Arabian Nights: Measuring the Arab Internet During the 2011 Events

Yuval Shavitt and Noa Zilberman, Tel-Aviv University, Israel

Abstract

The major turmoils in the Arab world since the beginning of 2011 were largely driven by social networks and are often referred to as the "Arab Spring." One of the methods used by rulers to mitigate the unrest is "shutting down" the Internet in their country. In this article we describe active measurements conducted during 2011 of several Arab countries, and analyze the changes in the network. These events provide a unique opportunity to measure features of the network that are otherwise hard to track, such as static or default BGP routes.

In December 2010, a wave of unrest shook the Arab world, starting in Tunisia and spreading to other Arab countries such as Egypt, Libya, Yemen, and Syria [1]. This unrest, often referred to as the Arab Spring, the Arab Awakening, or the Arab Revolt [1, 2], led in some of the countries to a change of regime. The events in the Arab world relied heavily on the Internet to incite people and coordinate their protest [3]. As a result, the oppressive regimes in these countries attempted to curb access to the Internet by various techniques. This presented us with a unique opportunity to study some aspects of the Internet that are otherwise hard to track.

Starting from the early events in Egypt, at the end of January 2011 and until May 2011, we conducted a large-scale traceroute measurement effort to several Arab countries. In this work we report our findings regarding the state of the Internet in Egypt, Libya, and Syria. The three countries show different approaches in their attempts to block access to the Internet, and thus give us a view of a wide range of such techniques. It is important to note that the three countries are far from homogeneous. While Egypt maintained fairly open access to the Internet before and after the peak of the unrest there, Syria has been tightly monitoring its citizens' access to the Internet for years, and in Libya the status is somewhere between the two. The countries also differ in the size and structure of their Internet: Egypt has a fairly large address space which is maintained by several operators, while Syria and Libya have a much smaller infrastructure. In Syria access is controlled by one government agency, and in Libya the main service provider had the President's son as a chairman.

In Egypt, the main method to disconnect the country's public Internet was by issuing Border Gateway Protocol (BGP) withdraw messages for large portions of the Egyptian address space. That is, the BGP protocol was told that there was no valid route to these portions of the address space. As a result, BGP routing tables, which are updated dynamically and age stale information, should quickly have no valid route to these destinations, and thus packets destined there should be dropped. This gave us a unique opportunity to examine the usage of default and static routing in the Internet, where routing tables to some (some in case of static routing, most in case of default routing) destinations are static. We also revealed many coherency problems in the routing tables of large providers. We report these findings as well in this article.

Data Set

The collected dataset for this work is taken from DIMES [4]. We use 3.63 million traceroute measurements from the end of January to the beginning of May 2011. The measurements were collected by 1137 DIMES agents, which are located in 74 countries around the world. About 16 percent of the agents are mobile.

The measurements are based on targeted experiments, separately running for each country, with the destination IP addresses selected by matching one of the following criteria: For the first, the IP address is located in the target country, based on Maxmind[1] or IPligence[2] Geolocation databass. For the second, the IP belongs to an AS that is registered in the target country. Each experiment was run daily by 100 agents, each was assigned to measure 1000 target IP addresses out of the pool of IPs described above. The measuring agents were altered every day, with the purpose of measuring to the same destinations through different paths, and with most agents taking part in the experiments repeatedly every few days.

Due to operational problems, during some of the days along the experiment period, measurements were not taken. These days are not taken into account in the analysis and do not affect the results.

Results

Reaching Destination

A traceroute measurement toward a destination is typically considered successful if the destination IP is reached. However, in our experiments the selection of target addresses is based on our attempt to cover address prefixes (APs), and the

[1] *Maxmind GeoIP, http://www.maxmind.com*

[2] *IPligence Max, http://www.ipligence.com*

0890-8044/12/$25.00

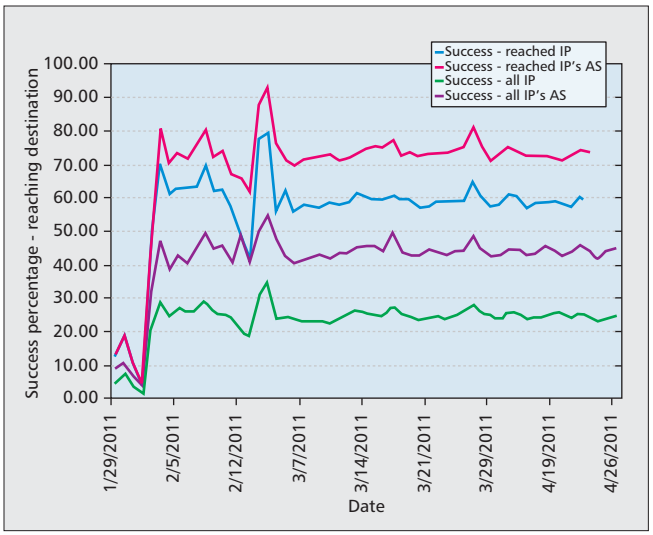

Figure 1. *Success rate in reaching destinations in Egypt, by IP and AS.*

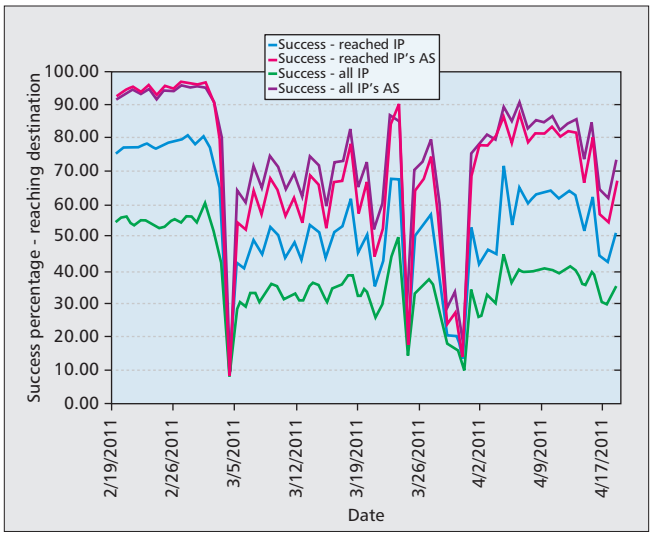

Figure 2. *Success rate in reaching destinations in Libya, by IP and AS.*

selection of an IP address in the AP is arbitrary. As a result, the selected IP address may not be active. Thus, we define a traceroute to be successful in the wide sense if the address of the last hop is within the target autonomous system (AS). In the following we use both the narrow (reaching the target IP) and wide sense measurement success definitions.

Egypt — The targeted measurements to Egypt show the Internet cutoff between January 28, a day after the government also closed the banks and cellular networks, and February 2, 2011, the bloodiest day in Tahrir Square to date, after which the ban was taken off [2]. During these dates, Fig. 1 shows the success rate for of the targeted measurements. While on average 25 percent of the measurements reach their final destination IP and 44 percent of them reach the destination AS, during the shutdown period only 4 percent of the measurements reached their destination IP, and 8 percent reached the destination AS. The low percentage of destination IPs reached may be attributed to bad selection of target IPs; thus, we selected the IP addresses that were reached at least once during the experiment's period. This group of IP addresses has an average of 60 percent of the traceroutes reaching their destination IP and 73 percent reaching the destination's AS. During the shutdown period, only 11 percent of these IP addresses were reached, and also only 11 percent of the destination AS. The ASs that are reached match ASs that continued to send their BGP updates, and whenever an AS stopped sending BGP messages (based on Route Views [5]), it was no longer reached. Among these ASs are Etisalat Egypt (AS36992), a class A service provider and NOOR (AS20928), a class B service provider.[3] We note that on February 12 through 15 there is a sudden drop and rise in the success rate. This was caused by insufficient DIMES measurements during these days, which distort the statistics.

Libya — The targeted measurements to Libya portray a different picture than in Egypt. While in Egypt the success rate in reaching destinations was rather steady, except for the shutdown period, in Libya there is a variation over time (Fig. 2): the average success rate in reaching destinations before March 5 is 54 percent in the narrow sense and 93 percent in the wide sense; after this date, the success percentage drops to 33 and 70 percent, respectively.

[3] *http://ntra.gov.eg/presentations/LicensedTelecomChart22122009_En.pdf*

Several Internet cutoff events were reported in Libya [6, 7]. The first, on February 19, a day after dozens of protesters were killed by security forces, lasted only six hours and did not significantly affect our results. A second cutoff event, on March 3, just as fighting in the city of Zawiyah intensified, did show a drastic effect on the results: only 8 percent success in the narrow sense and only 12 percent in the wide sense. Looking only at live IP addresses, the results are almost identical.

On March 19, coalition forces began military intervention in Libya. These actions initially targeted Libyan government ground forces, but spread to command and control installations. Indeed, it is easy to observe a drastic single day drop in traceroute success rates on March 25, and then again for a longer period at the end of March and the beginning of April. These dates correspond with heavy bombardments by NATO forces.

Syria — The Internet in Syria behaves differently than in Egypt and Libya. There are 12 service providers in Syria that are grouped under two ASs, AS29256 and AS29386; both belong to the Syrian Telecommunications Establishment (STE). The number of address ranges assigned to Syria is limited as well, which in turn limited the number of destinations measured by DIMES to around 50. The attempt to probe specific addresses in Syria proved more difficult than in other countries: only about 10 percent of the IP addresses were reached, and even when considering only the responding IP addresses, the traceroute success rate was only 30 percent. On the other hand, the success rate in reaching the destination AS is very high: about 88 percent on average. We also do not see Internet cutoffs: the behavior is quite constant throughout the measurement period (except for a slight increase in AS reach level), and even during February 5 and the following days, when the Internet in Syria was reported to be curbed by news agencies, there is no effect on traceroute measurements. This indicates that the nature of Syrian Internet blocking was likely based on site level, as indicated by Deibert [8]; thus, traceroutes may not have been affected. Figure 3 shows the routing success rate to Syria. Another aspect of interest is the difference in reaching different Internet service providers (ISPs) over the measurements' period. We observe a large variability between service providers, both over time and between ISPs, from a constant ISP reach rate to a bursty behavior, where the ISP is accessible on some days and unreachable on others. DIMES' traceroute measurements are created by a train of four consecutive probes (each with

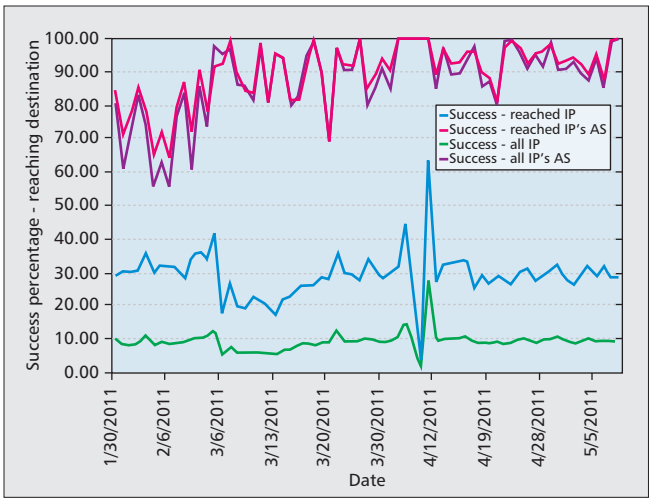

Figure 3. *Success rate in reaching destinations in Syria, by IP and AS.*

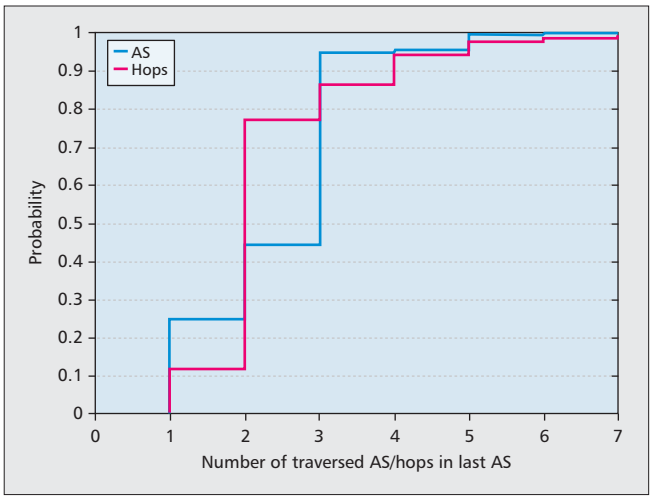

Figure 4. *CDF of number of ASes in a Traceroutes, and the number of IP hops in the last AS.*

increasing time to live [TTL] values). When a destination is reached, in 96.5 percent of the measurements all four consecutive traceroutes reach the destination, which shows that the routing success is not accidental.

Default Routing

Egypt — The Internet cutoff in Egypt presented a unique opportunity to examine default routing usage, since the means to cut off the network was by BGP routes withdrawal. In general, when an AP is not announce by BGP for sufficient time, its routing entry is aged out of the routing table. As a result, packets to a destination in this AP should be dropped, and an Internet Control Message Protocol (ICMP) packet with "destination unreachable" code should be sent to the origin. However, some ASs are using one upstream AS as their default routing, and only a small routing table for the few APs that are routed differently. For example, an AS in a small country can route all its traffic to the world via one designated "international" provider, except for traffic destined locally that require a significantly smaller routing table. Detecting default and static routes in an operating network is a hard task, which requires detection by techniques such as AS path poisoning [9].

The Internet blackout in Egypt started on January 28 [2], and ended on February 2. During this period, our measurements to Egypt included traceroutes through 163 different ASs, and they targeted IP addresses within 36 ASs. Out of the 163 ASs, traceroutes were terminated in 90 ASs, meaning traceroutes routed through the other 73 ASs were passed to the next AS and were not dropped, as one might expect. Among the ASs with default routing we found 23 universities and technological institutes (typically being the source of the measurement), which are expected to be heavy users of default routing. However, surprisingly, among the ASs with static routing we also found some tier 1 providers such as Sprint (AS1239), TeliaNet(AS1299), and Global Crossing (AS3549); and the EU research network GEANT (AS20965). The average number of ASs passed in a traceroute is 2.6, with some traceroutes being dropped at the originating AS while other traces traversed up to 7 different ASs before terminating. Figure 4 shows a CDF of the number of ASs included in a traceroute (solid line) when measuring to an unreachable destination. Only 24.6 percent of the traceroutes are dropped in the first AS and 19.7 percent in the second AS, but 94.7 percent of the traces are terminated after traversing at most three ASs. For comparison, when a destination is reached, only 13.6 percent of the traceroutes end by the third AS.

Concentrating on the AS that drops the traceroute, the minimal number of hops within the last AS is one, with an average of 2.39 hops. Figure 4 shows a CDF of the number of hops within the last AS (dotted line). Only 11.52 percent of the traceroutes are dropped in the first hop; this is unexpected when the dropping AS is not the first AS in the traceroute, meaning that the hop considered is typically in a point of presence (PoP) or IXP and is expected to be aware of BGP messages. Seventy-five percent of the traceroutes are dropped within the next two hops in the last AS, which in many cases is within the same PoP, as the nodes' DNS names indicate. The maximal number of hops until a traceroute was dropped was 42, and these rare cases were caused by loops in the last couple of hops.

The dominant ASs that terminated traceroutes were MTS (formerly Comstar, AS8359), Tata Communications (AS6453), Cogent (AS174), Australia's TPG Internet (AS7545) and Sparkle (the international IP backbone of Telecom Italia, AS6762). On the other end of the scale are ASs such as Level 3 (AS3356), where only a small number of the traceroutes were terminated. Out of the above ASs, Tata Communications, Cogent, and Level 3 are all major ASs where a message goes through 5 to 15 hops before it is dropped. As in most of these cases the AS is not the source of the measurement, the message is expected to be dropped sooner since the first hop is already part of major routing junction.

Aggregating ASs to countries nicely visualizes the static routing. Figure 5 shows the number of BGP static routes originating in each country, and the number of BGP static routes terminating in them. Each AS level path (origin AS, terminating AS) is only counted once regardless of the number on measurements, or the number of AS level routes in the path. The table shows only countries with two or more routes, and omits ten more counties with a single static route for clarity. In most countries there are a few originating ASs with a static route and a few terminating ones. These cases are often reflected in traceroutes that begin and end in the same country. The United States has a large number of routes, compared to other countries, in most cases between small providers and large ones, including tier 1 providers. Such routes are most likely default routes and not static. The large number of static routes terminated in Italy, and to a lesser extent in France, may be correlated to launching points of the submarine cables connecting to Egypt: SMW3, SMW4, and IMEWE. The terminating ASs in these cases, such as Sparkle (AS6762) and

Tata Communications (AS6453), are part of the submarine cables consortiums and the locations fit the cables launching points as well.

The loops detected during the cutoff period differ between providers. For example, in Sparkle (AS6762) probes passed 42 hops before being ; of these, 40 hops are between two nodes. This loop was detected only once, on January 31. On other dates, measurements to the same destination from the same source or on a similar path were routed over a different path starting from the second hop in this AS. A different type of loop was detected in the North Carolina Research & Education Network (NCERN — MCNC, AS81). This loop, 31 or 34 hops long, cycled the probes between four nodes, all of them configured as gateways by their DNS names. This loop appeared in traces to 12 distinct IP addresses in six different ASs in Egypt and did not reoccur after the Internet cutoff ended. An interesting loop was found between Bibliotheca Alexandrina (AS33782) and Reliance Globalcom (formerly Flag Telecom, AS15412). We find many traceroutes to two addresses in Bibliotheca Alexandrina that are looped back and forth between the two ASs, for about 30 hops (depending on entry point) until being dropped in AS15412. The measured IP address never replies, and the routing anomaly also exists after the Internet blackout event, indicating a problem in the routing policy of one of these ASs.

Libya — Libya's technique to cut off the Internet was different than in Egypt. While on February 19th and 20th there was a short attempt to cut off the Internet by withdrawing BGP messages too, these two attempts were short (approximately six hours). The main drops in Internet traffic and the reduction in the average traffic were claimed [6] to be caused by the international gating service provider, GPTC (AS21003). This claim could not be verified by us: while BGP updates are still announced (based on Routeviews [5], rib.20110303.0000 and rib.20110325.0000), the traceroutes do not reach the target AS. Most of the failing traceroutes (over 98 percent) terminate at Sparkle (AS6762), which is by BGP announcements the last AS before the Libyan GPTC AS.

Syria — As reported earlier, we did not detect in Syria traffic cutoff on specific dates. Internet cutoffs reported later in June 2011, were not covered by our experiment.

Routing in Syria

Syria presents a very interesting case, as all the traffic goes through a single AS, and there is a low reach rate of the destinations. Several unexpected phenomena characterize the aborted traceroutes. First, we characterize the IPs that are reached versus the IPs that fail. The IP addresses selected for the experiment are selected at random across a /24 CIDR range, so a postfix of .1 is as likely to be selected as a postfix of .135. Also, since Syria has a single AS, the address distribution between ISPs was random and uneven. As the experiment's intent was to detect default routes at the AS level, this was decided during the experiment's design stage.

Even though only 18 percent of the probed IP addresses are to SCS, the Syrian Computer Society, we manage to reach over 75 percent of its IP addresses, with very high success rates to some of the addresses (97–98 percent) and low success rate to others (0.05–14 percent). On the other hand, when measuring to ZAD, which included 65 percent of the measured IP addresses, only 14 percent of the IP addresses are reached and with low success rates — less than 0.2 percent. In the group of reached IP addresses

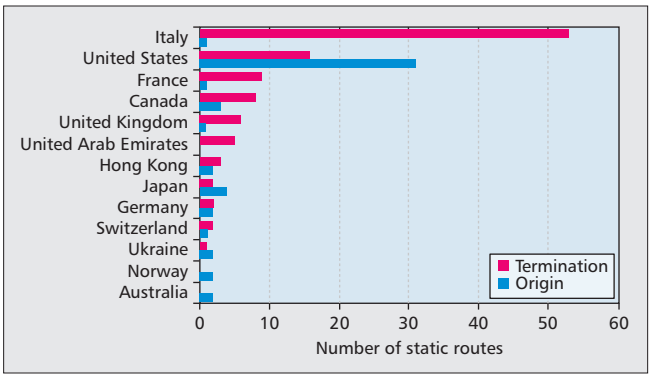

Figure 5. *Static Routing: the number of originating and terminating ASes in key countries.*

there is no preference for a range of addresses, say CIDR /25, over another.

Out of all the measurements to the ZAD network, only six measurements reached the destination's ISP IP range. On the other hand, 87.3 percent of the IP addresses were terminated at STE, the Syrian Telecommunications Establishment. It should be noted that ZAD is an operating ISP, but we cannot guarantee that it uses the address range assigned to it. We did track on the web at least one traceroute of a ZAD customer that connected from its home network to a network address assigned to STE and not to ZAD.

As previously mentioned, the success in reaching destinations in Syria is not accidental: a probed addressed is most of the time reached by all four traceroute probes, and every IP address is probed tens of times every day. The success rate in reaching the destination is also independent of the source AS: the success rate from all AS ranges from 8 to 11 percent. We do see a couple of cases or specific IP addresses where the success rate is lower; however, there is insufficient number of measurements or vantage points within the AS to correlate this to any kind of an AS blocking. The destination ISP, meaning destination IP range, is also considered. We find that for both SCS and Runnet, the reaching rate success is similar across all AS. This means that the source AS combined with the destination ISP does not influence the blocking. SCS and Runnet are representing examples, as one is always accessible (at a certain success rate), while for other service providers the probing succeeded only on occasional dates.

Private IP Addresses — A valid traceroute is expected to go through a series of hops, each either responding with a routable IP address, or not responding at all (anonymous hop). While in any traceroute study anomalies such as private IP addresses (e.g., 192.168.0.0/16 or 10.0.0.0/8) can be detected [10], we have found in this study a large percentage of traceroutes with this anomaly for Syrian targets.

Out of all the traceroutes to Syria, 5.6 percent of the measurements ended in a private IP (3.78 percent) or went through one (1.8 percent). When measuring to Egypt, 0.5 percent of the traceroutes terminate at an unknown IP address (close to zero pass through a private IP), and when measuring to a comparison group of 18 Arab countries, 0.21 percent of all traceroutes terminate or pass though a private IP. Interestingly, less than 0.09 percent of the measurements to Libya exhibit such a behavior. It is important to note that measurements passing through private IPs at the *beginning* of the traceroute are omitted from these numbers, since it is a common practice in home networks, and many public providers.

As Syria is conspicuous in the appearance of private IP address in its probing, we focus on it. Thirty-two private IP addresses are identified along the paths: 10 addresses as the

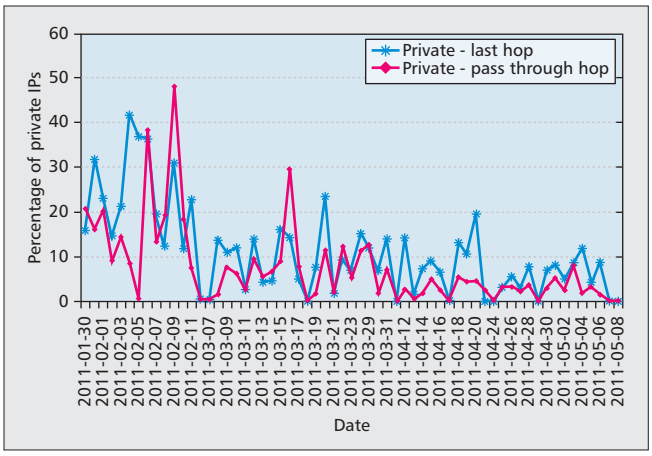

Figure 6. *Percentage of private IP addresses by date — Syria.*

final hop, 12 addresses as pass-through hops, and the rest appear in both cases. Most addresses that appear in both cases share the prefix 10.1.100.0/24 or 10.2.20.0/24, while addresses with the prefix 192.168.0.0/16 appear as either a last or pass-through hop, but not both. The prefix 10.50.19.0/24 reoccurs multiple times as a pass-through hop. Figure 6 shows that the appearance of private IP addresses in traceroutes is not stable, albeit the fact that the same target IP addresses reoccur every day: On some days, there are close to zero private addresses, while on others, private IP addresses appear in over 30 percent of the measurements, as either a last hop or pass-through. There is, however, one noticeable trend: during the first two weeks of February 2011, the average number of private IP addresses being the last hop was 24 percent, and 18 percent of the traceroutes encountered a private IP along the path, while during March 2011, these figures dropped to 9 and 7 percent, respectively, and reached 6 percent as the last hop and 2.8 percent en route. The reduction in traces with private IP addresses does not stem from a major change in the routing: all the terminating private IP addresses that appear in February also appear in March through May, and out of the en route hops' private IP addresses, only one is detected during February alone, and the number of measurements through it is small (12).

Many of the traceroutes to Syria contain a hop with a non-responding IP address, close to 24 percent. Out of those, a quarter of the traceroutes are followed by a private IP address. In the traceroutes that share both a non-responding hop and a private IP address, the private IP immediately follows the non-responding hop. Only 6.4 percent of the traceroutes that contain a private IP do not include a non-responding address along the path.

In 65 percent of the traceroutes that contain a private IP address, this IP address is the final hop. 15.8 percent of the traceroutes contain one more hop after the first encounter with a private IP, and 12.3 percent have two more hops following the first encounter with a private IP. Close to 53 percent of the traceroutes that terminate one hop after a private IP do not reach the destination IP, but rather reach an IP that belongs to SpeakEasy (AS23504), which is not a Syrian provider, but provides among others VoIP and software as a service for Internet cafes. Thirty-three percent stop at an IP address that belongs to STE address range (91.144.0.0/18). Among the traces that end two hops after a private IP address, 47 percent belong to AYA ISP and an equivalent number to STE. We note that some of the hops that are reached after a private IP addresses do so through an alternate route; thus, it is hard to make a judgement on the validity of the routing.

Discussion

The results described in the previous section shed light on both governments' attempts to control the Internet and otherwise hidden routing rules of the network. The importance of those is crucial when discussing the free and democratic usage of the network. We observe the diverse ways Arab governments use to control their citizen's web access (e.g., [8]). Even in a relatively open speech country such as Egypt, where tens of millions of people use the Internet via many ISPs, and where mobile Internet penetration rate is high, the government can disconnect the country from the Internet in a single act. The situation in less democratic societies, such as Syria and Libya, is worse, as access to the Internet is limited through a government controlled AS. In these countries, there is no need to stop BGP advertisement, as there are simpler means, less noticeable, to regulate traffic.

While static routes are expected to be scarce, we find that they are used in many ASs, small ones as well as tier 1 providers. It is hard to distinguish between default and static routes. A default route will be used only in a customer-provider relationship, complying with valley-free routing rules; thus, a default route may be used by a small ISP to route messages to their provider. When main ASs that are part of the Internet core are concerned, however, the route is for sure static and not default, going from the core of the net to its outskirts. We acknowledge that some of these routes may be backup routes, thus taking effect only when a dynamic route learned from BGP messaging is not available.

Static routing increases the load in the Internet: over half of the packets to an unreachable destination in our study traversed through three ASs or more instead of being dropped in the first couple of ASs, which translates to five to six additional IP hops. Furthermore, even in the last AS there are few hops before dropping the packet. Thus, minimizing static routing can decrease traffic and improve network performance.

For future civil unrest scenarios, BGP is shown to be a fairly strong tool to shut down a country. However, our study results show that adding static routing at a few points, together with the static and default routing that already exists, can be quite effective to restore much of the country's connectivity (at least in the case of countries like Egypt where a few submarine cables serve as the country entrance points).

Conclusion

To conclude, this work presents the results of a large-scale measurement effort of Arab countries during the turmoil of 2011. It showed how traffic was regulated during this period by the governments, thus blocking Internet access to these countries to varying extents. We also reported the existence of static and default routes in large and small autonomous systems and their mapping to countries. These findings can be used to understand possible risks to the free usage of the Internet as well as to improve day-to-day network management.

Acknowledgments

We would like to thank Emile Eben for suggesting the study of static routing with our data.

References

[1] K. M. Pollack, *The Arab Awakening: America and the Transformation of the Middle East*, Brookings Institution Press, 2011.
[2] Council on Foreign Relations, *The New Arab Revolt: What Happened,*

What It Means, and What Comes Next, 2011.
[3] R. B. Wright, *Rock the Casbah: Rage and Rebellion Across the Islamic World*, Simon & Schuster, 2011.
[4] Y. Shavitt and E. Shir, "DIMES: Let the Internet Measure Itself," *ACM SIGCOMM Computer Commun. Review*, vol. 35, Oct. 2005.
[5] University of Oregon Advanced Network Technology Center, Route Views Project, http://www.routeviews.org/.
[6] J. Cowie, What Libya Learned from Egypt, Renesys, 2011, http://www.renesys.com/blog/2011/03/what-libya-learned-from-egypt.shtml.
[7] G. Blight, S. Pulham, and P. Torpey, *Arab Spring: An Interactive Timeline of Middle East Protests*, The Guardian, 2011.
[8] R. Deibert, *Access Denied: The Practice and Policy of Global Internet Filtering*, Information Revolution & Global Politics, Mit Press, 2008.
[9] R. Bush *et al.*, "Internet Optometry: Assessing the Broken Glasses in Internet Reachability," *IMC '09: Proc. 9th ACM SIGCOMM Conf. Internet Measurement Conf.*, 2009, pp. 242–53.
[10] B. Donnet *et al.*, "Efficient Algorithms for large-Scale Topology Discovery," *ACM SIGMETRICS*, June 2005, pp. 327–38.

Biographies

YUVAL SHAVITT (shavitt@eng.tau.ac.il) received a B.Sc. in computer engineering (cum laude), an M.Sc. in electrical engineering, and a D.Sc. from the Technion – Israel Institute of Technology, Haifa, in 1986, 1992, and 1996, respectively. After graduation he spent a year as a postdoctoral fellow in the Department of Computer Science at Johns Hopkins University, Baltimore, Maryland. Between 1997 and 2001 he was a member of technical staff at Bell Labs, Lucent Technologies, Holmdel, New Jersey. Starting October 2000, he is a faculty member in the School of Electrical Engineering at Tel-Aviv University, Israel. His research interests include Internet measurements, mapping, and characterization; and data mining peer-to-peer networks.

NOA ZILBERMAN (noa@eng.tau.ac.il) received her B.Sc. and M.Sc. (both magna cum laude) in electrical engineering from Tel-Aviv University, Israel in 2003 and 2007, respectively. Since 1999 she has filled several development, architecture and managerial roles in the telecommunications industry. She is currently a Ph.D. candidate in the School of Electrical Engineering at Tel-Aviv University. Her research focuses on Internet measurements, mapping, and characterization.