Analysis of Country-Wide Internet Outages Caused by Censorship

Alberto Dainotti, Claudio Squarcella, Emile Aben, Kimberly C. Claffy, Marco Chiesa, Michele Russo, and Antonio Pescapé, *Senior Member, IEEE*

***Abstract*—In the first months of 2011, Internet communications were disrupted in several North African countries in response to civilian protests and threats of civil war. In this paper, we analyze episodes of these disruptions in two countries: Egypt and Libya. Our analysis relies on multiple sources of large-scale data already available to academic researchers: BGP interdomain routing *control plane* data, unsolicited *data plane* traffic to unassigned address space, active macroscopic traceroute measurements, RIR delegation files, and MaxMind's geolocation database. We used the latter two data sets to determine which IP address ranges were allocated to entities within each country, and then mapped these IP addresses of interest to BGP-announced address ranges (prefixes) and origin autonomous systems (ASs) using publicly available BGP data repositories in the US and Europe. We then analyzed observable activity related to these sets of prefixes and ASs throughout the censorship episodes. Using both *control plane* and *data plane* data sets in combination allowed us to narrow down which forms of Internet access disruption were implemented in a given region over time. Among other insights, we detected what we believe were Libya's attempts to test firewall-based blocking before they executed more aggressive BGP-based disconnection. Our methodology could be used, and automated, to detect outages or similar macroscopically disruptive events in other geographic or topological regions.**

***Index Terms*—Censorship, connectivity disruption, darknet, Internet background radiation, network telescope, outages.**

I. Introduction

On THE evening of January 27, 2011, Egypt—a population of 80 million, including 23 million Internet users[52]—vanished from the Internet. The Egyptian government ordered a complete Internet shutdown amidst popular anti-government protests calling for the resignation of Egyptian President Hosni Mubarak. The order followed reports on the previous day (January 25) of blocked access to Twitter [31], although an Egyptian government official denied blocking any social media Web sites [60]. In addition to mainstream media reporting of traffic disruption [61], several commercial Internet measurement companies posted technical data analyses on their blogs [20], [41], [62]. The heavy-handed attempt to block communications in the country did not quell the protests and may have even increased the number of people in the streets; protests intensified and continued even after Internet connectivity was restored 5 days later. Under political pressure from inside and outside Egypt, President Mubarak resigned, turning command over to the military on February 11.

Four days later, similar protests erupted in Libya, calling for an end to the Gaddafi regime. On February 17, major protests took place across the country [55], and that evening, YouTube became inaccessible from Libya [33]. On the night of February 18 (Friday), the government imposed an "Internet curfew," blocking all Internet access until morning (08:01 local time), and repeating it the next day (Saturday) [20], [41], [75]. In the following days, Libyan traffic to popular sites like Google increased steadily [34] until Internet access was disabled again, this time for nearly 4 days. Fig. 1 presents a brief chronology of the events.

Wide-scale Internet service disruptions are nothing new [14], [19], [69], [78]–[81], even politically motivated interference with Internet access in order to hinder anti-government organization [3], [9]. However, the scale, duration, coverage, and violent context of Egypt's and Libya's disruptions has brought the Internet "kill switch" issue into new critical light.

In this paper, we analyze the Internet disruptions that took place in Egypt and Libya in early 2011. These two events are of historical as well as scientific interest. Access to publicly available Internet measurement data that cover the outage intervals allows empirical study of what it takes to bring down an entire country's communications infrastructure, which has security relevance for every nation in the world. We were also able to observe surprisingly noticeable effects of such large-scale censorship on ongoing global measurement activities, suggesting how similar events could be detected and/or documented in the future. Our analysis relies on multiple measurements and vantage points.

Manuscript received April 17, 2013; revised September 10, 2013; accepted October 19, 2013; approved by IEEE/ACM Transactions on Networking Editor R. Mahajan. The UCSD network telescope operations and data collection, curation, analysis, and sharing were supported by the NSF CRI under Grant CNS-1059439, the DHS S&T under Grant NBCHC070133, and the UCSD. The Ark infrastructure operations and data collection, curation, and sharing, as well as K. C. Claffy's effort on this project, were supported by the DHS S&T under Grants NBCHC070133 and N66001-08-C-2029 and the NSF under Grant CNS-0958547. The work of A. Dainotti and K. C. Claffy was supported in part by the NSF under Grant CNS-1228994. The work of E. Aben was supported by RIPE NCC, although the majority of his contribution was on his own time so as to not interfere with his RIPE-NCC responsibilites. The work of C. Squarcella and M. Chiesa was supported in part by the MIUR of Italy under Project AlgoDEEP prot. 2008TFBWL4. Part of their research was conducted in the framework of the ESF project 10-EuroGIGA-OP-003 GraDR "Graph Drawings and Representations." The work of A Pescapè was supported in part by PLATINO (PON01_01007), the MIUR, and a Google Faculty Award for the UBICA project.

A. Dainotti and K. C. Claffy are with the CAIDA, University of California San Diego, La Jolla, CA 92093 USA (e-mail: alberto@caida.org).

C. Squarcella and M. Chiesa are with Roma Tre University, Rome 00146, Italy.

E. Aben is with RIPE NCC, Amsterdam 1016, The Netherlands.

M. Russo and A. Pescapé are with the Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Universitá degli Studi di Napoli Federico II, Naples 80125, Italy.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2013.2291244

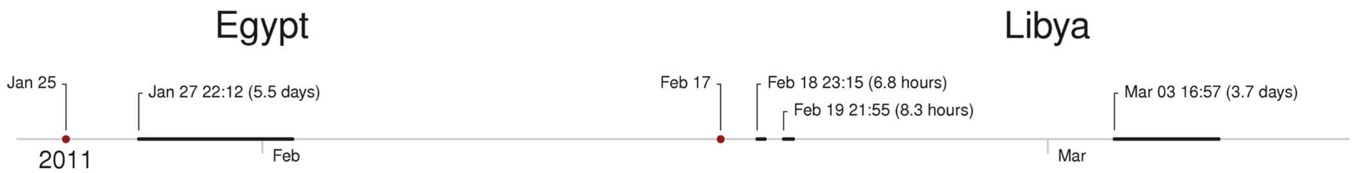

Fig. 1. Timeline of Internet disruptions described in the paper. Times in figure are UTC (Egypt and Libya are UTC+2). The pair of dots indicate the start of major political protests in the respective countries. The thicker horizontal segments indicate the approximate duration of the outages.

- *Traffic to unassigned address space*, specifically data collected from the UCSD network telescope [8], reveal changes in background Internet traffic from the affected regions.
- *BGP data* from Route Views [71] and RIPE NCC's Routing Information Service (RIS) [6] provide a view of BGP activity during the events.
- *Active traceroute probing* from Ark [2] and iPlane measurement infrastructures [49] reveals forward path information and bidirectional reachability. These data can reveal additional types of filtering, e.g., firewalling, not observable via BGP, as well as help calibrate BGP observations.

We focus on two episodes of Internet disruption, although the same global data sources and our proposed methodology could illuminate the study of similar events in other countries. The main contributions of this paper are as follows.

- We document a rich view of the disruptions using three types of measurement data sources.
- We demonstrate how to use unsolicited background traffic to identify country-level blocking events.
- We report previously unknown details of each event, including the use of packet filtering as well as BGP route withdrawals to effect the disruption.
- We sketch a general methodology for the analysis of some types of disruptions that combines multiple measurement data sources available to researchers.
- Our methodology and findings can form the basis for automated early-warning detection systems for Internet service suppression events.

The rest of this paper is organized as follows. Section II gives technical background on network disruption, limiting its focus to the paper's scope. Section III summarizes related work. Section IV describes our measurement and analysis methodology. Section V presents the results of our analysis. Section VI discusses our findings and concludes the paper.

II. BACKGROUND

Disabling Internet access is an extreme form of Internet censorship in which a population's Internet access is blocked completely, a coarse but technically more straightforward approach than the selective blocking used in most Internet censorship regimes. It can be implemented by simply powering down or physically disconnecting critical equipment, although this approach typically requires physical co-location with the communications equipment, which may be spread over a wide area. A more flexible approach is to use *software* to disrupt either the *routing* or *packet forwarding* mechanisms.

Routing Disruption: While *forwarding* is the mechanism that advances packets from source to destination, the *routing* mechanism determines which parts of the network are reachable and how. On the Internet, global routing is coordinated at the level of autonomous systems (ASs)—administratively distinct parts of the network—using BGP as the interdomain routing protocol. Routers exchange BGP information (*BGP updates*) regarding which destination addresses they can reach, and continually update their forwarding tables to use the best available path to each contiguous block of destination addresses.[1] Disabling the routing process on critical routers or suppressing BGP information transmission can effectively render large parts of a network unreachable.

Packet Filtering: Packet filtering intervenes in the normal packet forwarding mechanism to block (i.e., not forward) packets matching given criteria. Today, packet filtering of varying complexity is a standard security feature of network equipment from switches to dedicated firewalls.[2]

Disabling a network's connectivity to the Internet through BGP routing disruption is easily detectable since it entails changes in the *global* routing state of the network, i.e., in the *control plane*. Previously advertised prefixes must be withdrawn or readvertised with different properties in order to change global routing behavior. Detecting packet filtering is harder; it requires either active probing of the forward path, or monitoring traffic (the *data plane*) from the affected parts of the network for changes.

We also note that BGP-based traffic control and packet filtering are only two of many layers where censorship of connectivity and content may occur. A taxonomy of blocking technologies would include a range of approaches, including physical-layer disconnection, content blocking based on deep packet inspection, DNS-based blocking or manipulation at multiple granularities, injection (e.g., of TCP RST packets), and end-client software blocking in an enterprise or household.

III. RELATED WORK

An Internet blackout is put into effect by a government when more selective forms of censorship are either impractical or ineffective. Governments use selective censorship to restrict Internet traffic for political, religious, and cultural reasons. Recent studies have found that Internet censorship is widespread and on the rise [3], [9]. Unsurprisingly, there is also growing interest in analyzing the technical aspects of censorship and methods to circumvent them.

---

[1] Destination addresses are advertised as *prefixes*, ranges of addresses represented as a set of fixed high-order bits and varying low-order bits. A prefix is written as an IP address in dotted quad notation followed by a "/" and the prefix length. For example, the prefix 132.239.0.0/17 represents IP addresses 132.239.0.0 to 132.239.127.255.

[2] Because of their versatile packet filtering capabilities, firewalls can flexibly implement many types of selective censorship: address filtering, protocol filtering, and simple content filtering.

In [17], Clayton *et al.* analyzed the keyword filtering mechanism of the "Great Firewall of China" (GFC) and found that it was possible to circumvent it by simply ignoring the forged TCP packets with the reset flag set sent by the firewall. In [25], Crandall *et al.* disproved the notion that the GFC is a firewall strictly operating at the border of China's Internet. They showed that filtering can occur at different hops past the Chinese border (cases of up to 13 IP address hops were observed), suggesting that filtering happens at the AS level. They also proposed *ConceptDoppler*, a system for monitoring which keywords are filtered over time. In more recent work on censorship in China [76], Xu *et al.* explored the AS-level topology of China's network and probed the firewall to find the locations of filtering devices. They found that two major ISPs in China had different approaches to the placement of filtering devices, either decentralized or in the backbone. Results also showed that most of the filtering happened in "border ASs," that is, ASs that peer with foreign networks. In [65], Skoric *et al.* provided a detailed analysis of how social media platforms (e.g., Facebook), as well as traditional media were used to organize a student protest against censorship in Singapore. They found that activists used social media to engage rather than circumvent traditional media stakeholders, in order to amplify their impact. Trying to prevent such amplification was a likely motivation for the Internet blackouts we analyzed in this study.

The scientific study of Internet interdomain routing has a longer and richer history (thus far) than the study of Internet censorship. Academic researchers have used BGP data to support scientific study of Internet dynamics for over a decade, including Labovitz *et al.*'s studies in the late 1990s of several root causes of high-levels Internet routing instability [44], [45]. They traced a large fraction of unnecessary ("pathological") BGP update volume to router vendor software implementation decisions, and their research convinced router vendors to modify BGP parameters, decreasing the volume of Internet routing updates they observed a year later by an order of magnitude. More recently, researchers have explored the use of spatial and temporal correlations among the behaviors of BGP prefixes to detect and hopefully predict instabilities [30], [39].

In [63], Sahoo *et al.* used simulations to estimate BGP recovery time for networks with a range of topological characteristics undergoing large-scale failure scenarios, such as those caused by disastrous natural or man-made events. They showed that topological properties, especially the node degree distribution, can significantly affect recovery time. In [47], Li and Brooks proposed an *Internet seismograph* to consistently measure the impact of disruptive events such as large-scale power outages, undersea cable cuts, and worms to enable quantitative comparison of the impact of different events. Other researchers have also used BGP monitoring infrastructure to analyze the impact of these types of disruptive events on the routing system [23], [24], [46]. In [73], Wan and Van Oorschot used data from the Route Views project [71] during Google's outage in 2005, showing what was probably a malicious BGP advertisement of a Google network prefix by another AS.

Several analysis and real-time BGP monitoring tools have also been developed. In [68], Teoh *et al.* presented BGP Eye, a tool for root-cause analysis and visualization of BGP anomalies based on clustering of BGP updates into events that are compared across different border routers. The visualization helps to show events that may affect prefix reachability and hijacking. In [38], Katz-Bassett *et al.* introduced Hubble, a prototype system that operated continuously for a couple of years to find Internet reachability problems in real time, specifically for situations where routes exist to a destination but packets are unable to reach the destination. While the Hubble system detected unreachability of prefixes, it did not aggregate results beyond a single AS. Massey's group at Colorado State is maintaining BGPMon, a real-time BGP routing instrumentation system [77] designed to scalably monitor BGP updates and routing tables from many BGP routers simultaneously, while also allowing user-defined real-time "notification" feeds of specific subsets of the data.

Regarding the specific blackouts we describe in this paper, several commercial Internet measurement companies posted technical data analyses on their blogs during and following the outages [12], [20]–[22], [40]–[43], [72]. Most of them analyzed flow-level traffic data or BGP updates seen by their routers, providing timely news and technical insights, although omitting specifics about the proprietary data and analysis methodologies used.

Our work combines multiple different data sources available to academic researchers to analyze in detail the characteristics of two large-scale censorship episodes that occurred in early 2011. None of the data analysis studies described above was designed to detect phenomena at a country granularity, nor to integrate traffic, topology, and routing data, although their techniques could complement the ones we present in this paper to improve detection of countrywide network disruptions. Moreover, even if the original idea of using unsolicited traffic—sometimes called darknet or telescope traffic—for "opportunistic measurement" was introduced in 2005 by Casado *et al.* [16], we are not aware of any previous work using it to monitor Internet outages. Specifically, Casado *et al.* showed that by analyzing some categories of such traffic they could infer several properties of the infected hosts generating it (e.g., access link bandwidth, NAT usage).

Peer-to-peer (P2P) networks are another source of data to study Internet-wide disruptions. Bischof and Otto used data gathered from the Vuze BitTorrent client to analyze the disconnection of Egypt and Libya from the Internet [13]; their approach assumed widely distributed active P2P clients using a specific and far from ubiquitous application during events of interest. Since darknet traffic is generated mostly by malware, our passive measurements do not depend on end-user actions or even awareness that this traffic is being generated.

Another approach to the detection of outages based on passive traffic analysis was presented by Glatz *et al.* in [32]. By classifying one-way flows exiting from a live network (not unsolicited traffic), they were able to identify outages related to services on the Internet accessed by users of the monitored network. In an earlier version of this paper [29] and in [26], we used unsolicited traffic as a novel method to analyze large-scale Internet disruption. In [29], we also showed an example of how the visual tool BGPlay helped us in manually investigating the events that we analyzed, whereas in [26] we extended the darknet-based component of our approach to the analysis of outages caused by natural disasters. The analysis approach

TABLE I
IPv4 ADDRESS SPACE DELEGATED TO EGYPT (AS OF JANUARY 24, 2011) AND LIBYA (AS OF FEBRUARY 15, 2011) BY AFRINIC (TOP HALF) AS WELL AS ADDITIONAL IPv4 ADDRESS RANGES ASSOCIATED WITH THE TWO COUNTRIES BASED ON MAXMIND GEOLITE DATABASE (AS OF JANUARY 2011)

	Egypt	Libya
AfriNIC delegated IPs	5,762,816	299,008
MaxMind GeoLite IPs	5,710,240	307,225

presented in this paper facilitated the discovery of important technical aspects of the outages in Egypt and Libya not previously reported, including the combination of packet filtering techniques with BGP disruption and disruption of satellite Internet connectivity.

IV. DATA SOURCES AND METHODOLOGY

A. Internet Geography

To properly observe the effects of a blackout in a given country, we first need to identify which Internet numbering resources are under the control of, or related to, those countries. In this section, we explain how we used geolocation data to select the relevant set of IP addresses, BGP prefixes, and AS numbers to monitor for visibility into Egypt and Libya during the blackout intervals.

1) IP Addresses: Five regional Internet registries (RIRs) manage the allocation and registration of Internet resources (IP prefixes and autonomous system numbers) to entities within five distinct regions of the world and publish lists of the Internet resources they delegate, which include the country hosting the headquarters of the receiving entity. Egypt and Libya are in the AfriNIC (Africa's) region [1]. The first row of Table I lists the number of IPv4 addresses delegated to Egypt and Libya by AfriNIC. Additionally, IP addresses nominally allocated to a different country may be used within Egypt and Libya if a foreign ISP provides Internet access in those countries. The second row of Table I lists the IP addresses geolocated in Egypt and Libya according to the MaxMind GeoLite Country database [51]. We used these two sources of data to construct the list of IP address ranges (prefixes) that geolocated to Egypt and Libya.

Although there are accuracy issues in all geolocation databases [36], [58], [64], at a country granularity these databases almost always agree with (sometimes because they are based on) the RIR-delegation file information. Countries with either low Internet penetration or a small population (Libya has both) may only have few IPs officially RIR-delegated to them, in which case IP geolocation database information can provide useful additional information. Satellite connectivity, which at least one ISP uses to serve Libya, is another source of IP geolocation discrepancy.

2) AS Numbers and BGP-Announced Prefixes: Once we had derived a set of IP address ranges to monitor, we mapped these to BGP-announced prefixes and ASs announcing those prefixes, using a database constructed from publicly available BGP data from Route Views and RIPE NCC RIS [6], [71] for the week preceding the outages, up to and including the first day of the outage. The allocated address ranges might not map precisely to a BGP prefix boundary since ASs may implement complex routing policies to accomplish traffic engineering and other business objectives, which may involve splitting BGP prefixes into smaller chunks or aggregating prefixes into larger chunks to reduce routing table sizes. Thus, different views of the routing system (BGP monitors) may have different BGP prefixes covering a given set of IP addresses. Once we gather BGP events within the time window we want to observe, we compute the set of *covering* prefixes P for address space S as follows.

- We look up the address space in the BGP database described above to find an exactly matching BGP prefix.
- We find all the more specific (strict subset, longer) prefixes of this prefix.
- If the two previous steps yielded no prefix, we retrieve the longest BGP prefix entirely containing the address space S.

For each AS, we show results only for the IP ranges or BGP prefixes that are solely related to the country under analysis, e.g., traffic whose source addresses are included in prefixes announced by that AS *and* are geolocated or delegated to the country under analysis.

B. BGP Data

BGP routing changes can rapidly induce global effects, including coarse-grained filtering that may be indistinguishable from complete physical disconnection of infrastructure. Using BGP data in conjunction with data-plane measurements such as traceroute or traffic flows can yield a rich understanding of the type of censorship strategy being used.

The two main sources of BGP updates used throughout our analysis are the already mentioned Route Views Project [71] and RIPE NCC's Routing Information Service (RIS) [6], respectively maintained by the University of Oregon, Eugene, OR, USA, and RIPE NCC, Amsterdam, The Netherlands. Both services rely on routers that establish BGP peering sessions with many ISPs around the world. The available data reveal a broad and global though necessarily incomplete view of BGP connectivity over time, at an announced-prefix granularity. We analyzed this data at the finest possible time granularity—BGP updates—to detect and isolate events observed during the outages. However, BGP updates only provide information about *changes* in routing state. Each route collector also periodically dumps a snapshot of its entire control plane table, called an *RIB*, containing all known routing information related to prefixes that are reachable at that point in time. We used these periodic dumps in conjunction with the fine-grained updates to track a precise view of prefix reachability over the duration of the outage intervals.

Each source of data also has a graphical tool to query for specific prefixes, BGPlay [70] for Route Views and BGPviz [4] for RIS. Online services like REX [5] and RIPEstat [7] allow coarse-grained analysis of historical BGP events in RIS data, based on snapshots taken every 8 h of the RIB table dumps. (Route Views RIBs are dumped every 2 h.)

To perform chronological analysis of the outages, we first identified the time window during which disconnection activity was observed, using previous reports [20], [21], BGPlay, and BGPviz. We extended this window to 1 h before and 1 h after the main events we observed to detect possible early symptoms of the outage and late reconnection patterns. We used the last RIB table dumps from both Route Views and RIS just before this interval as the starting point, and used BGP updates and subsequent BGP table dumps to reconstruct the routing history during the time window under examination.

For each prefix, we processed the downloaded data to build a set of routing histories, one for each route collector peer (that is, an AS feeding its BGP updates to a route collector). We marked a prefix as *disappeared* if it was withdrawn during the blackout interval for each of the above routing histories, i.e., no longer observable from any route collector peer, and remained so for the duration of the interval. We chose the earliest of those withdrawals as the event representing the initial disconnection of the prefix. This approach provides an overview of the prefixes going down over time, as well as the first signs of disconnection for each withdrawn prefix.

We used a similar approach to study the end of the outage, focusing instead on when BGP prefixes become reachable again via announcements seen in BGP updates. We chose the earliest reannouncement for each prefix to determine a representative time instant for the reconnection. BGPlay and BGPviz were also useful to visualize the transition periods, to see which prefixes were still visible as the outage information propagated, and to see peers reconverge to secondary backup paths. We visualized the reconnection as well, with peers reverting to primary paths as prefixes were reannounced.

Note that the collected data are subject to uncertainty, especially regarding timing of events, so we cannot obtain a perfect understanding of BGP dynamics. BGP messages are sometimes delayed (aggregated) by routers in order to minimize control traffic. Furthermore, router clocks are not necessarily synchronized, so one cannot be sure of the exact interleaving sequence of events occurring at different routers. However, empirical analysis of data coming from different collectors generally shows good correlation in terms of time values (see, e.g., recent work by Lui *et al.* [48]). While imperfect, our methodology provides a consistent way to approximate the timing of BGP withdrawals during an outage.

C. Darknet Traffic

Unsolicited one-way Internet traffic, also called *Internet background radiation* (IBR) [56], has been used for years to study malicious activity on the Internet including worms, denial-of-service (DoS) attacks, and scanning address space looking for vulnerabilities to exploit. Such a vast number of computers generate such background radiation, mostly unbeknownst to their legitimate users, that the resulting traffic aggregate has proven a useful source of data for observing characteristics of the malware itself [28], [53], [54] not revealed by other types of data.

Researchers observe IBR traffic using *network telescopes*, often called *darknets* if the IP addresses are not being used by devices. We collected and analyzed traffic arriving at the UCSD network telescope [8], which observes a mostly unassigned /8, that is, 1/256 of the entire IPv4 address space. We used the same IP-AS databases described in Section IV-A to determine the levels of unsolicited traffic throughout the outages in Egypt and Libya. Although the unsolicited traffic from these countries exhibited a typical diurnal pattern, sudden changes in the packet rate suggest start and end times of several of the outages (see Fig. 2).

There are three primary causes of IBR traffic: 1) *backscatter* from spoofed DoS attacks; 2) scans; or 3) bugs and misconfiguration [56]. Different types of IBR traffic induce separate sources of packet rate dynamics, which can heavily affect the amount of traffic observed from a specific country to the network telescope. Our methodology identifies and separates these sources of IBR-induced dynamics to avoid misinterpreting them.

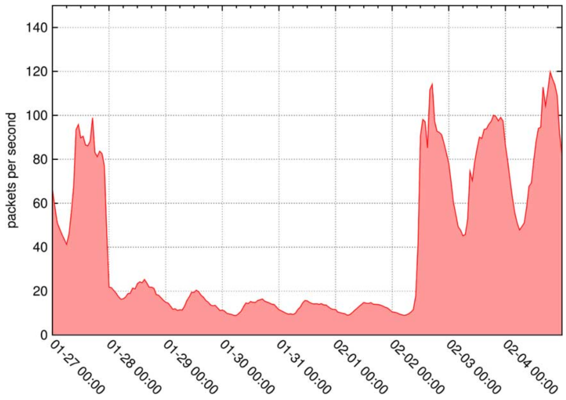

Fig. 2. Unsolicited packets from IPs geolocated in Egypt to UCSD's network telescope. The two dramatic changes in the packet rate respectively match the withdrawals and reannoumcements of BGP routes to Egyptian networks.

Packets associated with denial-of-service attacks represent a special category because they cause substantial packet rate variation, especially for countries with few IP addresses such as Egypt and Libya. A DoS attack attempts to overwhelm a victim with traffic or transaction requests in order to reduce or prevent his ability to serve legitimate requests. When the source IP addresses in attacking packets are randomly spoofed, the response packets (e.g., *SYN-ACK* TCP packets in reply to *SYNs* from the attacker) are sent back to the spoofed addresses, producing *backscatter*, which will be captured by telescopes [54] that happen to contain (and thus observe traffic to) those spoofed addresses. To identify and characterize these attacks, we used the methodology in [54], separating potential backscatter packet flows (i.e., TCP packets with SYN-ACK or RST flags on, ICMP echo replies) by sender (potential victims of the attack), then classifying any such flows above predefined volume or duration thresholds as backscatter reflecting a denial-of-service attack. Of course, DoS attacks to a country in civil unrest may themselves be of interest; we did notice attacks on some government Web sites approximately when the protests began and a short time before the outage (see Sections V-A.3 and V-B.3.)

Automated (e.g., from worms) or manually initiated random scanning of address space in search of victims is another component of IBR [28], [53]. On November 21, 2008, the amount of unsolicited traffic grew dramatically with the advent of the Conficker worm [10], which widely infected Windows hosts and actively scanned for hosts to infect on TCP port 445. Sufficiently pervasive network scanning such as done by Conficker reveals surprisingly detailed insights into global Internet behavior. In particular, geolocation of all IP source addresses of such scanning packets makes it easy to detect countrywide outages [11] since an entire country disappears from this category of traffic. We identified Conficker scanning traffic by selecting TCP SYN packets with destination port 445 and packet size 48 B.

Misconfiguration of systems can also induce IBR traffic, for example by setting the wrong IP address on a DNS or proxy server. Bugs in network applications and router firmware and software, e.g., getting byte ordering wrong, can assign incorrect network addresses to a device, triggering unsolicited traffic in response.

We further analyzed the unsolicited traffic by AS coming from within each of the two countries, revealing AS-specific behavior. IP address space that has been withdrawn from the global BGP routing table will not be able to receive traffic from the Internet default-free zone anymore, but may be able to successfully send outbound traffic in the absence of packet filtering. Analysis of background radiation, especially Conficker-like scanning, reveals some of this leaked traffic.

D. *Active Forward Path Probing*

We made use of active measurements taken during the outages in Egypt and Libya, toward address space in these countries. The measurements consisted of: 1) *ad hoc* measurements using standard ping and traceroute; 2) structured global IPv4 address probing from CAIDA's IPv4 Routed /24 Topology Dataset [37] collected on CAIDA's Ark infrastructure [2], and traceroutes collected by the iPlane project [49] through distributed probing from PlanetLab nodes [57].

Both the Ark and iPlane datasets show surprising two-way connectivity surviving the outage intervals that span more than a day. However, these data do not provide sufficient granularity to analyze in detail the shorter outages. Ark takes about 3 days to complete a full cycle probing all /24 IPv4 subnets from prefixes that have been announced before, whereas iPlane probes a single address per announced prefix and its data are organized in daily files without storing individual per-probe timestamps. Another problem is the relatively few IP prefixes in each country.

V. ANALYSIS

In this section, we present our analysis of the Internet blackouts from the perspectives of the control plane (BGP data) and the data plane (traffic and traceroute data). Section V-A discusses the outage in Egypt; Section V-B discusses the several disconnections in Libya. To prevent unintentional harm to those involved in these episodes of censorship or their circumvention, we have anonymized most AS numbers in this paper as described in the Appendix.

A. *Egypt*

1) Overview: According to Mahlknecht's cable map Web site [50], which might have dated information, Egypt's Internet infrastructure is dominated by state ownership, consisting primarily of a few large players with international connectivity through the major submarine cable systems that run through the Suez canal. Most, if not all, submarine fiber-optic circuits are backhauled to the Ramses Exchange [74], which is not only the main connection facility for the Egyptian-government-controlled telecommunications provider, but also the location of the largest Internet exchange point in North Africa or the Middle East. Both the small number of parties involved in international connectivity and the physical connectivity under control of the state telecommunications provider facilitate manipulation of the system by a state actor, as shown by the events described below.

Renesys reported that on January 27, 2011, around 22:34:00 GMT, they observed the "*virtually simultaneous withdrawal of all routes to Egyptian networks in the Internet's global routing table*" [20]. The packet rate of unsolicited traffic from Egypt seen by the UCSD telescope (Fig. 2) suddenly decreased at almost the same time, on January 27 around 22:32:00 GMT. In terms of BGP, the methodology explained in Section IV identifies the outage as a sequence of routing events between approximately 22:12:00 GMT and 22:34:00 GMT. The outage lasted for more than 5 days, during which more active BGP IPv4 prefixes in Egypt were withdrawn. In Fig. 3, each step represents a set of IPv4 prefixes at the point in time when they first disappeared from the network.

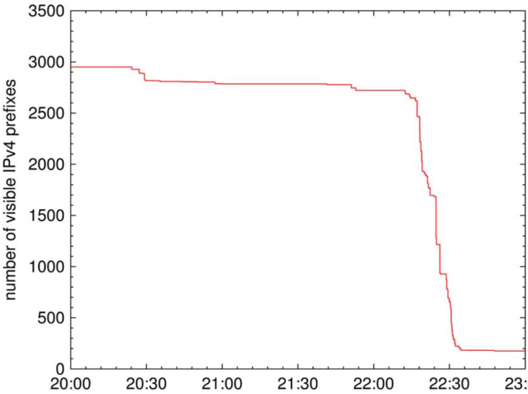

Fig. 3. Disconnection of Egyptian IPv4 prefixes via BGP during the outage on January 27, based on data from Route Views and RIPE NCC's RIS. For each disconnected prefix, the line drops down at the instant in which a lasting (i.e., not temporarily fluctuating) BGP withdrawal is first observed.

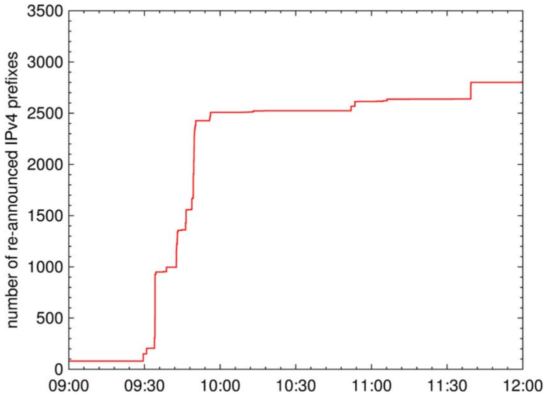

Fig. 4. Reannouncement of Egyptian IPv4 prefixes via BGP at the end of the outage on February 2, based on data from Route Views and RIPE NCC's RIS. For each reannounced prefix, the line goes up at the instant in which a stable BGP announcement is first detected.

The UCSD darknet traffic returned to packet rates comparable to those preceding the outage at 10:00:00 GMT on February 2. The unsolicited traffic level from Egypt to the telescope was roughly consistent with our BGP analysis, which found that the first set of reannouncements of Egyptian connectivity after the crisis occurred around 09:29:31 GMT. Fig. 4 shows the BGP connectivity reappearing.

2) Outages in Detail: BGP data reveal a dramatic drop in reachability for many Egyptian IPv4 prefixes during the outage. It is not obvious which event should be considered the first sign of the outage. The leftmost end of the graph in Fig. 3 shows 2928 Pv4 prefixes visible via BGP at 20:00:00 GMT. A noticeable loss of connectivity is first seen by Route Views and RIS route collectors on January 27 at 20:24:11 GMT, related to

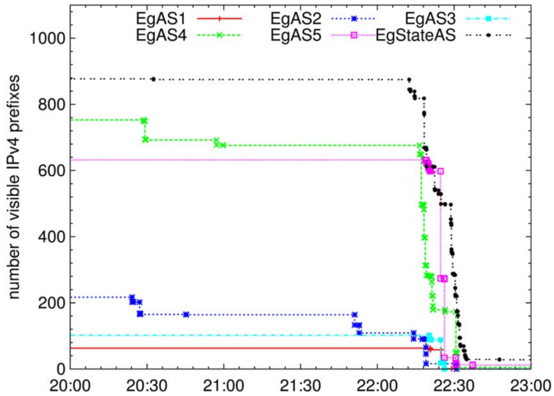

Fig. 5. Visibility of main Egyptian autonomous systems via BGP during the outage on January 27 (based on data from Route Views and RIPE NCC's RIS). Each AS is plotted independently; as in Fig. 3, each line drops down at the instant in which a lasting (i.e., not temporarily fluctuating) BGP withdrawal is first observed.

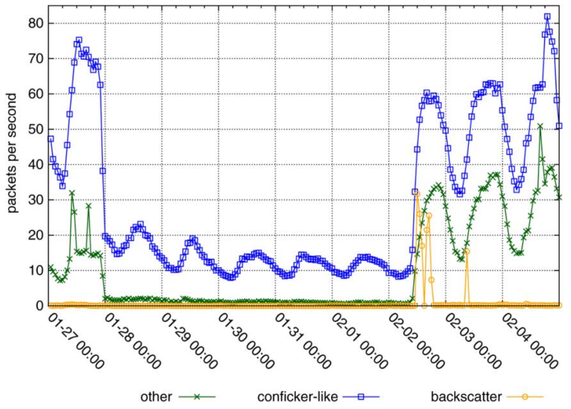

Fig. 6. Categories of unsolicited packets from IPs geolocated in Egypt to UCSD's network telescope: other, conficker-like, backscatter. Spikes in backscatter traffic reflect large denial-of-service attacks against hosts in Egypt.

15 IPv4 prefixes routed by EgAS2.[3] Further losses of BGP connectivity are visible in the next 2 h, summing up to 236 withdrawn IPv4 prefixes. The biggest disruption then appears as an almost vertical drop in Fig. 3, with the initial step at 22:12:26 GMT, after which roughly 2500 prefixes disappear within a 20-min interval. At 23:30:00 GMT, only 176 prefixes remain visible.

Fig. 5 shows the same sequence of events separated by the six main Egyptian ASs. Although the image seems to suggest a time sequence for the interleaving BGP withdrawals, we can make no safe assumption on the chronology of underlying decisions.

Contrary to IPv4 prefixes, there was no major change in visibility for IPv6 prefixes. Of the six IPv6 prefixes in AfriNIC's delegated file, only one is seen in RIS, and this prefix of length /32 is announced by IntAS1, a major international carrier. This prefix stayed visible during the outage, as did all its more specific prefixes seen in RIS (20 /48s announced by EgAS4 and one /48 by EgAS6).

Fig. 6 shows a breakdown of the traffic observed by the UCSD network telescope in three categories: *conficker*, *backscatter*, and *other*. Conficker refers to TCP SYN packets with destination port 445 and packet size 48 B. While we assume that these packets are generated by systems infected by the Conficker worm scanning for new victims, we cannot be absolutely certain, although our inferences are valid if the majority of packets satisfy this assumption. The most important implication of the assumption is that the source IP addresses are not spoofed; if they were, the geolocation mapping would be meaningless.

The backscatter category of traffic requires careful treatment. When an attacker uses fake source IP addresses in a denial-of-service attack targeting a victim in the address space we are trying to observe, backscatter traffic from the victim IP addresses can increase suddenly and dramatically, jeopardizing our inferences about background traffic levels coming from this address space. Thus, our methodology must identify and filter out this backscatter traffic.

The *other* category represents all other packets composing the background radiation [56] captured by the network telescope: worms, generic scanning and probing activity, misconfigurations, etc.

Fig. 6 reveals the diurnal patterns of activity in Conficker traffic, which are typical of (infected) PCs that are not kept on 24 h per day. Conficker traffic is the dominant component, and stays partially alive even during the outage, for two reasons: 1) some prefixes are still visible via BGP; and 2) outbound connectivity still works for some networks. For a given network, BGP withdrawals are a consequence of either propagation of a withdrawal from elsewhere in the network, or a data-plane failure immediately adjacent to the router. In the former case, the network is unreachable from the outside world, but may still be able to send packets in the outbound direction. In the same figure, the *backscatter* traffic has some spikes related to a couple of denial-of-service attacks that we discuss in Section V-A.3.

The *other* category of traffic in Fig. 6 is most interesting: Soon after the Egyptian networks are again BGP-reachable, the packet rate of this traffic grows much higher than before the outage. By analyzing traffic from the entire Internet reaching the UCSD darknet, we found that a large UDP/TCP scan targeted toward a specific service was conducted from thousands of hosts all around the world. The diurnal pattern suggests this traffic was a coordinated scan operated by a botnet, which started on January 31 (based on global traffic to the telescope) and lasted several days [27]. It looks like the Egyptian hosts infected by the botnet lost communication with the botnet control channel during the outage, but after BGP connectivity returned, they started to participate in the coordinated scan. The interesting insight is that scanning activities under botnet control cannot operate in the absence of bidirectional connectivity (since the bots cannot communicate with their controller), but random scans from worm-infected hosts still do and are still visible by the telescope when the senders are not BGP-reachable but still connected to the network. Such gaps between what the telescope can see globally versus from a specific country can help define criteria for the automated detection and classification of such events.

More than 3165 IPv4 prefixes are delegated to Egypt and managed by 51 ASs. In order to sketch per-AS observations by the network telescope, we classify the traffic observed from Egyptian IP addresses by the AS responsible for ("originating") the IP addresses. Fig. 7 shows the packet rate of traffic observed by the telescope from two major ASs in Egypt: EgStateAS,

---

[3] AS numbers anonymized; see the Appendix.

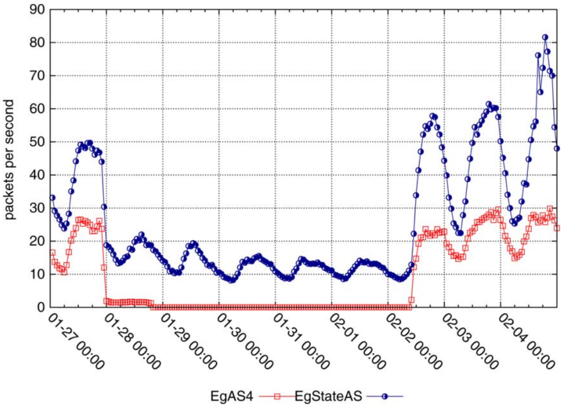

Fig. 7. Unsolicited packets from IPs geolocated in Egypt to UCSD network telescope: EgAS4, EgStateAS. Traffic from EgStateAS is still significant during the outage because: 1) some prefixes remain visible; 2) some networks probably retain outbound connectivity. The decay observable in the first days of the outage matches the progressive withdrawal of further routes.

and EgAS4. EgStateAS is the largest Egyptian Internet service provider.

Fig. 5 shows that many of the prefixes announced by EgStateAS via BGP were withdrawn on or shortly after January 27. A small set of IPv4 prefixes remained visible during the outage, including some not announced in the previous months. For this reason, we still observe darknet traffic coming from this AS, whereas the prefixes of EgAS4 were all withdrawn. A closer look at EgStateAS reveals that several of the visible IPv4 prefixes were reachable through IntAS2 or IntAS3, either because they were already served by those two autonomous systems or they rerouted to paths using those ASs after the massive disconnection.

Finally, we ran *ad hoc* active measurements during the outage to some related prefixes. In particular, we sent ICMP echo requests on February 1 at 09:00:00 GMT from GARR (the Italian Research and Academic Network), the replies to which revealed that at least three IPv4 prefixes, among those announced by EgStateAS and not withdrawn during the outage, were actually reachable. Traceroute probes simultaneously issued toward the same destinations went through IntAS2.

Another interesting case is that of EgAS7. As also reported by Renesys [20], the 83 prefixes managed by this AS remained untouched for several days during the Egyptian Internet outage. There was speculation that this AS retained Internet connectivity due to its high-profile, economically relevant customers, including the Egyptian stock exchange, the National Bank of Egypt, and the Commercial International Bank of Egypt. However, at a certain point, the censorship was tightened in Egypt: We observed the withdrawals of all 83 prefixes, almost simultaneously, on Monday, January 31, 20:46:48 GMT until the end of the outage, when all the Egyptian routes were restored. Fig. 8 shows a perfect match between our telescope observation of Egyptian traffic from EgAS7 and the BGP reachability of its prefixes.

Fig. 9 shows what percentage of active measurements from Ark and iPlane infrastructures reached a responding target geolocated to Egypt. Data from both infrastructures show a clear drop in bidirectional reachability of targets in Egypt during the outage. Examination of the specific IP addresses that retained bidirectional connectivity throughout the outage confirms they are all part of prefixes that were not withdrawn from the global routing table.

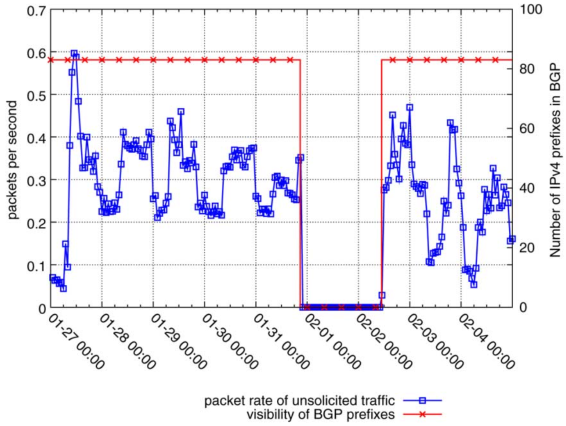

Fig. 8. Case of EgAS7: a perfect match across data sources. Unsolicited traffic to UCSD's network telescope versus BGP reachability of its 83 prefixes.

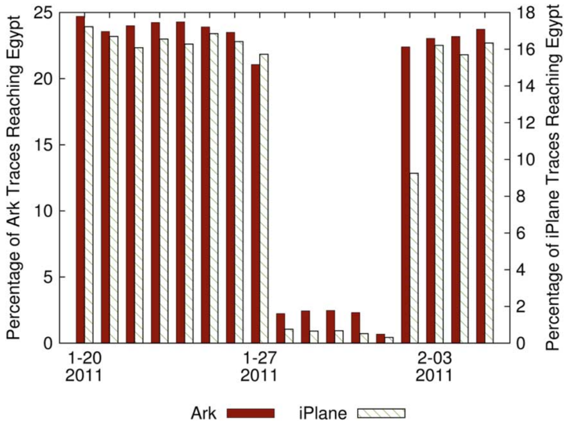

Fig. 9. Fraction of Ark and iPlane traceroutes, directed to IP addresses geolocated to Egypt, that terminated (either at the destination or the last reachable hop) in Egypt. The few Egyptian IP addresses that were seen in traceroutes throughout the outage were in BGP prefixes that were not withdrawn.

The absolute percentage of hosts for which we detected bidirectional connectivity differs between iPlane and Ark. Both the differences of absolute signal and relative drop (i.e., the ratio between before and during outage) could be explained by the differing measurement methodologies in Ark and iPlane. iPlane targets the .1 address in every routed IPv4 prefix in BGP snapshots, while Ark probes a random IP address in every /24 of the globally routed IPv4 prefixes.

At the end of the outage, a steady reconnection is observed via BGP. Figs. 4 and 10 respectively show time-series of BGP announcements in aggregate and for each of the six larger ASs. Fig. 10 shows each AS reinjecting sets of previously withdrawn routes, with most of them globally visible within 20 min. The process began with a first step at 09:29:31 GMT; by 09:56:11 GMT more than 2500 Egyptian IPv4 prefixes are back in BGP tables around the world. BGP data suggest that the key decisions on the outage were quite synchronized and produced dramatic globally observable consequences.

3) Denial-of-Service Attacks: Analysis of the UCSD darknet traffic also allowed us to identify some denial-of-service attacks to institutional sites of the Egyptian government, which, because of the timing and victims, look strongly related to protests in the country. The Web site of the Ministry of Communications

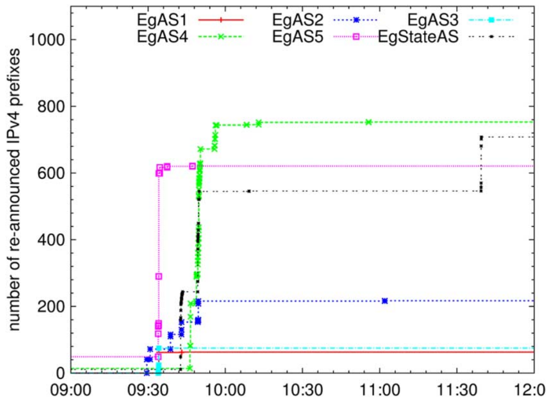

Fig. 10. Reconnection of main Egyptian autonomous systems via BGP at the end of outage on February 2, based on data from Route Views and RIPE NCC's RIS. Each AS is plotted independently; as in Fig. 4, each line rises at the instant in which a stable BGP announcement from that AS is first observed.

(mcit.gov.eg) was attacked with a randomly spoofed DoS attack just before the outage started, on January 26 at different times: 15:47 GMT (for 16 min), 16:55 GMT (17 min), and 21:09 GMT (53 min). Analysis of the *backscatter* traffic to the darknet allows estimation of the intensity of the attack in terms of packet rate, indicating average packet rates between 20k and 50k packets per second.

On February 2, the Web site of the Egyptian Ministry of Interior (www.moiegypt.gov.eg) was targeted by two DoS attacks just after the end of the censorship from 11:05 to 13:39 GMT and from 15:08 to 17:17 GMT. The same IP address was attacked another time the day after, from 08:06 to 08:42 GMT. In this case, the estimated packet rates were smaller, around 7k packets per second.

B. Libya

1) Overview: Libya's Internet infrastructure is even more prone to manipulation than Egypt's, judging from its physical structure. International connectivity is provided by only two submarine cables, both ending in Tripoli [50], and the Internet infrastructure is dominated by a single state-owned AS. We only found two other ASs having a small presence in Libya, as described in Section V-B.2.

In Libya, three different outages in early 2011 were identified and publicly documented (Fig. 1). Fig. 11 shows the traffic observed by the UCSD network telescope from Libya throughout an interval encompassing the outages. The points labeled A, B, and C indicate three different blackout episodes; points D1 and D2 refer to two denial-of-service attacks discussed in Section V-B.3. Toward the right of the graph, it is difficult to interpret what is really happening in Libya because of the civil war.

2) Outages in Detail: The first two outages happened during two consecutive nights. Fig. 12(a) shows a more detailed view of these two outages as observed by the UCSD telescope. Fig. 12(b) shows BGP data over the same interval: In both cases, within a few minutes, 12 out of the 13 IPv4 prefixes associated with IP address ranges officially delegated to Libya were withdrawn. These 12 IPv4 prefixes were announced by LyStateAS, the local telecom operator, while the remaining IPv4 prefix was managed by IntAS2. As of May 2011, there were no IPv6 prefixes in AfriNIC's delegated file for Libya.

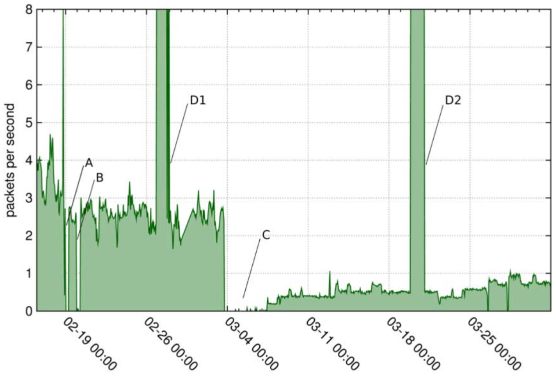

Fig. 11. UCSD darknet's traffic coming from Libya. Labels A, B, C indicate the three outages. Spikes labeled D1 and D2 are due to backscatter from two denial-of-service attacks.

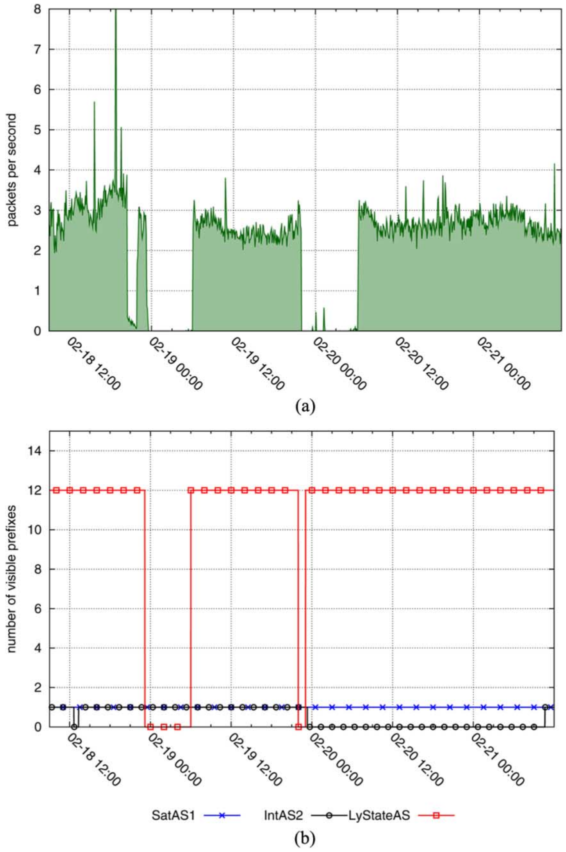

Fig. 12. First two Libyan outages: (a) unsolicited traffic to UCSD darknet coming from Libya; (b) visibility of Libyan IPv4 prefixes in BGP data from Route Views and RIPE NCC RIS collectors. Note that the control-plane and data-plane observations of connectivity do not match, suggesting that different techniques for censorship were being used during different intervals.

The MaxMind IP geolocation database further puts 12 noncontiguous IP ranges in Libya, all part of an encompassing IPv4 prefix announced by SatAS1, which provides satellite services in the Middle East, Asia, and Africa. The covering IPv4 prefix also contained 180 IP ranges in several other countries predominantly in the Middle East. We considered this additional AS because the UCSD darknet generally observed a significant amount of unsolicited traffic coming from IPs in those 12 ranges before the first outage (about 50k packets each day). This level of background traffic indicates a population of customers using PCs likely infected by Conficker or other

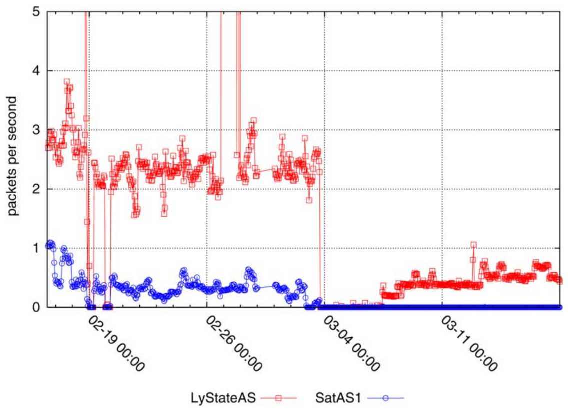

Fig. 13. UCSD darknet's traffic coming from Libya: traffic from selected ASs. The connectivity of satellite-based provider SatAS1 was probably disrupted through deliberate jamming of the satellite signal.

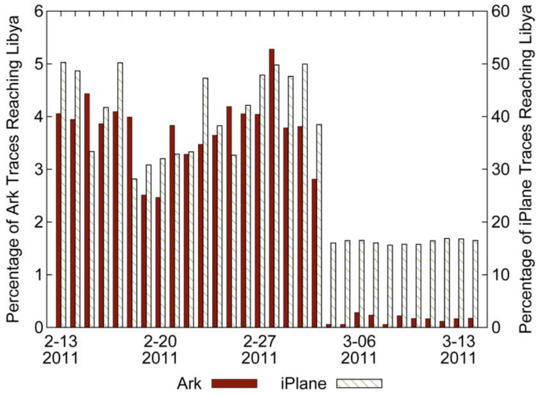

Fig. 14. Fraction of Ark and iPlane traceroutes, directed to IP addresses geolocated to Libya, which terminated (either at the destination or the last reachable hop) in Libya.

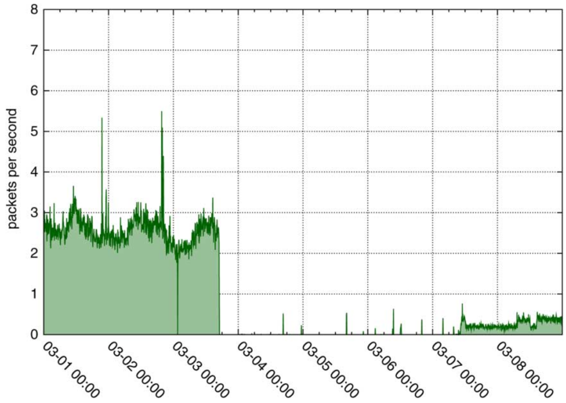

Fig. 15. UCSD darknet's traffic coming from Libya: detail of the third outage. The small but visible amount of traffic during the third outage (coming from a small number of /24 networks) is consistent with the use of selective packet filtering, instead of BGP withdrawals, to effect the outage.

malware, allowing inference of network conditions. Traffic from this network also provided evidence of what happened to Libyan Internet connections based on satellite systems not managed by the local telecom provider.

Comparing Fig. 12(a) and (b) reveals a different behavior that conflicts with previous reports [21]: The second outage was not entirely caused by BGP withdrawals. The BGP shutdown began on February 19 around 21:58.55 UTC, exactly matching the sharp decrease of darknet traffic from Libya (and in accordance with reports on Libyan traffic seen by Arbor Networks [40]), but it ended approximately 1 h later, at 23:02.52. In contrast, the Internet outage as shown by the telescope data and reported by the news [21] lasted until approximately February 20 at 6:12 UTC. This finding suggests that a different disruption technique—a packet-blocking strategy apparently adopted subsequently in the third outage and recognized by the rest of the world—was already being used during this second outage. The firewall configuration may have been set up during the BGP shutdown, and the routes were restored once the packet blocking was put in place.

Fig. 12(b) shows that the IPv4 prefix managed by SatAS1, the satellite company, was not withdrawn, which seems reasonable considering that this IPv4 prefix was managed by a company outside of Libya. However, the darknet traffic from both the local telecom and SatAS1 plummeted when the two outages occurred (see Fig. 13). A tiny amount of traffic still reached UCSD's darknet from SatAS1 IPs in Libya, especially during the second outage (Fig. 13), suggesting that the government could have used signal jamming to disrupt the satellite service for Internet connectivity, as they did for satellite TV news and mobile communication services [59], [66].

As for IntAS2, there was not enough unsolicited traffic reaching the darknet preceding and during the outages to usefully analyze, likely due to lack of end-users in this network. However, the only Libyan IPv4 prefix announced by IntAS2 was withdrawn twice: 1) on the same day of the first outage, but several hours before it started (for approximately 40 min, from 12:38.58 to 12:41.25 UTC); 2) approximately 10 min after the BGP routes of the local telecom were withdrawn in the second outage. The matching times in the latter case suggest a form of coordination or forcing the common loss of BGP connectivity. Fig. 12(b) shows that the BGP disruption of the Libyan IPv4 prefix of IntAS2 lasted for about 2 days (from February 19, 23:20.22 UTC, to February 21, 10:38.15 UTC), far longer than the duration of the second outage.

The third outage in Libya happened several days later. We verified, by analyzing all BGP updates collected by Route Views and RIPE NCC RIS, that all BGP routes stayed up without interruption. However, Fig. 15 shows that the darknet traffic sharply dropped at March 3, 16:57:00 UTC. Perhaps not surprisingly given their earlier experimenting with different censorship techniques, the third and longest Libyan outage was not caused by BGP disruption, but by packet filtering, confirmed by other sources [22].

While probing in the Ark data was not frequent enough to see the first two Libyan outages, the third and longer outage caused a significant drop in the fraction of reachable destinations in IPv4 prefixes geolocated in Libya, as seen in Fig. 14. The remaining reachable destinations in Libya were both from wired and satellite-connected ASs, showing that bidirectional communication for some hosts in both types of networks was still possible during this longer outage.

Our analysis revealed three discoveries.

- We established the potential of network telescopes to detect countrywide filtering phenomena, even phenomena that cannot be detected by monitoring BGP connectivity. The sharp decrease in traffic shown in Fig. 15 suggests that a simple change point detection algorithm would automatically raise an alert in this case, similar to how others used sharp drops in observed BGP announcements.

- We confirmed that packet filtering techniques for censorship were used because we still had visibility of a few packets from a few subnets, suggesting that perhaps the regime wanted to preserve connectivity for some sites.
- We discovered that packet filtering techniques were also used for previous outages that were reported as BGP-only disruptions. Moreover, we captured a retrospective of what happened, also explaining the short gap (February 18 from 20:24 to 21:57 UTC) in traffic visible on the UCSD network telescope [Fig. 12(a)] a few hours before the first outage. This short drop in traffic was also visible in data published by others [40], but was never discussed. We verified [Fig. 12(b)] that all the BGP routes were up during this gap in observed traffic, which suggests that Libya was already testing firewall blocking during this interval. The fact that the first two outages were BGP-based may indicate that the censors were unsatisfied with the results of these tests and used BGP-based disruption for the first two outages as an alternative while they further tested packet filtering techniques.

3) Denial-of-Service Attacks: In addition to reflecting the outages, our analysis of the UCSD darknet traffic shows two denial-of-service attacks to systems located in Libya. Because these attacks used randomly spoofed source addresses, we do not know if the attackers were inside or outside the country (or both), or how many machines were used to source the attack. The first attack, labeled D1 in Fig. 11, started on February 26 at 20:27 UTC, targeted a few IPs in a subnet of the Libyan Telecom (ltt.ly), and lasted approximately 24 h and 19 min. Analysis of the *backscatter* traffic allows us to estimate an average packet rate of 30 390 packets per second.

The second attack, labeled D2, started on Saturday, March 19, 2011, at 20:31 UTC, and the victim was a single IP assigned to the Libyan telecom provider. The attack lasted for about 27 h and 51 min with an estimated average packet rate of 30 280 packets per second.

VI. DISCUSSION AND CONCLUSION

Political events in the Middle East in 2011, as well as political discussions in the US Congress [67] have inspired popular as well as technical interest in possible mechanisms, impact, circumvention, and detection of Internet filtering at different layers. Our study of Egypt's and Libya's government-ordered Internet outages have revealed a number of challenges and opportunities for the scientific study of Internet filtering and disruption. Given the growing interest and expanding circumstances that will give rise to large-scale Internet filtering behaviors, and the need to inform policy development with the best available empirical data and analysis of such behavior, we believe the topic will necessarily merit its own discipline. This study offers an initial contribution in this direction.

We used multiple types of large-scale data in this analysis, all from data sets already available to academic researchers. The first type of data—BGP interdomain routing *control plane* data—was already widely analyzed and reported on during the outages. Our analysis of BGP data suggested that key decisions related to the outage were quite synchronized and produced dramatic, globally observable consequences.

We have not seen the second type of data—unsolicited *data plane* traffic to unassigned address space (darknet or telescope data)—previously used for this purpose, and we were surprised at the range of insights it yielded. Unsolicited and unwanted traffic on the Internet has grown to such significant levels that instrumentation capturing such traffic can illuminate many different types of macroscopic events, including but not limited to broad-scale packet-filtering-based censorship, which is not observable in BGP data. From this unidirectional traffic data, we detected what we believe were Libya's attempts to test firewall-based blocking before they executed more aggressive BGP-based disconnection. These data also revealed Libya's use of such packet filtering technology during the second BGP-based connection. Interestingly, the backscatter component of these traffic data enabled us to identify some denial-of-service attacks against Egyptian government Web sites before and after the censorship interval.

We also made limited use of active ping and macroscopic traceroute measurements toward address space in these countries during the outages. We used CAIDA's IPv4 topology data set and iPlane traceroute measurements to observe, surprisingly, a limited amount of two-way connectivity surviving the outage intervals that span more then a day.

Using both *control plane* and *data plane* data sets in combination allowed us to narrow down which form of Internet access disruption was implemented at different times in a given region. Our methodology required determining which IP address prefixes were in each country using RIR-delegation data and public geolocation database (MaxMind) data, and then mapping those prefixes of interest to origin ASs using publicly available BGP data repositories in the US and Europe. Looking deeper into all sources of data can reveal different filtering approaches, possible satellite jamming, ranges of IPs not filtered by the firewall, and different forms of (or lack of) coordination with other authorities. These techniques could also be used to improve the accuracy of geolocation databases, e.g., detecting errors in geolocation databases that map IP addresses to completely censored countries, but such IP addresses still show up in measurements.

Since IPv6 is not as widely deployed as IPv4, there is a lack of feature parity in IPv4 and IPv6 technologies, including censorship technologies such as deep packet inspection. This disparity means that IPv6 may offer a time-limited opportunity for evasion of layer-3 IP censorship. The fact that all IPv6 prefixes in Egypt were unaffected by the outage, shows that currently data flows using IPv6 are easily overlooked or ignored. Whether it was a willful act of disobedience by the network operators or (more likely) an oversight, the effect is the same: Unless or until IPv6 gains considerably more traction, IPv6 data flows may remain "under the radar" (or more accurately, may continue to slip through the firewall).

Additional data sources would deepen the analysis, especially since only a subset of data sources may be available at any time. We used Ark and iPlane data as supporting evidence for bidirectional layer-3 reachability of prefixes in Egypt and Libya, but the same data also contain forward path information that we have not explored in detail. These data could allow us to see path changes at the data plane that happened due to the censorship we described. Comparing path changes observed in the data plane to those observed in the control plane, i.e., BGP, may expose non-BGP routing phenomena, such as use of default routes in addition to BGP-based routing [15].

Automating this relatively manual post-event analysis methodology involves several technical and intellectual challenges we are now pursuing.[4] In designing a prototype system for the continuous and combined monitoring of extracted metrics, we found we need to first detect outages within national boundaries since IP geolocation at a finer granularity is still much less accurate by the best available methods.

Effective distributed and continuous active probing also requires properly managing the inherent tradeoff between temporal granularity, address space and geographic scope by the measurements, and computational and network traffic overhead. We are evaluating an approach to optimizing this tradeoff using precompiled hitlists of responsive probing targets from the ISI Internet Address Census project [35].

Analysis of darknet traffic is also challenging to automate, although we have found that observing the number of distinct source IP addresses per unit time sending packets to the darknet can effectively detect and quantify the impact of macroscopic Internet outages [26]. Unlike packet rate, this metric is not significantly affected by backscatter or other anomalies, although it is susceptible to spoofing of source IP addresses. We are currently developing methods to automatically filter spoofed traffic reaching our darknet.

Combining metrics from the three different types of measurement sources used in this paper will enable diagnosis of connectivity disruption happening at different layers of the protocol stack. The resolution of BGP updates and darknet traffic data offers finer time granularity, which facilitates early detection of anomalies that could trigger further active probing to provide additional timely insight into the event. We plan to leverage on-demand active probing capabilities developed for our related Internet infrastructure mapping project [18].

Real-time monitoring supporting geographical regions not delimited by national borders poses additional challenges. We are investigating approaches to geolocate IP prefixes advertised on the BGP plane with higher granularity than country-level, in order to create a common aggregation set with active and passive measurements on the data plane. More importantly, the need to identify an optimal geographical granularity introduces another dimension to the challenges we mentioned, such as the tradeoffs in continuous probing. Additionally, since our metrics extracted from darknet traffic are essentially based on opportunistic measurements, their applicability is a function of the granularity of geographical aggregation (e.g., Internet penetration and density of infected hosts in the areas involved affect the "quality" of the signal). We are currently working on rigorously identifying the limits within it is possible to extract a stable signal from darknet traffic that is still effective to infer connectivity disruption in the monitored regions.

APPENDIX

We recognize that this paper could possibly expose network operators' disobedience to government orders, with potential harmful consequences to the operators. Therefore, we believe it would be ethical to anonymize the ASs that might be at risk for harm. Here, we describe and justify the reasoning for our anonymization policy.

Our main concern is that we do not know the nature of the relationships between the operators of these networks to their governments, whether they disobeyed direct orders and what the penalties might be. We assume that the national incumbent telecom operators will follow orders from their respective governments; our research did not reveal behavior that suggests otherwise. Furthermore, the behavior of both incumbent telecom operators has been described in detail elsewhere, e.g., [20]. Hence, in this paper we label these ASs *EgStateAs* and *LyStateAs*. For the remaining ASs, we anonymized ASs that were operating nationally in Egypt with an *Eg*-prefix, e.g., *EgAs1*, *EgAs2*, etc. We did not need to do so for Libya because we did not find ASs operating at the national level in Libya, except for the state-owned telecom operator. As the upstreams of networks in Libya and Egypt may have personnel operating in one of these countries, we decided to anonymize these as well and labeled them *IntAs1*, *IntAs2*, etc. One distinguishing characteristic of censorship activity we observed is whether it occurs via wireless satellite connectivity; depending on who owns the satellite links, disrupting such connectivity may require jamming technology rather than firewall-based filtering. To acknowledge this distinction, which is revealed in the data anyway, we labeled the one international satellite connectivity provider we show data for as *SatAs1*. We realize this anonymization policy cannot guarantee to prevent deanonymization by careful examination of publicly available data, but it raises the threshold for doing so without compromising the scientific value of this paper.

REFERENCES

[1] AfriNIC, "AfriNIC: The registry of Internet number resources for Africa," [Online]. Available: http://www.afrinic.net

[2] CAIDA, La Jolla, CA, USA, Ark Measurement Infrastructure"," [Online]. Available: http://www.caida.org/projects/ark

[3] OpenNet Initialive, "OpenNet Initialive," [Online]. Available: http://opennet.net

[4] RIPE NCC, Amsterdam, The Netherlands, "RIPE NCC: BGPviz," [Online]. Available: http://www.ris.ripe.net/bgpviz/

[5] RIPE NCC, Amsterdam, The Netherlands, "RIPE NCC: REX, Resources EXplained," [Online]. Available: http://albatross.ripe.net/cgi-bin/rex.pl

[6] RIPE NCC, Amsterdam, The Netherlands, "RIPE NCC: Routing Information Service (RIS)," [Online]. Available: http://www.ripe.net/data-tools/stats/ris/routing-information-service

[7] RIPE NCC, Amsterdam, The Netherlands, "RIPEstat: Internet Measurements and Analysis," [Online]. Available: https://stat.ripe.net

[8] CAIDA, La Jolla, CA, USA, "UCSD Network Telescope," 2010 [Online]. Available: http://www.caida.org/data/passive/network_telescope.xml

[9] Freedom House, Washington, DC, USA, "Freedom on the net 2011: A global assessment of Internet and digital media," Apr. 2011 [Online]. Available: http://www.freedomhouse.org/report/freedom-net/freedom-net-2011

[10] E. Aben, "Conficker/Conflicker/Downadup as seen from the UCSD network telescope," 2008 [Online]. Available: http://www.caida.org/research/security/ms08-067/conficker.xml

[11] E. Aben, "Unsolicited Internet traffic from Libya," Mar. 2011 [Online]. Available: http://labs.ripe.net/Members/emileaben/unsolicited-internet-traffic-from-libya

[12] BGPmon, "Internet in Egypt offline," Jan. 2011 [Online]. Available: http://bgpmon.net/blog/?p=450

[13] Z. Bischof and J. Otto, "Egypt and Libya Internet disconnections," Mar. 10, 2011 [Online]. Available: http://www.aqualab.cs.northwestern.edu/blog/egypt-libya-peers.html

---

[4] This activity is supported from a recent grant from the US NSF's Security and Trustworthy Cyberspace (SATC) program (CNS-1228994) "Detection and analysis of large-scale Internet infrastructure outages."

[14] J. Burris, “Comcast: Internet service restored after regional outage,” Nov 29, 2010 [Online]. Available: http://articles.baltimoresun.com/2010-11-29/news/bs-md-comcast-20101128_1_outage-internet-service-disruptions

[15] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, “Internet optometry: Assessing the broken glasses in internet reachability,” in *Proc. ACM SIGCOMM Conf. Internet Meas.*, 2009, pp. 242–253.

[16] M. Casado, T. Garfinkel, W. Cui, V. Paxson, and S. Savage, “Opportunistic measurement: Spurious network events as a light in the darkness,” in *Proc. 4th ACM HotNets-IV*, New York, NY, USA, 2005.

[17] R. Clayton, S. Murdoch, and R. Watson, “Ignoring the great firewall of China,” in *Proc. Privacy Enhancing Technol. Workshop*, 2006, pp. 4258:20–4258:35.

[18] Cooperative Association for Internet Data Analysis, La Jolla, CA, USA, “Cartographic capabilities for critical cyberinfrastructure,” 2013 [Online]. Available: http://www.caida.org/funding/c4/

[19] J. Cowie, “Strange changes in Iranian transit,” Jun 14, 2009 [Online]. Available: http://www.renesys.com/blog/2009/06/strange-changes-in-iranian-int.shtml

[20] J. Cowie, “Egypt leaves the Internet,” Jan. 27, 2011 [Online]. Available: http://www.renesys.com/blog/2011/01/egypt-leaves-the-internet.shtml

[21] J. Cowie, “Libyan disconnect,” Feb. 2011 [Online]. Available: http://www.renesys.com/blog/2011/02/libyan-disconnect-1.shtml

[22] J. Cowie, “What Libya learned from Egypt,” Mar. 2011 [Online]. Available: http://www.renesys.com/blog/2011/03/what-libya-learned-from-egypt.shtml

[23] J. Cowie, A. Ogielski, B. Premore, E. Smith, and T. Underwood, “Impact of the 2003 blackouts on Internet communications,” Nov. 2003.

[24] J. Cowie, A. Ogielski, B. Premore, and Y. Yuan, “Internet worms and global routing instabilities,” in *Proc. SPIE Int. Symp. Convergence IT Commun.*, Jul. 2002.

[25] J. Crandall, D. Zinn, M. Byrd, E. Barr, and R. East, “ConceptDoppler: A weather tracker for internet censorship,” in *Proc. 14th ACM Conf. Comput. Commun. Security*, 2007, pp. 352–365.

[26] A. Dainotti, R. Amman, E. Aben, and K. C. Claffy, “Extracting benefit from harm: Using Malware pollution to analyze the impact of political and geophysical events on the internet,” *Comput. Commun. Rev.*, vol. 42, no. 1, pp. 31–39, 2012.

[27] A. Dainotti, A. King, K. Claffy, F. Papale, and A. Pescapè, “Analysis of a “/0” stealth scan from a botnet,” in *Proc. ACM IMC*, New York, NY, USA, 2012, pp. 1–14.

[28] A. Dainotti, A. Pescapé, and G. Ventre, “Worm traffic analysis and characterization,” in *Proc. IEEE ICC 2007*, Jun. 2007, pp. 1435–1442.

[29] A. Dainotti, C. Squarcella, E. Aben, K. C. Claffy, M. Chiesa, M. Russo, and A. Pescapé, “Analysis of country-wide internet outages caused by censorship,” in *Proc. ACM SIGCOMM IMC*, New York, NY, USA, 2011, pp. 1–18.

[30] S. Deshpande, M. Thottan, T. K. Ho, and B. Sikdar, “An online mechanism for BGP instability detection and analysis,” *IEEE Trans. Comput.*, vol. 58, no. 11, pp. 1470–1484, Nov. 2009.

[31] S. Garret, “We can confirm that Twitter was blocked in Egypt around 8am PT today. . .,” Jan. 25, 2011 [Online]. Available: http://twitter.com/#!/twitterglobalpr/status/30063209247408128

[32] E. Glatz and X. Dimitropoulos, “Classifying internet one-way traffic,” in *Proc. ACM IMC*, New York, NY, USA, 2012, pp. 37–50.

[33] Google, Mountain View, CA, USA, “Google transparency report,” [Online]. Available: http://www.google.com/transparencyreport/traffic/?r=LY&l=YOUTUBE&csd=1296723600000&ced=1299142800000

[34] Google, Mountain View, CA, USA, “Google transparency report,” [Online]. Available: http://www.google.com/transparencyreport/traffic/?r=LY&l=EVERYTHING&csd=1296862200000&ced=1299281400000

[35] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos, G. Bartlett, and J. Bannister, “Census and survey of the visible internet,” in *Proc. 8th ACM SIGCOMM IMC*, 2008, pp. 169–182.

[36] B. Huffaker, M. Fomenkov, and K. C. Claffy, “Geocompare: A comparison of public and commercial geolocation databases,” in *Proc. NMMC*, May 2011 [Online]. Available: http://www.caida.org/publications/papers/2011/geocompare-tr/

[37] Y. Hyun, B. Huffaker, D. Andersen, E. Aben, C. Shannon, M. Luckie, and K. C. Claffy, “The IPv4 routed /24 topology dataset,” 2011 [Online]. Available: http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml

[38] E. Katz-Bassett, H. V. Madhyastha, J. P. John, A. Krishnamurthy, D. Wetherall, and T. Anderson, “Studying black holes in the internet with hubble,” in *Proc. USENIX NSDI*, 2008, pp. 247–262.

[39] R. Khosla, S. Fahmy, and Y. C. Hu, “BGP molecules: Understanding and predicting prefix failures,” in *Proc. IEEE INFOCOM Mini-Conf.*, 2011, pp. 146–150.

[40] C. Labovitz, “Libya firewall begins to crumble?,” Feb. 2011 [Online]. Available: http://monkey.org/~labovit/blog//viewpage.php?page=libya_firewall_crack

[41] C. Labovitz, “Egypt loses the Internet,” Jan. 28, 2011 [Online]. Available: http://monkey.org/~labovit/blog/viewpage.php?page=egypt_loses_internet

[42] C. Labovitz, “Egypt returns to the Internet,” Feb. 2011 [Online]. Available: http://asert.arbornetworks.com/2011/02/egypt-returns-to-the-internet/

[43] C. Labovitz, “Middle East Internet scorecard (February 12–20),” Feb. 2011 [Online]. Available: http://arbornetworks.com/asert/2011/02/middle-east-internet-scorecard-february-12-20/

[44] C. Labovitz, G. R. Malan, and F. Jahanian, “Origins of internet routing instability,” in *Proc. IEEE INFOCOM*, 1999, pp. 218–226.

[45] C. Labovitz, G. R. Malan, and F. Jahanian, “Internet routing instability,” *IEEE/ACM Trans. Netw.*, vol. 6, no. 5, pp. 515–528, Oct. 1997.

[46] S. LaPerriere, “Taiwan earthquake fiber cuts: A service provider view,” presented at the NANOG 39, Feb. 2007.

[47] J. Li and S. Brooks, “I-seismograph: Observing and measuring internet earthquakes,” in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2624–2632.

[48] Y. Liu, X. Luo, R. K. C. Chang, and J. Su, “Characterizing inter-domain rerouting after Japan earthquake,” in *Proc. Netw.*, 2012, vol. 2, pp. 124–135.

[49] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, “iPlane: An information plane for distributed services,” in *Proc. USENIX OSD*, Berkeley, CA, USA, 2006, pp. 367–380.

[50] G. Mahlknecht, “Greg’s cable map,” Retrieved Sep. 12, 2011 [Online]. Available: http://www.cablemap.info/

[51] MaxMind, “MaxMind GeoLite country: Open source IP address to country database,” 2013 [Online]. Available: http://www.maxmind.com/app/geolitecountry

[52] Ministry of Communications and Information Technology, Arab Republic of Egypt, “ICT indicators in brief,” Feb. 2011 [Online]. Available: http://www.mcit.gov.eg/Upcont/Documents/ICTinBriefFeb2011-E.pdf

[53] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, “Inside the slammer worm,” *IEEE Security Privacy*, vol. 1, no. 4, pp. 33–39, Jul. 2003.

[54] D. Moore, C. Shannon, D. J. Brown, G. M. Voelker, and S. Savage, “Inferring internet denial-of-service activity,” *Trans. Comput. Syst.*, vol. 24, pp. 115–139, May 2006.

[55] S. Nasrawi, “Libya protests: Anti-government protesters killed during day of rage,” Feb. 2011 [Online]. Available: http://www.huffingtonpost.com/2011/02/17/libya-protests-antigovern_0_n_824826.html

[56] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson, “Characteristics of internet background radiation,” in *Proc. 4th ACM SIGCOMM IMC*, New York, NY, USA, 2004, pp. 27–40.

[57] L. Peterson, A. Bavier, M. E. Fiuczynski, and S. Muir, “Experiences building planetlab,” in *Proc. USENIX OSDI*, Berkeley, CA, USA, 2006, pp. 351–366.

[58] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye, “IP geolocation databases: Unreliable?,” *Comput. Commun. Rev.*, vol. 41, no. 2, pp. 53–56, Apr. 2011.

[59] “Thuraya satellite telecom says jammed by Libya,” Feb. 2011 [Online]. Available: http://af.reuters.com/article/libyaNews/idAFLDE71N2CU20110224

[60] “Egypt govt denies disrupting websites—Cabinet,” Jan 26, 2011 [Online]. Available: http://www.reuters.com/article/2011/01/26/egypt-web-idUSLDE70P28720110126

[61] “Facebook says has seen drop in traffic from Egypt,” Jan 2011 [Online]. Available: http://www.reuters.com/article/2011/01/27/facebook-egypt-idUSN2727880720110127

[62] RIPE, Amsterdam, The Netherlands, “Analysis of Egyptian Internet outage 27th January–2nd February 2011,” 2011 [Online]. Available: http://stat.ripe.net/egypt

[63] A. Sahoo, K. Kant, and P. Mohapatra, “Characterization of BGP recovery time under large-scale failures,” in *Proc. IEEE Int. Conf. Commun.*, 2006, vol. 2, pp. 949–954.

[64] Y. Shavitt and N. Zilberman, “A geolocation databases study,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 10, pp. 2044–2056, Dec. 2011.

[65] M. Skoric, N. Poor, Y. Liao, and S. Tang, "Online organization of an offline protest: From social to traditional media and back," in *Proc. 44th HICSS*, Jan. 2011, pp. 1–8.

[66] "Libya source of jamming of Lebanese news channels: TRA," Feb. 23, 2011 [Online]. Available: http://www.dailystar.com.lb/News/ Politics/ Feb/23/Libya-source-of-jamming-of-Lebanese-news-channels-TRA. ashx#axzz1Ev3JP000

[67] J. Swartz, "'Kill switch' Internet bill alarms privacy experts," Feb 15, 2011 [Online]. Available: http://usatoday30.usatoday.com/tech/news/ internetprivacy/2011-02-15-kill-switch_N.htm

[68] S. T. Teoh, S. Ranjan, A. Nucci, and C.-N. Chuah, "BGP eye: A new visualization tool for real-time detection and analysis of BGP anomalies," in *Proc. 3rd ACM VizSEC*, New York, NY, USA, 2006, pp. 81–90.

[69] T. Underwood, "Con-Ed steals the 'net," Jan. 22, 2006 [Online]. Available: http://www.renesys.com/2006/01/coned-steals-the-net

[70] R. T. University, "BGPlay—Graphical visualisation of BGP updates," [Online]. Available: http://bgplay.routeviews.org/

[71] University of Oregon, Eugene, OR, USA, "University of Oregon Route Views Project," [Online]. Available: http://www.routeviews.org

[72] I. van Beijnum, "How Egypt did (and your government could) shut down the Internet," Feb. 2011 [Online]. Available: http:/ /arstechnica.com/tech-policy/news/2011/01/howegypt-or-how-your-government-could-shut-down-the-internet.ars

[73] T. Wan and P. van Oorschot, "Analysis of BGP prefix origins during Google's May 2005 outage," in *Proc. Parallel Distrib. Process. Symp.*, Apr. 2006, pp. 25–29.

[74] "Ramses Exchange," Sep. 12, 2011 [Online]. Available: http://en.wikipedia.org/wiki/Ramses_Exchange

[75] B. Woodcock, "Overview of the Egyptian Internet shutdown," Feb. 2011 [Online]. Available: http://www.pch.net/resources/misc/Egypt-PCH-Overview.pdf

[76] X. Xu, Z. M. Mao, and J. A. Halderman, "Internet censorship in China: Where does the filtering occur?," in *Proc. PAM*, 2011, pp. 133–142.

[77] H. Yan, R. Oliveira, K. Burnett, D. Matthews, L. Zhang, and D. Massey, "BGPmon: A real-time, scalable, extensible monitoring system," in *Proc. IEEE Conf. Cybersecurity Appl. Technol. Conf. Homeland Security*, 2009, pp. 212–233.

[78] E. Zmijewski, "Georgiaclings to the 'net," Aug. 10, 2008 [Online]. Available: http://www.renesys.com/blog/2008/08/georgia-clings-to-the-net.shtml

[79] E. Zmijewski, "Mediterranean cable break," Jan. 30, 2008 [Online]. Available: http://www.renesys.com/blog/2008/ 01/mediterranean_cable_break.shtml

[80] E. Zmijewski, "Mediterranean cable break—Part II," Jan 2008 [Online]. Available: http://www.renesys.com/blog/2008/ 01/mediterranean_cable_break_part_1.shtml

[81] E. Zmijewski, "Mediterranean cable break—Part III," Feb 2008 [Online]. Available: http://www.renesys.com/blog/2008/ 02/mediterranean_cable_break_part.shtml

Alberto Dainotti received the Ph.D. degree in computer engineering and systems from the University of Napoli "Federico II," Naples, Italy, in 2008.

He is a Research Scientist with the Cooperative Association for Internet Data Analysis (CAIDA), SDSC, University of California, San Diego, La Jolla, CA, USA. His main research interests are in the field of Internet measurement and network security, with a focus on the analysis of large-scale Internet events.

Dr. Dainotti was awarded the IRTF Applied Networking Research Prize in 2012.

Claudio Squarcella is currently pursing the Ph.D. degree in engineering at Roma Tre University, Rome, Italy.

He previously worked with the RIPE NCC, Amsterdam, The Netherlands, and the Cooperative Association for Internet Data Analysis (CAIDA), La Jolla, CA, USA. His current research interests include network visualization and graph drawing.

Emile Aben received the Master's degree in chemistry from the Radboud University Nijmegen, Nijmegen, The Netherlands, in 1998.

He is a System Architect with the research and development group of the RIPE NCC, Amsterdam, The Netherlands. Before that, he worked with CAIDA, La Jolla, CA, USA. Since then, he has been working on Internet-related things, as a System Administrator, Security Consultant, Web Developer, and Researcher. His research interests include network measurement, technology changes (IPv6), and network security.

Kimberly C. Claffy received the Ph.D. degree in computer science from the University of California (UC), San Diego, La Jolla, CA, USA, in 1994.

She is Director of the Cooperative Association for Internet Data Analysis (CAIDA), which she founded at the UC San Diego Supercomputer Center in 1996. CAIDA provides Internet measurement tools, data, analyses, and research to promote a robust, scalable global Internet infrastructure. She has been at SDSC since 1991 As a Research Scientist with the SDSC and Adjunct Professor of computer science and engineering at UCSD, her research interests include Internet (workload, performance, topology, routing, and economics) data collection, analysis, visualization, and enabling others to make use of CAIDA data and results.

Marco Chiesa received the Master's degree in computer science from the Roma Tre University, Rome, Italy, in 2010, and is currently pursuing the Ph.D. degree in computer science and automation at the same university.

His research activity is primarily focused on routing protocols, including theoretical analysis of convergence, security, and traffic-engineering problems.

Michele Russo received his Master's degree in computer engineering from the University of Napoli "Federico II," Naples, Italy, in 2011.

He works in the field of network monitoring as an SQA Engineer with NetScout Systems, Modena, Italy. He also worked in the field of network monitoring for Accanto Systems, Modeno, Italy, and railway signaling for Ansaldo STS, Genova, Italy.

Antonio Pescapé (M'00–SM'09) received the M.S. Laurea degree in computer engineering and Ph.D. degree in computer engineering and systems from the University of Napoli Federico II, Naples, Italy, in 2000 and 2004, respectively.

He is an Assistant Professor with the Electrical Engineering and Information Technology Department, University of Napoli Federico II, and an Honorary Visiting Senior Research Fellow with the School of Computing, Informatics and Media, University of Bradford, Bradford, U.K. His research interests are in the networking field with focus on Internet monitoring, measurements and management, and network security.