Censorship and Co-option of the Internet Infrastructure

Michael Bailey
Electrical Engineering and Computer Science Department
University of Michigan, Ann Arbor, MI 48109
mibailey@eecs.umich.edu

Craig Labovitz
DeepField Networks
Ann Arbor, MI 48104
labovit@umich.edu

Abstract

Over a few short years, the Internet has grown to play an integral part of daily economic, social, and political life in most countries. From the Egyptian "Velvet Revolution" to the last US presidential campaign, Internet communication shapes public opinion and fosters social change. But despite its immense social importance, the Internet has proven remarkably susceptible to disruption and manipulation, including government induced multi-week outages (e.g. Libya and Egypt) and multi-year campaigns by autocratic regimes to render web sites and large address blocks unreachable. While parents, enterprises, and governments have always placed restrictions on end-user communication to meet social or legal goals we argue recent years have seen the beginning of a new trend—the co-option of the Internet infrastructure itself to affect large-scale censorship. In this paper, we use Internet routing, allocation, and backbone traffic statistics to explore several recent and ongoing infrastructure-based efforts to disrupt Internet communication. We focus our discussion on the risks of this infrastructure corruption trend to long-term evolution of the Internet.

1 Introduction

The Internet plays a critical role in the economic, political, and social fabrics of global society. The current global Internet has roughly 1.7 billion users [26], fosters an estimated $1.5 trillion in annual global economic benefits [6], and is widely agreed to offer a staggering array of societal benefits, from improving the efficiency of our institutions [4] to enhancing our individual intelligence and decision-making [3]. As its growth continues (interdomain traffic has an annual growth rate of 44.5% [22]), the Internet has proven susceptible to emerging patterns of overt and more subtle disruption such as the loss of nearly all of Egyptian Internet traffic around this year's election [20] and persistent concerns over the "great firewall of China" (GFC) [28].

The primary motivation for the co-option of Internet infrastructure to effect filtering is political. Whether the filtering of content in China [28] and Iran [16], or the wholesale blocking of traffic in Burma or in Egypt [20], this form of censorship seeks to influence the spread of ideas and limits communication with the global community. Still other motivations call for meeting compliance with moral standards or laws as child porn filters in the UK [8] or the blocking of nazi promoting materials in Germany [12]. Yet a third such motivation, is filtering to meet engineering or commercial goals to gain economic advantage, realize profits, or assure the availability of a resource such as blocking or limiting P2P [10] or Skype.

A variety of techniques are employed to implement these goals (e.g., filtering URLs or packet filtering) and numerous have been studied in depth [8, 12, 28, 9]. While many of these edge-based techniques have been around since the early days of the Internet (e.g., NetNanny), we focus on a new trend in large-scale censorship—co-opting the core Internet infrastructure. In this form of censorship, weaknesses in underlying routing, naming, and transport protocols are employed to perform a censorship by blocking specific CIDR addresses or ASNs, blocking specific destinations by name, or violating the confidentiality or integrity of end-to-end communication. Whether motivated by the need for scalable (e.g., manageable, low-cost) filtering, robustness (i.e., censorship at multiple lay-

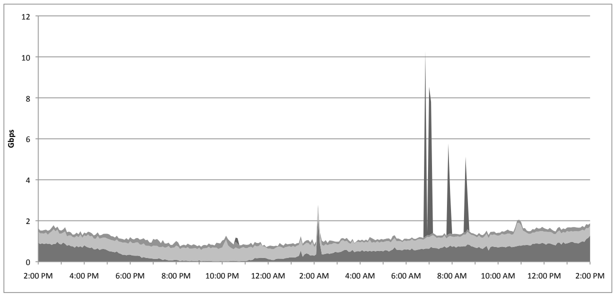

Figure 1: The second day of "hacktivists" attempting censorship of the Wikileaks sites through Distributed Denial of Services (DDoS) attacks.

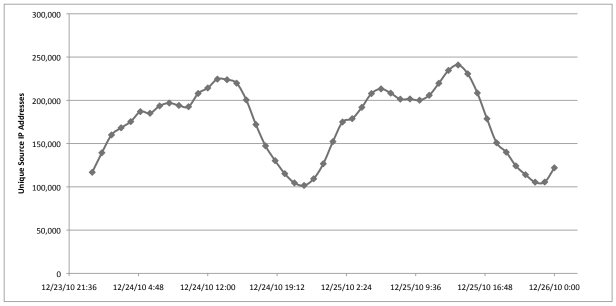

Figure 2: Two day's worth of TCP port 80 traffic to 37.61.54.0/24 representing unique source IP addresses attempting to contact facebook with bogus names supplied by the GFC.

ers), or deceit (i.e., who is censoring me?), such censorship by rogue autonomous systems (AS) poses significant challenges for both detection and mitigation.

2 Recent Events

In this section we discuss several recent incidents of large-scale Internet traffic disruption to gain insight into both the technical basis (e.g., DNS manipulation, vendor cooperation, BGP route manipulation, router filters blocking all traffic, payload and web filters) and impact of the events.

2.1 Wikileaks DDoS

In the second round of what turned out to be a protracted Internet skirmish [17], a denial of service attack briefly blocked access to the cablegate.wikileaks.org web site around 8:00 am EST, Tuesday, November 30th, 2010 [18]. On twitter, Wikileaks pegged the DDoS as exceeding 10 Gbps (significantly larger than our 2-4 Gbps estimate for the first round of attacks [19]). As compared with the initial attack, data from commercial traffic monitors [22] around the world suggest the second round of attacks was both larger and more sophisticated. Specifically, this attack involved several different components, including a low bandwidth application level DDoS and a 2-3 Gbps Syn attack against the primary "cablegate" IP addresses (the hosted web site is load balanced across data center locations in Europe and the US West Coast). Figure 1 graphs traffic from 110 ATLAS carriers around the world to address blocks (BGP prefixes) used by Wikileaks. The attack began around 7am EST though a smaller traffic spike occurs around 2am. Based on Netcraft and other reports, the outage was brief though cablegate web site performance was moderately impacted throughout the day. Interestingly, the attack appears to originate from a relatively small number of source IPs, including machines in Russia, eastern Europe and Thailand.

Ultimately, we argue the DDoS attacks surrounding Wikileaks supporters and opponents falls far short of a "cyberwar". While it makes a far less dramatic headline, cyber-vandalism may be a more apt description. This is not to say DDoS is not a serious problem. The number and firepower of botnets grows dramatically each year as well as the sophistication of application attack toolsets. Succeeding generations of volunteer botnet controlled PCs may evolve to pose a significant Internet-wide threat. However, traditionally the DDoS threat has come more from increasingly professional criminal hackers than volunteer activists.

2.2 China Facebook Filtering

Clayton et el. [9] discuss well known content blocking systems including packet dropping, DNS poisoning, and content inspection and discuss these methods in the context of the Great Firewall of China (GFC). While much of that work focused on TCP connection disruption tech-

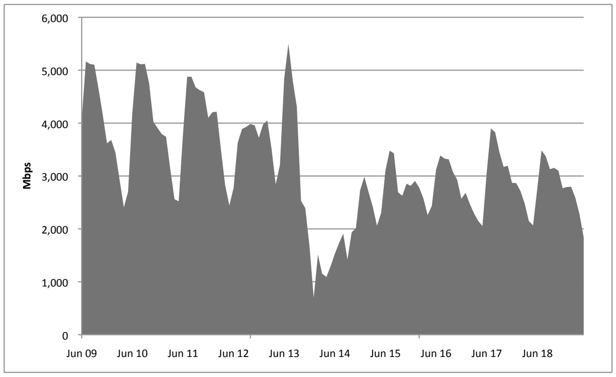

Figure 3: One day after the elections on June 13th at 1:30pm GMT Iran dropped off the Internet. Within a few hours, a trickle of traffic returned and as of 6:30am GMT June 16, traffic levels returned to roughly 70% of normal.

niques, Lowe et el. [24] explored the DNS filtering and modification in the GFC. More recently, Brown et el, noted that DNS-based censorship techniques leaked information about their operation [7]. In particular, they showed that some of the chinese nodes that make up part of I-root returned seemingly bogus IP addresses for a variety of sites being censored in China (i.e., youtube, facebook, twitter). For example,

```
dig @dns1.chinatelecom.com.cn. www.facebook.com.
   ...
   www.facebook.com.       11556   IN      A       37.61.54.158
   www.facebook.com.       24055   IN      A       78.16.49.15
   www.facebook.com.       38730   IN      A       203.98.7.65
```

Figure 2 illustrates the pervasiveness of the technique. During a continuing study of unallocated, but reachable IP addresses, first introduced in Wustrow et el. [27], we caught connection attempts to these bogus addresses. 200,000 such IP addresses, nearly 100% of whom are from within china, attempt to contact these bogus Facebook IPs every hour.

Application	Average percentage decrease
SSH	84.05
Flash	82.23
Bittorrent	82.06
POP	73.6
Alternative Web Ports	70.22

Table 1: The average percentage decrease in application traffic in the days before and after the election. Iranian firewalls appear to be selectively impacting application traffic.

2.3 Iran

In 2009 we investigated the drop in Iranian Internet traffic surrounding the time of the Iranian elections [16]. The state owned Data communication Company of Iran (or DCI) acts as the gateway for all Internet traffic entering or leaving the country. Historically, Iranian Internet access has enjoyed some level of freedom despite government filtering and monitoring of web sites. Iran normally sees around 5 Gbps of traffic, through a reported capacity of 12 Gbps from through 6 upstream regional and global Internet providers, and demonstrates typical diurnal and weekly curves (though Iran sees dips both on Iranian weekend of Thursday / Friday as well as during western Saturday / Sunday weekends). One the day after the elections on June 13th at 1:30pm GMT (9:30am EDT and 6:00pm Tehran / IRDT), Iran dropped off the Internet. All six regional and global providers connecting Iran to the rest of the world saw a near complete loss of traffic. Initially, DCI severed most of the major transit connections into Iran. Within a few hours, a trickle of traffic returned and as of 6:30am GMT June 16, traffic levels returned to roughly 70% of normal with Reliance traffic climbing by more than a Gigabit. Figure 3 shows this event and surrounding traffic from the perspective of the Arbor Internet observatory.

We can only speculate on proximate cause of this significant reduction in traffic volumes. DCI's Internet changes suggest piecemeal migration of traffic flows. Typically off the shelf / inexpensive Internet proxy and filtering appliances can support 1 Gbps or lower. If DCI needed to support higher throughput (say, all Iranian Internet traffic), then redirecting subsets of traffic as the filtering infrastructure comes online would make sense. Unlike Burma, Iran has significant commercial and technological relationships with the rest of the world. In other words, the government cannot turn off the Internet without impacting business and perhaps generating further social unrest. In all, this represents a delicate balance for the Iranian government and a test case for the Internet to impact democratic change. In analysis of the top applica-

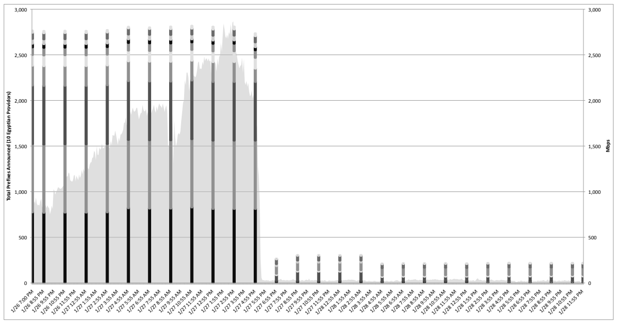

Figure 4: At 5:20 pm EST on January 27th, 2011 traffic to Egypt world drops precipitously and the majority of routes are withdrawn.

tions now blocked by the DCI firewall(s) [14], we showed the Iranian firewalls appear to be selectively impacting application traffic. Examining Table 1 it is interesting to note that ssh (a secure communication protocol) tops the list followed by video streaming and file sharing.

2.4 Egypt

Following a week of growing protests and periodic telecommunication disruption, we have shown that [20, 21] Egypt suddenly lost all Internet connectivity at approximately 5:20pm EST Thursday January 27, 2011. Figure 4 shows traffic to and from Egypt based on commercial probe statistics [22]. Between 3 and 5pm EST, Egyptian traffic rapidly climbed to several Gigabits. At 5:20pm, the all Egyptian transit providers abruptly withdrew the majority of Egypt's several thousand BGP routes and traffic dropped to a handful of megabits per second. After the week long Internet outage following widespread social unrest and political protest, Egyptian Internet traffic returned to near normal levels at approximately 5:30am EST, Wednesday, February 2nd, 2011.

While other countries, including Iran and Myanmar, experienced telecommunication disruptions following social unrest in the past, the Egyptian outage represents a new Internet milestone. For the region, Egypt enjoys one of the largest and most robust Internet infrastructures with four major national providers and a hundred or more smaller consumer and web hosting providers. Put simply, we have never seen a country as connected as Egypt completely lose Internet connectivity for such an extended period. Also as a sign of the growing importance of social media, and web sites, it is telling that Egyptian telecommunications block largely focused on the Internet mobile and fixed line service returned earlier in the week. Today, the Internet is as an integral part the Egyptian economy and society. Unlike periods as recent as a decade ago, governments of technically developed countries cannot disrupt telecommunication without incurring significant economic cost and social / political pressures.

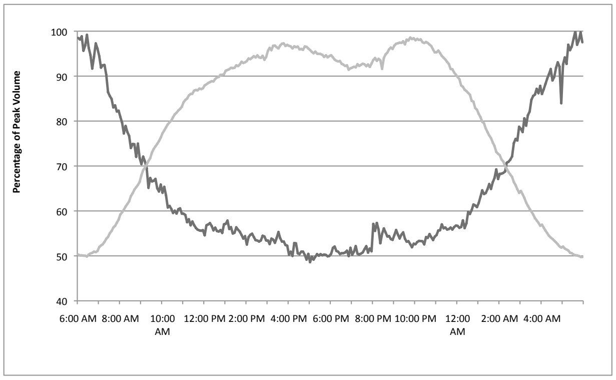

Figure 5: The traffic from North American consumer / regional providers and P2P traffic across several large large US and Canadian cable operators. Traffic is show as a percentage of peak traffic levels.

2.5 P2P

The debate on whether Internet Service Providers can, for economic reasons, block, curtail, or modify the content of Internet communications is commonly referred to as Net Neutrality [25]. While fairly representing the extent of this debate is outside the scope of a single paper, we note that such economically motivated modification of content speaks to the feasibility of such blocking and modification when the goals are overt censorship.

In a 2009 study, we explored the diurnal traffic patterns of Internet application in North America and Europe [15]. In particular, we looked for statistical traffic variations that might be indicative of intentional per application traffic manipulation. Figure 5 shows the unusual inverted behavior of p2p traffic relative to normal traffic on the Internet. It shows the daily average traffic fluctuations of 40 North American consumer / regional providers (taking the average of 10 weekdays in July). Since most P2P does not use standard ports and/or includes encryption, we extrapolate the data below using a combination of port data with statistics from application payload characterization across several large US and Canadian cable operators. We graph P2P as an average daily percentage of peak North American P2P Internet traffic and we show traffic as a percentage of peak traffic levels. The way to interpret the graph above is that at 6am EDT North American traffic volumes are at 50% of their daily peaks. Traffic then climbs to a local maxima at 4pm and then a daily peak around 11pm EDT before again dropping during the early morning hours. P2P reaches it low at 4pm when web and overall Internet traffic approaches its peak. The cyclical inverted traffic pattern of P2P is interesting in its own accord. While the inversion may be the result of unique P2P demand patterns or congestion, we believe it is largely due to the aforementioned provider manipulation of P2P traffic rates.

3 Challanges

While all of the incidents described in the last section involve different geographies and technologies, we argue the ongoing trend of co-option and outright corruption of the Internet infrastructure components represents risk for the long-term future of the Internet. Table 2 briefly summarizes these events. In this section we discuss both challenges that detection of these censorship efforts pose as well as well as technologies to ensure consistency and reliability of Internet.

3.1 Challenges in detection

In many cases, end users have little difficulty detecting censorship of their Internet traffic, although the specific censoring methodology may obscure the censoring party or technique. In other instances, the censorship is more subtle and difficult for end users to detect (e.g. Google Mail returning errors [1]) or BGP routes intentionally corrupted. In general, given the Internet's transcendence of historical boundaries makes analysis difficult. We discuss some of these challenges below.

Identifying Targeted Populations There are 6.7 billion people in the world living in 194 countries. 1.7 billion of these are on these Internet, which itself consists of 4 billion possible IPv4 addresses, 37 thousand "active" autonomous system numbers, and 359 thousand routed prefixes. Without prior knowledge (i.e., self identification), tracking changes in these populations are nearly impossible, even with global visibility. Complications in identifying organizations and other entities (e.g., 46,678 companies large enough that they have stocks traded in world stock exchanges) as well as changing shape of the Internet (e.g., IPv6) only further complicate identifying the target of censorship.

Locality The Internet is, in its trivial definition, a network of networks. As such, while globally interconnected, policy or censorship can be localized to a single autonomous system. This posses challenges to visibility, especially when identifying such censorship in edge ASes. Similar challenges have arisen in, for example, determining private routing policies within networks.

Assigning Intentionality Even with global visibility into infrastructure changes within a rogue ASes, ISP community detection of censorship is hampered by an inability to assign intentionality to a specific behavior. As such ISPs often find it difficult to disambiguate censorship from exogenous events like natural disasters, DDoS, cable cuts, etc. For example, what appeared to be the complete, instantaneous global disappearance of Egypt from the Internet in fact was much more controlled (ISPs left one by one in a period of a few minutes) and targeted (the ISP servicing the Egyptian stock market remained operational) [11].

3.2 Risks

In a world of rogue ASes, trusting your upstream provider is clearly a losing proposition. With sufficient resources any provider can block access to destinations, ports, protocols, or names (e.g., the GFC) and there is little to nothing that can be done about it short covert channel encoding of communications. In such cases, then, the goal of mitigation involves assuring the consistency and resilience of the core infrastructure.

Event	Filtering	Actor	Rationale
Egypt	block everything, block specific CIDR or ASN	State	Political
Iran	block everything, block specific CIDR or ASN, block specific port and or protocol	State	Political
China	block specific port and or protocol, block traffic matching payload or DNS contents	State	Political
P2P	block specific port and or protocol, block traffic matching payload or DNS contents	Company	Economic
Wikileaks	DDoS	Citizens	Political

Table 2: Various forms of censorship and their coarse grained implementation.

Routing In a notable 2008 incident, Pakistan government effort to block local access objectionable YouTube content went awry when corrupted BGP routes leaked to create global YouTube outage[2]. In a similar 2009 incident, Chinese IPSs announced US routes redirecting some traffic to China. In both cases, the likely primary intent of BGP announcements was localized censorship. Today, Internet routing like much of infrastructure relies primarily on trust between carriers. The challenge and risk of corrupting Internet routing integrity is unintended outages and balkanization of Internet topology. Unlike local security vulnerability which networks can secure at their borders, many of these threats are remote and require cooperation amongst many different parties.

For last twenty years, the IETF actively exploring several proposals for securing BGP[23]. Unfortunately, industry has make little progress with few production deployments. Unfortunately, the main challenge is economic. Unlike firewall or other endpoint security, threats to routing consistency are against a global, shared resource and require large-scale distributed effort to protect.

Naming Like routing, several notable recent incidents have renewed focus on weakness in DNS infrastructure. Increasingly, these weakness are exploited by criminals (e.g. fastflux spam and phising) ad governments to affect censorship. As with routing, significant risk of corruption leaking beyond intent local scope as happened with China I-root.

Also like routing security, IETF exploring solutions for better part of two decades. However, recent threats to DNS (e.g., Kaminsky) and significant security risks posed by untrusted DNS have motivated enterprises and carriers to begin adopting DNSSec[5]. As of this year, root zone signed and deployments underway.

Transport Finally, recent years have seen intentional disruption and manipulation of traffic at transport layer, including previously discussed P2P rate limiting and China reportedly introducing errors to block external Gmail access. While most of these transport layer corruptions have remained localized unlike DNS and routing leakage incidents, we argue the lack of transport security equally poseses a long-term threat to reliability and trust in Internet infrastructure.

Overall, IPSec [13] provides a long-term solution to protect against transport layer traffic manipulation or censorship. Unfortunately, less than one percent of Internet traffic uses IPSec today [22] and many countries and enterprises place restrictions on the use of encryption.

4 Conclusion

The Internet is a great thing(tm). It has enable the global-scale democratization of communication and effectively turn hundreds of millions of computing devices into private printing presses and libraries. The debate around free and open access to the Internet will continue to be of growing importance as networks becomes more and more intertwined in the economic, political, and social fabric of our societies. While the limiting of this free and open access for moral political, or economic reasons, is ultimately a societal challenge, this paper has examined how governments, citizens, and companies have sought to co-opt the core infrastructure to meet their goals. Such co-opting, used to scale, obscure, or more robustly filter, directly challenges core properties of the Internet including availability and security. Through an examination of global routing, allocation, and transport visibility, we have explored several recent and ongoing infrastructure co-opting events (e.g., Iran, China, and Egypt) and discussed the broad challenges in Internet community-based policing of such activity.

References

[1] http:http://tech.blorge.com/Structure:%20/2011/03/22/google-and-china-battling-again.

[2] http://www.ripe.net/internet-coordination/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study.

[3] Janna Quitney Anderson and Lee Rainie. Future of the Internet IV. http://www.pewinternet.org/Reports/2010/Future-of-the-Internet-IV.aspx, February 2010.

[4] Janna Quitney Anderson and Lee Rainie. The Impact of the Internet on Institutions in the Future. http://www.pewinternet.org/Reports/2010/Impact-of-the-Internet-on-Institutions-in-the-Future.aspx, March 2010.

[5] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. RFC 4035: Protocol modifications for the dns security extensions. http://datatracker.ietf.org/doc/rfc4035/, March 2005.

[6] Robert D. Atkinson, Stephen Ezell, Scott M. Andes, Daniel Castro, and Richard Bennett. The Internet Economy 25 Years After .com. http://www.itif.org/publications/internet-economy-25-years-after-com, March 2010.

[7] Martin A. Brown, Doug Madory, Alin Popescu, and Earl Zmijewski. DNS Tampering and Root Servers. http://www.renesys.com/tech/presentations/DNS-Tampering-and-Root-Servers.pdf, 2010.

[8] Richard Clayton. Failures in a hybrid content blocking system. In George Danezis and David Martin, editors, *Privacy Enhancing Technologies*, volume 3856 of *Lecture Notes in Computer Science*, pages 78–92. Springer, 2005.

[9] Richard Clayton, Steven Murdoch, and Robert Watson. Ignoring the great firewall of china. In George Danezis and Philippe Golle, editors, *Privacy Enhancing Technologies*, volume 4258 of *Lecture Notes in Computer Science*, pages 20–35. Springer Berlin / Heidelberg, 2006.

[10] Comcast Corporation. Before the federal communications commission in the matter of broadband industry practices. http://fjallfoss.fcc.gov/ecfs/document/view?id=6519840991, 2008.

[11] James Cowie. Egypt Leaves the Internet - Renesys Blog. http://www.renesys.com/blog/2011/01/egypt-leaves-the-internet.shtml, January 2011.

[12] Maximillian Dornseif. Government mandated blocking of foreign web content. In Jan von Knop, Wilhelm Haverkamp, and Eike Jessen, editors, *Security, E-Learning, E-Services: Proceedings of the 17. DFN-Arbeitstagung uber Kommunikationsnetz*, Lecture Notes in Informatics, pages 617–648, 2003.

[13] S. Kent and R. Atkinson. RFC 2401: Security architecture for the internet protocol. http://datatracker.ietf.org/doc/rfc2401/, November 1998.

[14] Craig Labovitz. A deeper look at the iranian firewall — security to the core — arbor networks security. http://asert.arbornetworks.com/2009/06/a-deeper-look-at-the-iranian-firewall/, 2009.

[15] Craig Labovitz. The internet after dark (part 1) — security to the core — arbor networks security. http://asert.arbornetworks.com/2009/08/the-internet-after-dark/, 2009.

[16] Craig Labovitz. Iranian traffic engineering — security to the core — arbor networks security. http://asert.arbornetworks.com/2009/06/iranian-traffic-engineering/, 2009.

[17] Craig Labovitz. The internet goes to war — security to the core — arbor networks security. http://asert.arbornetworks.com/2010/12/the-internet-goes-to-war/, 2010.

[18] Craig Labovitz. Round 2: Ddos versus wikileaks — security to the core — arbor networks security. http://asert.arbornetworks.com/2010/11/round2-ddos-versus-wikileaks/, 2010.

[19] Craig Labovitz. Wikileaks cablegate attack — security to the core — arbor networks security. http://asert.arbornetworks.com/2010/11/wikileaks-cablegate-attack/, 2010.

[20] Craig Labovitz. Egypt loses the internet — security to the core — arbor networks security. http://asert.arbornetworks.com/2011/01/egypt-loses-the-internet/, 2011.

[21] Craig Labovitz. Egypt returns to the internet — security to the core — arbor networks security. http://asert.arbornetworks.com/2011/02/egypt-returns-to-the-internet/, 2011.

[22] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. Internet Inter-Domain Traffic. In *Proc. ACM SIGCOMM*, 2010.

[23] M. Lepinski and S. Turner. Internet Draft an overview of bgpsec. http://tools.ietf.org/html/draft-lepinski-bgpsec-overview-00, March 2011.

[24] Graham Lowe, Patrick Winters, and Michael L. Marcus. The Great DNS Wall of China. http://cs.nyu.edu/~pcw216/work/nds/final.pdf, 2007.

[25] The New York Times. Net neutrality. http://topics.nytimes.com/topics/reference/timestopics/subjects/n/net_neutrality/index.html, 2010.

[26] International Telecommunication Union. Measuring the Information Society. http://www.itu.int/ITU-D/ict/publications/idi/2010/Material/MIS_2010_without%20annex%204-e.pdf, 2010.

[27] Eric Wustrow, Manish Karir, Michael Bailey, Farnam Jahanian, and Geoff Houston. Internet Background Radiation Revisited. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement (IMC '10)*, Melbourne, Australia, November 2010.

[28] Xueyang Xu, Z. Morley Mao, and J. Alex Halderman. Internet censorship in china: Where does the filtering occur? In *12th Passive and Active Measurement (PAM '11)*, Atlanta, GA, March 2011.