7] have all taken steps to limit the export of these technologies to countries under sanctions such as Syria and Iran. Further, as part of the OpenNet Initiative (ONI), we have documented numerous cases of products developed by Western companies being used to limit freedom of speech online via censorship and potentially surveillance [11, 13, 14, 21, 32] over the past ten years.

With the stakes so high, it is important that we have techniques for monitoring the use of specific technologies for censorship. These tools can help inform policymakers and even vendors, who may be unaware that their technology is being used for censorship. In 2009, Websense actually withdrew software update support once the ONI informed them that their technology was being used for censorship by the government of Yemen [35]. While our methodologies have sufficed thus far, recent legal implications of our observations beg for repeatable methodologies that produce high confidence results.

Challenges of measuring URL filtering deployments. Measuring URL filtering products is complicated by the fact that censorship is difficult to observe without vantage points located within the country of interest. Through our involvement in ONI, we have managed to gain access to measurements from within many countries with restrictive filtering regimes. However, performing client-based measurements in some countries is considered too risky (e.g., Cuba, North Korea), thus we cannot claim global coverage. Further, identifying installations of specific URL filtering products requires an understanding of distinct properties of the product under consideration (e.g., relevant HTTP headers) and careful validation to avoid false positives. Finally, even if a given product is installed in a country, it does not necessarily mean it is being used for censorship/surveillance purposes.

Our contribution. With these challenges in mind, we sought to design a simple, repeatable, methodology for *identifying* installations of URL filtering products and *confirming* their use for censorship. Our method for identifying URL filter installations (§3) serves to identify installations of filtering products where we can apply our confirmation methodology (§4); however, the confirmation methodology alone is enough to verify that a specific product is used for censorship in a given network. Our identification method-

*Data available at: `http://goo.gl/Cc1v0`

IMC'13, October 23–25, 2013, Barcelona, Spain.

http://dx.doi.org/10.1145/2504730.2504763 .

Table 1: Summary of products we consider.

Company	Headquarters	Product description	Previously observed
Blue Coat [1]	Sunnyvale, CA, USA	Web proxy (ProxySG) and URL Filter (Web Filter)	Burma, Egypt, Kuwait, Qatar, Saudi Arabia, Syria, UAE [14, 32]
McAfee SmartFilter [16]	Santa Clara, CA, USA	Filtering of Web content for enterprises	Bahrain, Iran, Kuwait, Oman, Saudi Arabia, Tunisia, UAE [11, 21, 22]
Netsweeper [18]	Guelph, ON, Canada	Netsweeper Content Filtering	Qatar, UAE, Yemen [21]
Websense [34]	San Diego, CA, USA	Web proxy gateways including features to monitor for corporate data leakage.	Yemen (prior to 2009) [21]

ology hinges on the observation that some URL filter installations are visible on the global Internet (likely due to inexperienced network administrators). We came to this observation during initial study of Syria where external facing IP addresses were used to host Blue Coat products [32]. We develop methodologies to locate a set of these externally visible IPs and verify that (1) they are hosting the suspected product and (2) confirm the product is indeed used for censorship (when we have access to in-country measurements). Our study highlights the human rights implications of these products and can provide ground truth for Web proxy fingerprinting (as is done by Netalyzr [12, 17]).

Key insights. Our study highlights many properties of URL filters that underscore the importance of studying their use. Specifically, we observe the use of these products in multiple North American ISPs which raises issues about monitoring how these products are being used on a global scale, and not just in jurisdictions commonly considered in censorship studies (§3). We further highlight complications in terms of how URL filters are implemented (e.g., inconsistent blocking and use of multiple products) that make characterizing their use challenging (§4). Finally, we demonstrate our proposed method by confirming the use of Netsweeper and McAfee SmartFilter for censorship of content protected by human rights norms in Qatar, Saudi Arabia, the UAE and Yemen (§5).

Limitations. We note that our methodology for identifying installations of URL filtering products requires that these installations be visible on the global Internet, thus the identification method is likely identifying installations that are not maintained by a technically sophisticated administrator. Further, we are not robust to products that attempt to evade our profiling. Thus, our results should be viewed as presenting a high confidence subset of URL filter deployments. We discuss these limitations in more detail in Section 6.

2. BACKGROUND

2.1 URL filtering products

There are multiple methods that may be used for implementing URL filtering systems. These products usually include a database of pre-categorized URLs, that allow the network operator to configure which categories to block within their network, and the ability to create custom categories for blocking. The products may also include a subscription/update component to push newly categorized URLs to the product's database. Depending on the functionality provided by the product (*e.g.,* whether it proxies all Web traffic), the product may be sold as software to be installed by the administrator or as a stand-alone middlebox in the case of proxies. Table 1 summarizes the products we consider.

2.2 Prior work by the OpenNet Initiative

The OpenNet Initiative has studied Internet censorship for the past decade [11, 13, 14, 21, 32]. Over the course of the project, we began to observe products developed by Western companies used for censorship by repressive regimes, starting with SmartFilter products in Tunisia in 2005. By 2010, we had documented extensive use of Western technologies for censorship in the Middle East and North Africa (MENA) region [21]. Table 1 highlights the products we have discovered, as well as the countries where we have located them. Initially, our methods for identifying these products consisted of manual analysis of block pages for company logos/branding and product names in HTTP headers. Over time, however, we observed Western vendors obscure the use of their products (*e.g.,* by omitting logos on block pages [21]). In response, we have been developing novel techniques to confirm use of URL filtering products. This paper expands upon these efforts by describing how we identify networks containing URL filter installations (§3) and how we confirm that these installations are in fact being used for censorship (§4).

The policy impacts of our efforts thus far, have been mixed. In 2009, our identification of Websense in Yemen led to the vendor discontinuing support of their product for the Yemen government [35]. In contrast, Netsweeper stated that is not against their company policies to aid foreign governments in implementing Internet censorship. Finally, we have observed Blue Coat withdraw update support from Syria [32], as a result of legal sanctions against the country [26]. However, the company still plays a role in Internet censorship in many countries around the world, and has even been named an "enemy of the Internet" by Reporters Without Borders [6]. Part of our goal in this study is to present a repeatable methodology for identifying and confirming the use of these products for censorship to inform future discussions with vendors and policy makers.

3. IDENTIFYING URL FILTERS

We begin by presenting our methodology to identify installations of URL filtering products. As described in Section 2.2, we previously leveraged user reports to identify URL filtering products. However, as vendors remove branding from block pages, it becomes more difficult for non-technical users to identify these products. Further, the individuals we engage with tend to be biased towards certain regions of interest (*e.g.,* the MENA region). In this section, we present an identification method that does not depend

Table 2: Summary of our methodology for identifying URL filtering products.

Product	Shodan keywords	WhatWeb signature
Blue Coat	"proxysg","cfru="	Built in detection or `Location` header contains hostname "www.cfauth.com"
McAfee SmartFilter	"mcafee web gateway", "url blocked"	`Via-Proxy` header or HTML title contains "McAfee Web Gateway"
Netsweeper	"netsweeper", "webadmin", "webadmin/","webadmin/deny", "8080/webadmin/"	Built in detection
Websense	"blockpage.cgi", "gateway websense"	`Location` header redirects to a host on port 15871 with parameter "ws-session"

on user reports and is more scalable than manual inspection, by examining HTTP headers and Web directory structure for evidence of filtering installations.

3.1 Methodology

Our methodology leverages the observation that URL filtering products are sometimes configured such that they are visible on the global Internet. Since they are visible on the global Internet, these products can be located via external scans of IP address space. Indeed, our group has previously leveraged network scans to identify Blue Coat installations in Syria [14, 32]. As a proof of concept, we demonstrate our techniques using the Shodan search engine [27] to locate IP addresses, but are working towards applying it on a larger scale with the Internet Census [10] data in ongoing work.

Locating potential installations. The Shodan search engine [27] indexes the IP addresses of externally visible devices on the Internet. Entries in Shodan consist of an IP address, along with meta-data and HTTP headers observed when the IP address was accessed by the search engine. By manually analyzing results from the ONI tests, we were able to identify commonly appearing keywords and headers for the products we consider (summarized in Table 2). These keywords include HTTP headers (e.g., "ProxySG" for Blue Coat) and paths known to be associated with the management console (e.g., "8080/webadmin" for Netsweeper). We search for these keywords, in combination with each of the two letter country-code top-level domains, to maximize the set of results we obtain from Shodan.

Validating URL filter installations. When locating IP addresses of the URL filters, we are not conservative, and rely on the following step to confirm that a given product is indeed installed on the identified host. We use the WhatWeb profiling tool [9] to confirm the product that is installed on a given host. For some products (e.g., Netsweeper) WhatWeb contains a pre-existing signature that we leverage in our validation, whereas in other cases we create signatures based on HTTP headers observed when running the WhatWeb tool on an IP address. Table 2 also summarizes how we identify the various products using WhatWeb.

Finally, we use geolocation data from MaxMind [15] and whois data from TeamCymru [31] to map the IP addresses matching WhatWeb signatures to country-level location and autonomous system (AS) number.

3.2 Networks with URL filtering installations

Figure 1 summarizes the countries where each product was observed. While ONI tends to focus on installations by national ISPs, our new methodology allows us to uncover URL filtering installations in a wider set of countries and networks. Indeed, we observe Blue Coat in many new countries and regions: South America (Argentina and Chile), Europe (Finland, Sweden), Asia (Philippines, Thailand and Taiwan) and the Middle East (Israel, Lebanon). Further, for the remaining three products all the installations we discover (with the exception of McAfee SmartFilter in Pakistan) were previously unknown.

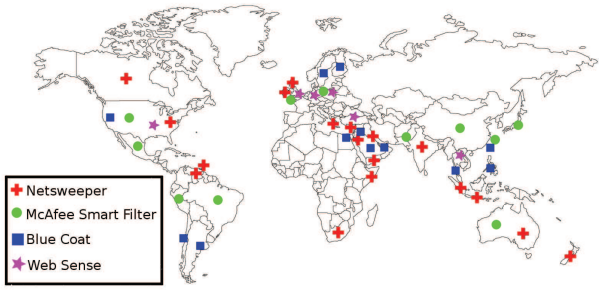

Figure 1: Locations of URL filter installations

As expected, we observe installations on a diverse range of networks in the US, such as Websense in two Texas utilities' networks and Netsweeper installations in educational networks in West Virginia, Oklahoma and Missouri. However, we also observe Netsweeper installations in large ISP networks such as Global Crossing, AT&T, Verizon, and Bell South; and Blue Coat installations in Comcast and Sprint. Interestingly, we also observe an installation of Blue Coat on an IP address registered to the United States Information Systems Command (USAISC). The dual-use of these products for network management and censorship, requires confirming *how* they are used before drawing conclusions.

4. CONFIRMING USE OF URL FILTERS

Many of the products we identify play a legitimate role in network management. Thus, when evaluating the human rights implications of these technologies, it is important to validate that they are actually being used for censorship. Further, vendors may obscure the identities of their products by removing the headers we identified in Table 2. The confirmation method we present, is robust to a lack of signatures and does not require the IP address of the URL filter be externally visible. However, we use networks identified via the techniques in Section 3 as a case study of our technique.

4.1 In-network testing

To confirm that a URL filter is being used for censorship, we perform experiments from within the network under consideration, using our global network of testers. Tests of Web page accessibility are performed using a measurement client

that accesses a specified list of URLs in the "field" i.e., the location where censorship is suspected. This client software also triggers the same set of URLs to be accessed from a server in our lab at the University of Toronto (which does not censor the type of content tested). The results of the Web page accesses in the field and lab are compared to determine if the page was blocked in the field location. For our measurements of URL filtering, we test short lists of URLs that are amenable to manual analysis of results. Further, the products we test tend to use block pages that explicitly state that content has been censored. Thus, we avoid ambiguities such as censorship via dropped packets or TCP resets.

4.2 Methodology

Our methodology seeks to answer the question: is the given URL filtering product used for censorship in the measured ISP? Since many URL filters provide a mechanism for users to submit sites that should be blocked, we wondered if we could use this mechanism to confirm the use of a specific URL filter. The basic idea is to test sites (under our control) that are not blocked within the ISP, and then submit a subset of these sites to the appropriate URL filter vendor. After 3-5 days, we retest the sites and observe whether or not the submitted sites are blocked. If they are blocked, it is highly likely that the URL filter under consideration is being used for censorship, and our submission of the sites triggered the blocking. We present case studies that explore the effectiveness of this idea to confirm the use of a variety of products in networks where we have in-country testers. Table 3 summarizes these case studies.

4.3 Case study: McAfee SmartFilter in UAE and Saudi Arabia

In 2009, the ONI identified McAfee SmartFilter being used in UAE's national ISP, Etisalat, and in a centralized blocking implementation in Saudi Arabia (effectively used for all ISPs) [23, 24]. We use our proposed methodology to confirm whether these technologies are still deployed in these networks in 2012 and 2013.

ONI had previously observed Etisalat using SmartFilter to block content related to anonymizing proxies [24]. Thus, we created a set of 10 domains providing proxy services to test whether SmartFilter was still being used. These domains had the form of two random (non-profane) words registered with the ".info" top-level domain (e.g., `starwasher.info`) and contained the Glype proxy script [8] as their index page. We first ran measurements in the country to verify that these 10 domains were accessible. We then submitted five of these domains to SmartFilter for blocking. Within a few days we observed that the five submitted sites were blocked on Etisalat, thus confirming that the product was still in use within the country (Table 3).

Challenge 1: Access to sites that will be blocked. Our methodology requires access to Web sites that are blocked by the studied ISP. Unlike UAE, we found Web sites classified as proxies by SmartFilter were accessible in Saudi Arabia. Thus it appears that Saudi Arabia is not using the proxy category provided by SmartFilter in their deployment. However, Web sites classified as pornography by SmartFilter are blocked in Saudi Arabia. Thus, we perform a similar experiment as we did for UAE, except that the 10 created domains hosted an adult image found via a Google image search. (The image was only used for the duration of our experiment and then removed.) Using the ISP Bayanat Al-Oula, we verified that the 10 domains were accessible in Saudi Arabia. We then submitted five of the domains to SmartFilter for blocking. After four days, we observed that these five domains were blocked (Table 3). We repeated this methodology on Nournet, also in Saudi Arabia, and Etisalat to confirm SmartFilter is still used within these ISPs in 2013.

4.4 Case study: Netsweeper in Qatar, UAE and Yemen

Implementation details of censorship platforms can impact the ability of our method to confirm censorship. For example, we have observed Netsweeper queuing Web sites for categorization once they have been accessed within the country (to expand the set of categorized sites [19]). As a result, once we have validated that our set of URLs is accessible, they may be queued for categorization by Netsweeper, and eventually may be blocked. Thus, it is not possible for us to validate that our sites are accessible prior to submitting a subset of them to be blocked. As a result, we operate on the assumption that none of our sites will be blocked prior to submission.

Prior study by the ONI identified the use of Netsweeper in YemenNet in Yemen [25], Du, in UAE [24] and Ooredoo, in Qatar [21] to block content related to anonymizing proxies. Thus, we use these ISPs to test our proposed methodology. We created a list of 12 domains providing proxy services and submitted six of them to Netsweeper's "test-a-site" service for classification [20]. We then accessed these 12 domains in YemenNet, Du and Ooredoo, and observed whether the six submitted sites were blocked. In all three ISPs, the methodology was successful with 5 of the 6 sites blocked in Du and all six blocked in YemenNet and Ooredoo. (Table 3).

Challenge 2: Inconsistent blocking. Validating censorship in Yemen was complicated by inconsistent blocking. We observed cases where the blocking technology appeared to be temporarily "offline" within the country. For example, some proxy URLs are accessible on runs where other proxy URLs are blocked, while in later runs the reverse is true for the same set of URLs. Indeed, prior work by the ONI observed a Yemeni ISP using Websense with a limited number of concurrent user licenses. When the number of users exceeded the number of licenses no content would be filtered [25].

Inconsistent blocking means that we need to repeat the tests numerous times and require a larger set of domains for testing (as we cannot be sure that previously-accessed sites are not queued for classification). This inconsistency limits the scalability of our approach for validating Netsweeper installations in Yemen.

However, we identified another way to validate that Netsweeper is being used for censorship. To help with configuration of the middlebox, Netsweeper provides a Web site for operators to validate that censorship is working within their network by querying a set of 66 category-specific URLs (e.g., `denypagetests.netsweeper.com/category/catno/23` for pornography). While this method is only viable in networks where the tool has not been disabled, a manual test of this tool in YemenNet, in January 2013, indicated that five categories were blocked: adult images, phishing, pornography, proxy anonymizers, and search keywords.

Table 3: Summary of URL filter case studies.

Product	Country	ISP	Date	Sites submitted	Category	Sites blocked	Confirmed?
Blue Coat	UAE	Etisalat (AS 5384)	4/2013	3/6	Proxy Avoidance	0/3	N
Blue Coat	Qatar	Ooredoo (AS 42298)	4/2013	3/6	Proxy Avoidance	0/3	N
McAfee SmartFilter	Qatar	Ooredoo (AS 42298)	4/2013	5/10	Pornography	0/5	N
McAfee SmartFilter	Saudi Arabia	Bayanat Al-Oula (AS 48237)	9/2012	5/10	Pornography	5/5	Y
McAfee SmartFilter	Saudi Arabia	Nournet (AS 29684)	5/2013	5/10	Pornography	5/5	Y
McAfee SmartFilter	UAE	Etisalat (AS 5384)	9/2012	5/10	Anonymizers	5/5	Y
McAfee SmartFilter	UAE	Etisalat (AS 5384)	4/2013	5/10	Pornography	5/5	Y
Netsweeper	Qatar	Ooredoo (AS 42298)	8/2013	6/12	Proxy anonymizer	6/6	Y
Netsweeper	UAE	Du (AS 15802)	3/2013	6/12	Proxy anonymizer	5/6	Y
Netsweeper	Yemen	YemenNet (AS 12486)	3/2013	6/12	Proxy anonymizer	6/6	Y

Table 4: Summary of Web content blocked by URL filtering products.

Category		Media Freedom	Human Rights	Political Reform	LGBT	Religious Criticism	Minority Groups and Religions
Product	Where						
McAfee SmartFilter	UAE (AS 5384)	X	X	X	X	X	X
Netsweeper	Yemen (AS 12486)	X			X		X
Netsweeper	UAE (AS 15802)	X	X	X	X	X	X
Netsweeper	Qatar (AS 42298)	X	X	X	X	X	X

4.5 Case study: Blue Coat in UAE

Challenge 3: URL filtering and network management tools used in combination. Some of the products we consider in this study can be used in tandem to achieve the goals of an ISP. For example, software products such as SmartFilter can be configured to run on proxy appliances, such as Blue Coat's ProxySG. We observe this situation in Etisalat in the UAE, where we confirm that the ISP is using SmartFilter for URL filtering (§4.3). However, we also identified installations of Blue Coat products in Etisalat using our methodology in Section 3. We created a set of test URLs and submitted them to the "Proxy avoidance" category for filtering by Blue Coat (since we had previously observed proxies blocked using SmartFilter). Upon testing these URLs in Etisalat, we found that none of them were blocked. While it may be the case that the network is using the two vendors to block different types of content, the more likely cause of this discrepancy is Etisalat using SmartFilter for URL filtering atop a Blue Coat proxy which can provide further traffic management capabilities (e.g., policy enforcement, traffic shaping).

4.6 Ethical considerations

The case studies in the prior sections raise ethical issues concerning use of the interfaces provided by the vendors to submit Web sites. We emphasize that our approach does not harm the intended performance of the URL filter as we do not impact the classification of legitimate Web sites. Further, since the Web sites we submit are under our control, there is no collateral damage to existing sites. Finally, in the use of a pornographic image in Saudi Arabia we took care to remove the image promptly after our tests were done. To limit the testers' exposure to the pornographic content, we had them access a benign image file located on the host, rather than the page containing the offensive content. Since our tests indicate that even the benign content on the host was blocked, we conclude that blocking was at the granularity of hostname, thus this method of mitigating user risk does not impact the results.

5. CHARACTERIZING CENSORED CONTENT

Now that we have designed a methodology to identify and confirm the usage of URL filters, we consider the type of content these products are blocking. The types of content found blocked by URL filters was determined by querying lists of URLs through the measurement client (§4.1). Two lists of URLs were tested in each country; a "global list" of internationally relevant content which is constant for all countries, and a "local list" of locally relevant content which is designed for each country by regional experts and is unique for each country tested. Each of the URLs on these lists was assigned to one of 40 content categories (e.g. "human rights" or "gambling") under four general themes: political, social, Internet tools and conflict/security content.

Tests using the measurement client to characterize the type of content censored by URL filtering products were performed within 30 days of the confirmations in Section 4. Manual analysis identified regular expressions corresponding to the vendors' block pages and automated analysis identified all URLs which matched a given block page regular expression.

Results of these tests, presented in Table 4, show that all products are used to block a wide variety of content, including oppositional and critical political content, non-pornographic gay and lesbian content, human rights content, independent media, as well as content relating to minority groups and religious discussion. The blocking of such content contradicts internationally recognized human rights frameworks protecting freedom of expression, such as Article 19 of the Universal Declaration of Human Rights [33].

Table 5: Summary of methods presented in this paper, their limitations, and potential techniques vendors may use to evade them.

Step	Technique	Limitations	Evasionary tactics
Identify installations (§3.1)	Port scans (*e.g.,* Shodan [27], Internet Census [10])	Can only identify externally visible installations.	Do not allow device to be accessed externally
Validate installations (§3.1)	WhatWeb [9]	Requires distinctive use of protocol headers	Can remove evidence of product from headers.
Confirm censorship (§4)	In-country testing and URL submission	Requires in-country testers, knowledge of what categories are blocked, and a set of domains for submission.	Vendors may identify and disregard our submissions (non-trivial)

6. DISCUSSION

In this paper, we present an initial methodology for identifying and confirming the use of URL filtering products around the world and highlight challenges faced when applying this methodology. We now elaborate on how our techniques would fare in the face of vendors that attempt to mask the use of their products. Table 5 summarizes these limitations. We emphasize that the identification of products and confirmation that they are used for censorship are independent; thus the confirmation (§4) is robust even if the techniques presented in Section 3 are evaded.

6.1 Identifying and validating installations.

While our identification method serves as a useful filter when determining where to apply the techniques of Section 4, we observe that URL filter vendors can take simple steps to evade discovery. To prevent identification of their products, vendors could provide ISPs with scripts and configuration instructions that prevent the product from being visible on the global Internet, however they are still reliant on the ISP to implement this correctly. URL vendors may also take steps to remove evidence of their products from protocol headers which is fairly simple to do, but would require having all ISPs running the product perform a software update before the change would take effect.

If both of these tactics are implemented, we would become more reliant on local contacts to provide reports of products being used (*e.g.,* because of access to internal IP address space) in their region. Alternatively, we could apply the techniques of Section 4 more widely, but scalability issues would make this time consuming.

6.2 Confirming censorship

URL filtering products view their database of URLs as a key differentiator to their business, and many even advertise the number of URLs they have classified and the rate at which they add to their databases [19]. By allowing individuals/administrators to submit sites to be blocked in different categories, they effectively crowdsource the database maintenance process. As a result of this, vendors are unlikely to reject all submissions based on our approach.

However, they may attempt to identify our submissions and disregard them. The can be accomplished by identifying either (1) our IP or e-mail address when we submit URLs, or (2) hosting services we use to host the domains under our control (§4). The first is easy for us to evade using proxy services or Tor and many e-mail addresses from free Webmail providers. The second can be evaded by using a popular cloud or hosting provider for our URLs, thus making blocking URLs from the given provider too damaging to the vendor's database.

Even though we may be able to counter evasion tactics by vendors, such a cat-and-mouse game is undesirable. Thus, the design of more scalable techniques, that can use our methodology to provide ground truth, is critical for continued study of URL filtering technologies.

7. CONCLUSIONS

We have presented a methodology for identifying installations of specific URL filtering products and confirmed their use for censorship in networks with in-country testers. Through our case studies, we have confirmed the use of North American products to block content protected by human rights norms in Qatar, Saudi Arabia, UAE, and Yemen.

Future work. While our methodology moves us beyond manual analysis of these products, it still poses many challenges in terms of scalability. Specifically, the methods in Section 4 require that we identify which categories are blocked in each ISP before creating test sites. Further, these methods also require vantage points in the network being considered. We hope this paper spurs dialog in the network measurement community about how to characterize URL filtering products in a high confidence, yet scalable, way. Indeed, our methodology can provide a useful ground truth for more general identification of transparent proxies (*e.g.,* [12, 17]) to yield a more complete picture of URL filtering deployments.

8. REFERENCES

[1] Blue Coat. `http://www.bluecoat.com/`.

[2] B. Elgin. Israel didn't know high-tech gear was sent to Iran, 2011. `http://www.bloomberg.com/news/2012-02-15/syria-blocks-texts-with-dublin-made-gear.html`.

[3] B. Elgin and V. Silver. Syria crackdown gets Italy firm's aid with U.S.-Europe spy gear, 2011. `http://www.bloomberg.com/news/2011-11-03/syria-crackdown-gets-italy-firm-s-aid-with-u-s-europe-spy-gear.html`.

[4] B. Elgin and V. Silver. Syria disrupts text messaging of protesters with made-in-Dublin equipment, 2012. `http://www.bloomberg.com/news/2012-02-15/syria-blocks-texts-with-dublin-made-gear.html`.

[5] B. Elgin, V. Silver, and A. Katz. Iranian police seizing dissidents get aid of Western companies, 2011. `http://www.bloomberg.com/news/2011-10-`

31/iranian-police-seizing-dissidents-get-aid-of-western-companies.html.

[6] Era of the digital merceneries. *Reporters Without Borders*, 2013. http://surveillance.rsf.org/en/.

[7] J. Ferziger. Israeli lawmaker calls for investigation of Iran equipment sales, 2011. http://www.bloomberg.com/news/2011-12-23/israeli-lawmaker-calls-for-investigation-of-iran-equipment-sales.html.

[8] Glype proxy script. http://www.glype.com/.

[9] A. Horton. WhatWeb, 2011. http://www.morningstarsecurity.com/research/whatweb.

[10] Internet census 2012: Port scanning /0 using insecure embedded devices, 2012. http://internetcensus2012.bitbucket.org/paper.html.

[11] ONI research profile: Tunisia. http://opennet.net/studies/tunisia, 2005.

[12] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. Netalyzr: Illuminating the edge network. In *The Internet Measurement Conference*, 2010.

[13] M. Marquis-Boire. From Bahrain with love: FinFisher's spy kit exposed? *The Citizen Lab*, 2011. https://citizenlab.org/wp-content/uploads/2012/08/09-2012-frombahrainwithlove.pdf.

[14] M. Marquis-Boire, J. Dalek, S. McKune, M. Carrieri, M. Crete-Nishihata, R. Deibert, S. Khan, H. Noman, J. Scott-Railton, and G. Wiseman. Planet Blue Coat: Mapping global censorship and surveillance tools. *The Citizen Lab*, 2013. https://citizenlab.org/wp-content/uploads/2013/01/Planet-Blue-Coat.pdf.

[15] Maxmind. http://www.maxmind.com/.

[16] McAfee Web protection. http://www.mcafee.com/us/products/web-protection.aspx.

[17] The ICSI Netalyzr. http://netalyzr.icsi.berkeley.edu/.

[18] Netsweeper content filtering. http://www.netsweeper.com/.

[19] Netsweeper by the numbers. http://www.netsweeper.com/what-we-do/netsweeper-by-the-numbers.

[20] Netsweeper Test-a-site. http://www.netsweeper.com/support/test-a-site.

[21] H. Noman and J. York. West censoring East: The use of Western technologies by Middle East censors 2010-2011. *The OpenNet Initiative Bulletin*, 2011. https://opennet.net/sites/opennet.net/files/ONI_WestCensoringEast.pdf.

[22] ONI research profile: Iran. http://opennet.net/research/profiles/iran, 2009.

[23] ONI research profile: Saudi arabia. http://opennet.net/research/profiles/saudi-arabia, 2009.

[24] ONI research profile: United arab emirates. http://opennet.net/research/profiles/united-arab-emirates, 2009.

[25] ONI research profile: Yemen. http://opennet.net/research/profiles/yemen, 2009.

[26] R. Runningen. Obama moves to block technology used by regimes on rebels, 2012. http://www.bloomberg.com/news/2012-04-23/obama-moves-to-block-technology-used-by-regimes-against-protests.html.

[27] Shodan - Computer Search Engine. http://www.shodanhq.com/.

[28] V. Silver. HP Computers underpin Syria surveillance, 2011. http://www.bloomberg.com/news/2011-11-18/hewlett-packard-computers-underpin-syria-electonic-surveillance-project.html.

[29] V. Silver. European Union bans exports to Syria of systems for monitoring web, phones, 2013. http://www.bloomberg.com/news/2011-12-01/european-union-bans-exports-to-syria-of-systems-for-monitoring-web-phones.html.

[30] V. Silver and B. Elgin. Torture in Bahrain becomes routine with help from Nokia Siemens, 2011. http://www.bloomberg.com/news/2011-08-22/torture-in-bahrain-becomes-routine-with-help-from-nokia-siemens-networking.html.

[31] IP to ASN Mapping, 2013. http://www.team-cymru.org/Services/ip-to-asn.html.

[32] Behind Blue Coat: Investigations of commercial filtering in Syria and Burma. *The Citizen Lab*, 2011. https://citizenlab.org/wp-content/uploads/2012/07/01-2011-behindbluecoat.pdf.

[33] Universal declaration of human rights: Article 19, 2013. http://www.un.org/en/documents/udhr/index.shtml#a19.

[34] Websense. http://www.websense.com/.

[35] J. York. Websense bars Yemen's government from further software updates, 2009. https://opennet.net/blog/2009/08/websense-bars-yemens-government-further-software-updates.