Arabian Nights - measuring the Arab Internet during the 2011 events

Yuval Shavitt
School of Electrical Engineering
Tel-Aviv University, Israel
shavitt@eng.tau.ac.il

Noa Zilberman
School of Electrical Engineering
Tel-Aviv University, Israel
noa@eng.tau.ac.il

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 by "shutting down" the Internet in their country. In this paper we describe active measurements conducted during 2011 to 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.

1. INTRODUCTION

On December 2010, a wave of unrest has shaken the Arab world, starting in Tunisia and spreading to other Arab countries such as Egypt, Libya, Yamen 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 the 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 in 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 had 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 a fairly open access to the Internet before and after the peak of the unrest there, Syria is tightly monitoring its citizens' access to the Internet for years, and in Libya the status is somewhere in 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 has 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 BGP withdraw messages for large portions of the Egyptian address space. Namely, the BGP protocol was told that there is 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 packet 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 is static. We also revealed many coherency problems in the routing tables of large providers. We report these findings as well in this paper.

2. DATA SET

The collected dataset for this work is taken from DIMES [4]. We use 3.63 million traceroute measurements from 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% 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: The first, the IP address is located in the target country, based on Maxmind* or IPligence† Geolocation databases. Or second, the IP belongs to an AS that is registered in the target country. Each experiment was run daily by one hundred agents, each was assigned to measure one thousand target IP addresses out of the pool of IPs described above. The measuring agents were altered every day, with the purpose to measure to the same destinations through different paths, and with most agents taking part in the experiments

*Maxmind GeoIP, http://www.maxmind.com
†IPligence Max, http://www.ipligence.com

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.

3. RESULTS

3.1 Reaching Destination

A traceroute measurement towards 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 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 AS. In the following we use both the narrow (namely, reaching the target IP) and wide sense measurement success definitions.

Egypt.

The targeted measurements to Egypt show the Internet cutoff between January 28th, a day after the government also closed the banks and cellular networks, and February 2nd, 2011, the bloodiest day in Tahrir Square to date, after which the ban was taken off [2]. During these dates Figure 1 shows the success rate for of the targeted measurements. While on average 25% of the measurements reach their final destination IP and 44% of them reach the destination AS, during the shutdown period only 4% of the measurements reached their destination IP and 8% 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% of the traceroutes reaching their destination IP and 73% reaching the destination's AS. During the shutdown period only 11% of these IP addresses were reached, and also only 11% of the destination AS. The ASes that are reached match ASes 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. Amongst these ASes are Etisalat Egypt (AS36992), a class A service provider and NOOR (AS20928), a class B service provider [‡]. We note that on February 12th through 15th 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 (see Figure 2): The average success rate in reaching destinations before March-5th is 54% in the narrow sense and 93% in the wide sense, after this date the success percentage drops to 33% and 70%, respectively.

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

On March 19th, coalition forces began the 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 25th, 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 twelve service providers in Syria that are grouped under 2 ASes, 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 fifty. The attempt to probe specific addresses in Syria proved more difficult than in other countries: only about 10% of the IP addresses were reached, and even when considering only the responding IP addresses the traceroute success rate was only 30%. On the other hand, the success rate in reaching the destination AS is very high: about 88% on the average. We also do not see Internet cut offs: the behavior is quite constant throughout the measurement period (except for a slight increase in AS reach level), and even during February 5th 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 hints that the nature of Syrian Internet blocking is likely have been 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 ISP over the measurements' period. We observe a large variability

[‡] *http : //ntra.gov.eg/presentations/ LicensedTelecomChart*22122009*_En.pdf*

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 increasing TTL values). When a destination is reached, in 96.5% of the measurements all four consecutive traceroutes reach the destination, which comes to show that the routing success is not accidental.

3.2 Default Routing

Egypt.

The Internet cutoff in Egypt presented a unique opportunity to examine default routing usage, since the mean to cutoff 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 ICMP packet with "destination unreachable" code should be sent to the origin. However, some ASes 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 28th [2], and ended on February, 2nd. During this period, our measurements to Egypt included traceroutes through 163 different ASes, and they targeted IP addresses within 36 ASes. Out of the 163 ASes, traceroutes were terminated in 90 ASes, meaning traceroutes routed through the other 73 ASes were passed to the next AS and were not dropped, as one might expect. Amongst the ASes with default routing we found twenty three universities and technological institutes (typically being the source of the measurement) which are expected to be heavy users of default routing. However, surprisingly amongst the ASes 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 ASes passed in a traceroute is 2.6, with some traceroutes being dropped at the originating AS while other traces traversed up to 7 different ASes before terminating. Figure 4 shows a CDF of the number of ASes included in a traceroute (solid line) when measuring to an unreachable destination. Only 24.6% of the traceroutes are dropped in the first AS and 19.7% in the second AS, but 94.7% of the traces are terminated after traversing at most three ASes. For comparison, when a destination is reached only 13.6% 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% 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 PoP or IXP and is expected to be aware of BGP messages. 75% of the traceroutes are dropped within the next two hops in the last AS, which in many cases is within the same point of presence (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 ASes 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 ASes such as Level 3 (AS3356), where only a small number of the traceroutes were terminated. Out of the above ASes, Tata Communications, Cogent and Level 3 are all major ASes 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 ASes 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 ASes 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 ones. 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 ASes 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 dropped, out of them forty hops are between two nodes. This loop was detected only once at January 31st. On other dates, measurements to the same destination from the same source or in similar path were routed in a different path starting from the second hop in this AS. A different type of loop was detected in 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 twelve distinct IP addresses in six different ASes 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 traceroute to two addresses in Bibliotheca Alexandrina that are looped back and forth between the two ASes, for about twenty hops (depending on entry point) until being dropped in AS15412. The measured IP address never replies, and the routing anomaly exists also after the Internet blackout event, indicating a problem in the routing policy of one of these ASes.

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%) terminate at Sparkle (AS6762), which is by BGP announcements the last AS before the Libyan GPTC AS.

Syria.

As reported in Section 3.1, we did not detect in Syria traffic cut off on specific dates. Internet cutoffs reported later in June 2011, were not covered by our experiment.

3.3 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 in 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 on AS level, this deemed during the experiment's design stage.

Even though only 18% of the probed IP addresses are to SCS, the Syrian Computer Society, we manage to reach over 75% of its IP addresses, with very high success rates to some of the addresses (97%-98%) and low success rate to others (0.05% to 14%). On the other hand, when measuring to ZAD, which included 65% of the measured IP addresses, only 14% of the IP addresses are reached and with low success rates - less than 0.2%. In the grouped of reached IP addresses there is no preference of a range of addresses, say CIDR /25, over another. Out of all the measurements to ZAD network, only six measurements reached the destination's ISP IP range. On the other hand, 87.3% of the IP addresses were terminated at STE, the Syrian Telecommunications Establishment. It should be noted that ZAD is an operating ISP, but we can not guarantee that it uses the address range that was 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 was previously mentioned in section 3.1, 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%. 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.

3.3.1 Private IP addresses

A valid traceroute is expected to go through a series of hops, each responding either with a routable IP address, or not responding at all (anonymous hop). While in any traceroute study anomalies such as private IP addresses (such as 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% of the measurements ended in a private IP (3.78%) or went through one (1.8%). When measuring to Egypt, 0.5% 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 eighteen Arab countries, 0.21% of all traceroute terminate or pass though a private IP. Interestingly, less than 0.09% 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: ten addresses as the final hop, twelve addresses as a 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 either as a last or a 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% of the measurements, either as a last hop or pass through. There is, however, one noticeable trend: during the first 2 weeks of February 2011 the average number of private IP addresses being the last hop was 24% and 18% of the traceroutes encountered a private IP along the path, while during March 2011, these figures dropped to 9% and 7%, respectively, and reached 6% as the last hop and 2.8% 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%. 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% of the traceroutes that contain a Private IP do not include a non-responding address along the path.

In 65% of the traceroutes that contain a private IP address, this IP address is the final hop. 15.8% of the traceroutes contain one more hop after the first encounter with a private IP, and 12.3% have 2 more hops following the first encounter with a private IP. Close to 53% of the traceroutes that terminate one hop after a private IP do not reach the destination IP, rather it reaches an IP that belongs to SpeakEasy (AS23504), which is not a Syrian provider, but provides amongst others VoIP and Software as a Service for internet cafes. 33% stop at an IP address that belongs to STE address range (91.144.0.0/18). Amongst the traces that end two hops after a private IP address 47% 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 judgement on the validity of the routing.

4. DISCUSSION

The results described in the previous section shed light both on governments attempts to control the Internet and on 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 (see 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 Syria and Libya, is worse, as the 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 ASes, small ones as well as tier-1 providers. It is hard to distinct 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 small ISP to route messages to their provider. When main ASes 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 message 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 ASes or more instead of being dropped in the first couple of ASes, which translate 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 the 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 in a

few points, together with the static and default routing that already exists, can be quite effective to restore much of the country connectivity (at least in the case of countries like Egypt where a few submarine cables serve as the country entrance points).

5. CONCLUSION

To conclude, this work presented the results of a large scale measurements effort to 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 extent. 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.

6. ACKNOWLEDGEMENTS

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

7. 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] Robin B. Wright. *Rock the Casbah: Rage and Rebellion Across the Islamic World*. Simon & Schuster, 2011.

[4] Yuval Shavitt and Eran Shir. DIMES: Let the internet measure itself. In *ACM SIGCOMM Computer Communication Review*, volume 35, October 2005.

[5] University of Oregon Advanced Network Technology Center. Route views project. http://www.routeviews.org/.

[6] James Cowie. What libya learned from egypt. Renesys, 2011. http://www.renesys.com/blog/2011/03/what-libya-learned-from-egypt.shtml.

[7] Garry Blight, Sheila Pulham, and Paul 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] Randy Bush, Olaf Maennel, Matthew Roughan, and Steve Uhlig. Internet optometry: assessing the broken glasses in internet reachability. In *IMC '09: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 242–253, 2009.

[10] Benoit Donnet, Philippe Raoult, Timur Friedman, and Mark Crovella. Efficient algorithms for large-scale topology discovery. In *ACM SIGMETRICS*, pages 327–338, June 2005.

Yuval Shavitt received the B. Sc. in Computer Engineering (cum laude), M. Sc. in Electrical Engineering and D. Sc. from the Technion — Israel Institute of Technology, Haifa, Israel in 1986, 1992, and 1996, respectively. After graduation he spent a year as a Postdoctoral Fellow at the Department of Computer Science at Johns Hopkins University, Baltimore, MD. Between 1997 and 2001 he was a Member of Technical Stuff at Bell Labs, Lucent Technologies, Holmdel, NJ. 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 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.

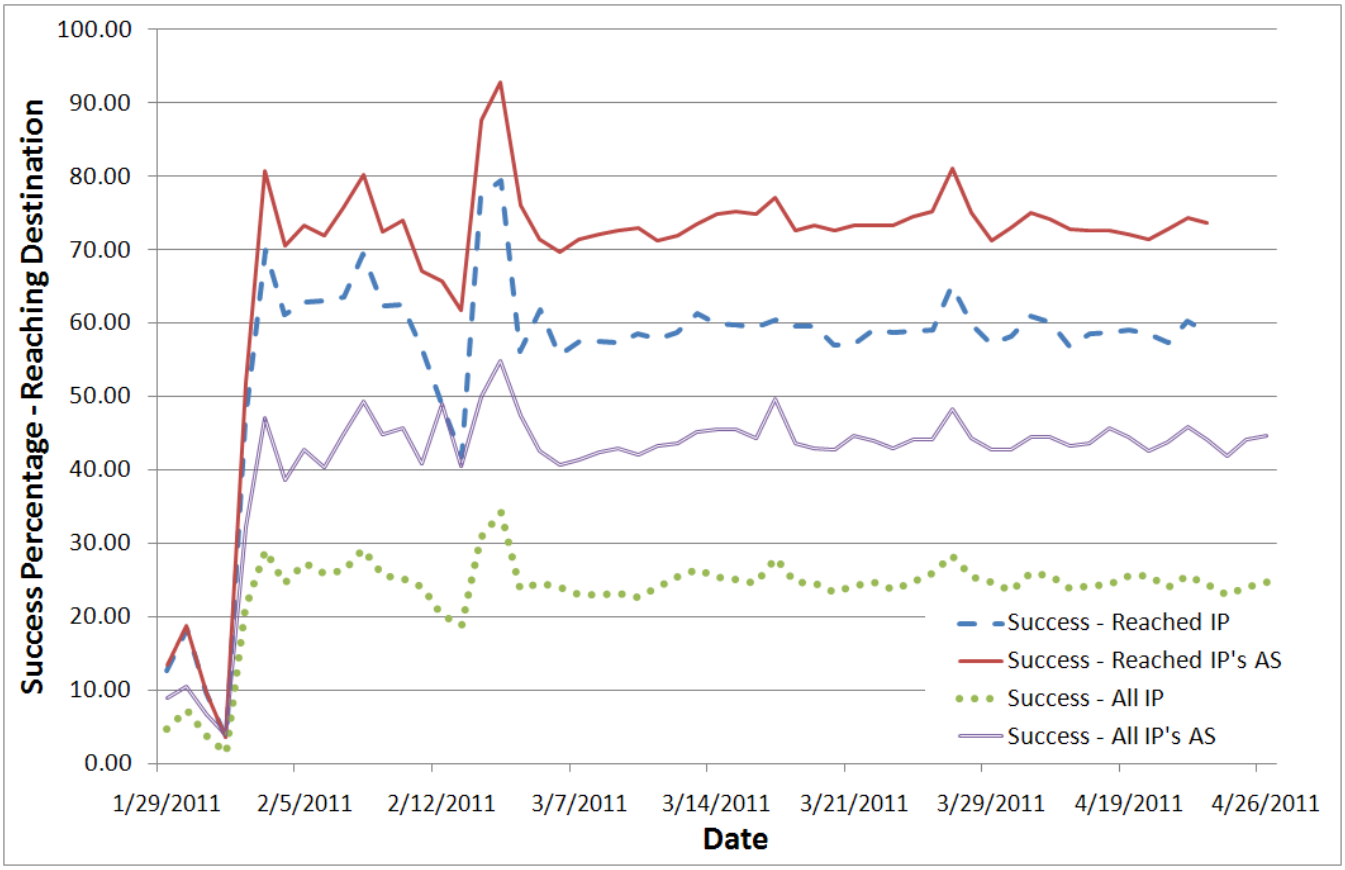

Figure 1: Success Rate in Reaching Destinations in Egypt, by IP and AS

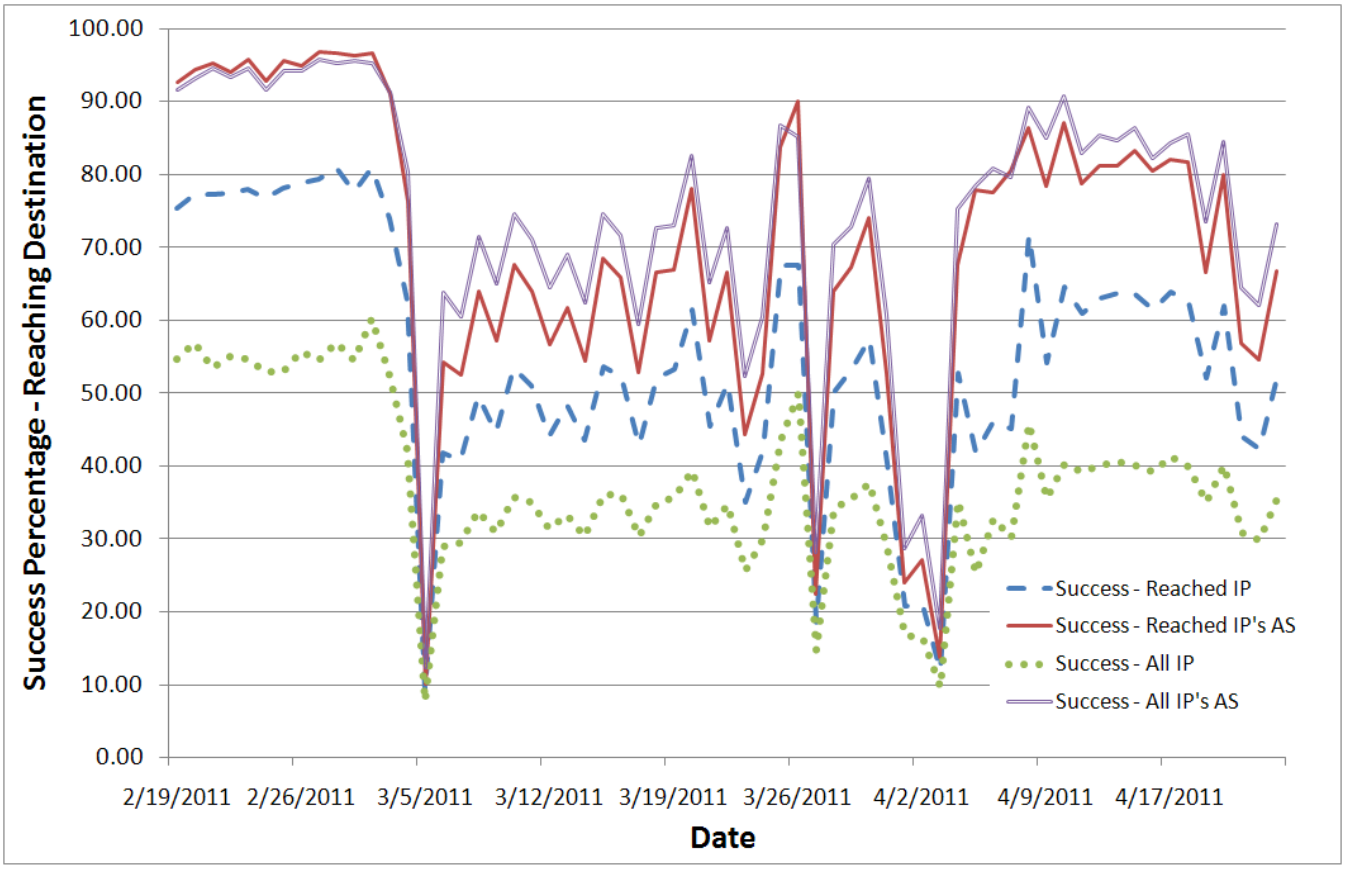

Figure 2: Success Rate in Reaching Destinations in Libya, by IP and AS

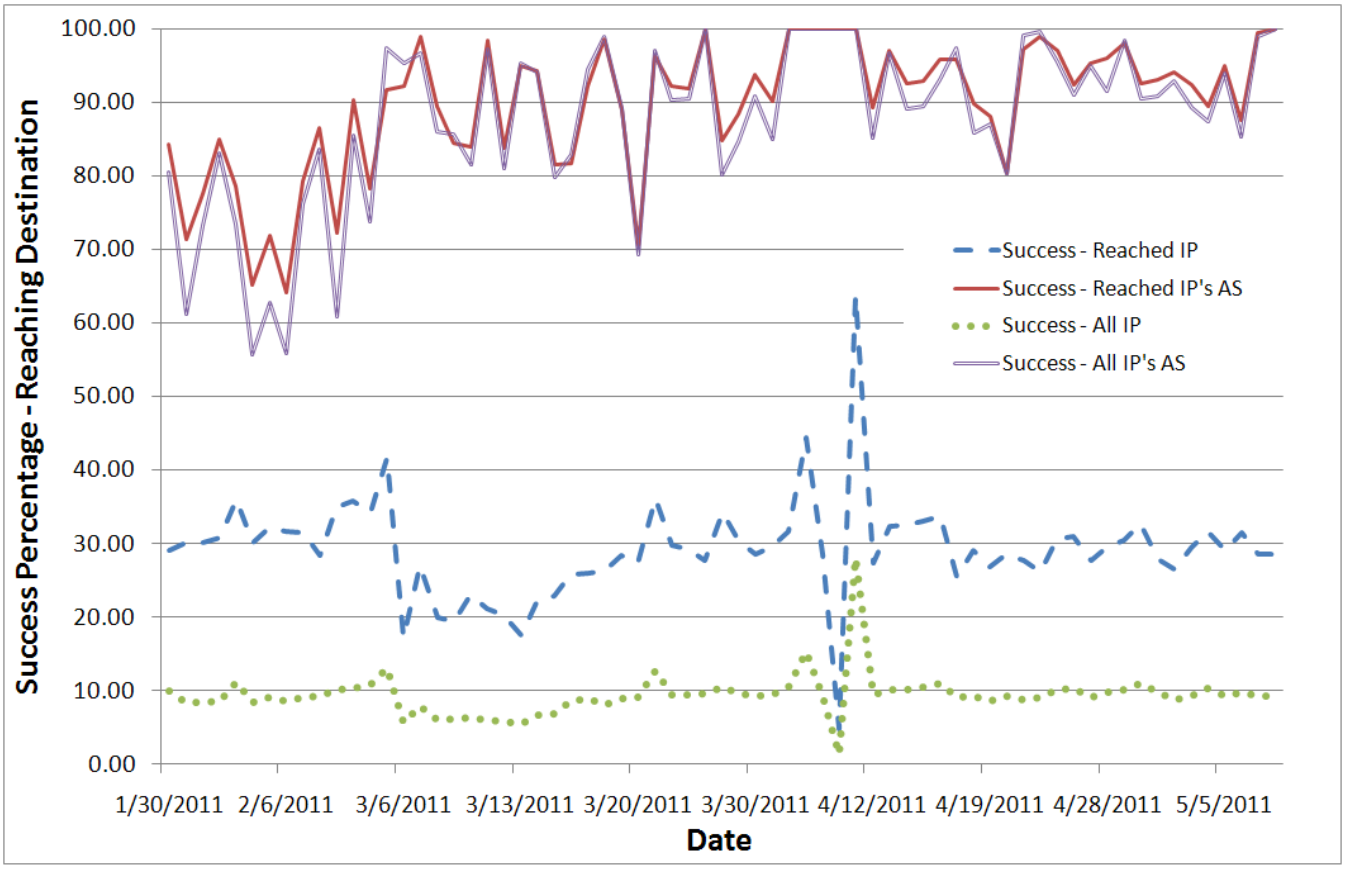

Figure 3: Success Rate in Reaching Destinations in Syria, by IP and AS

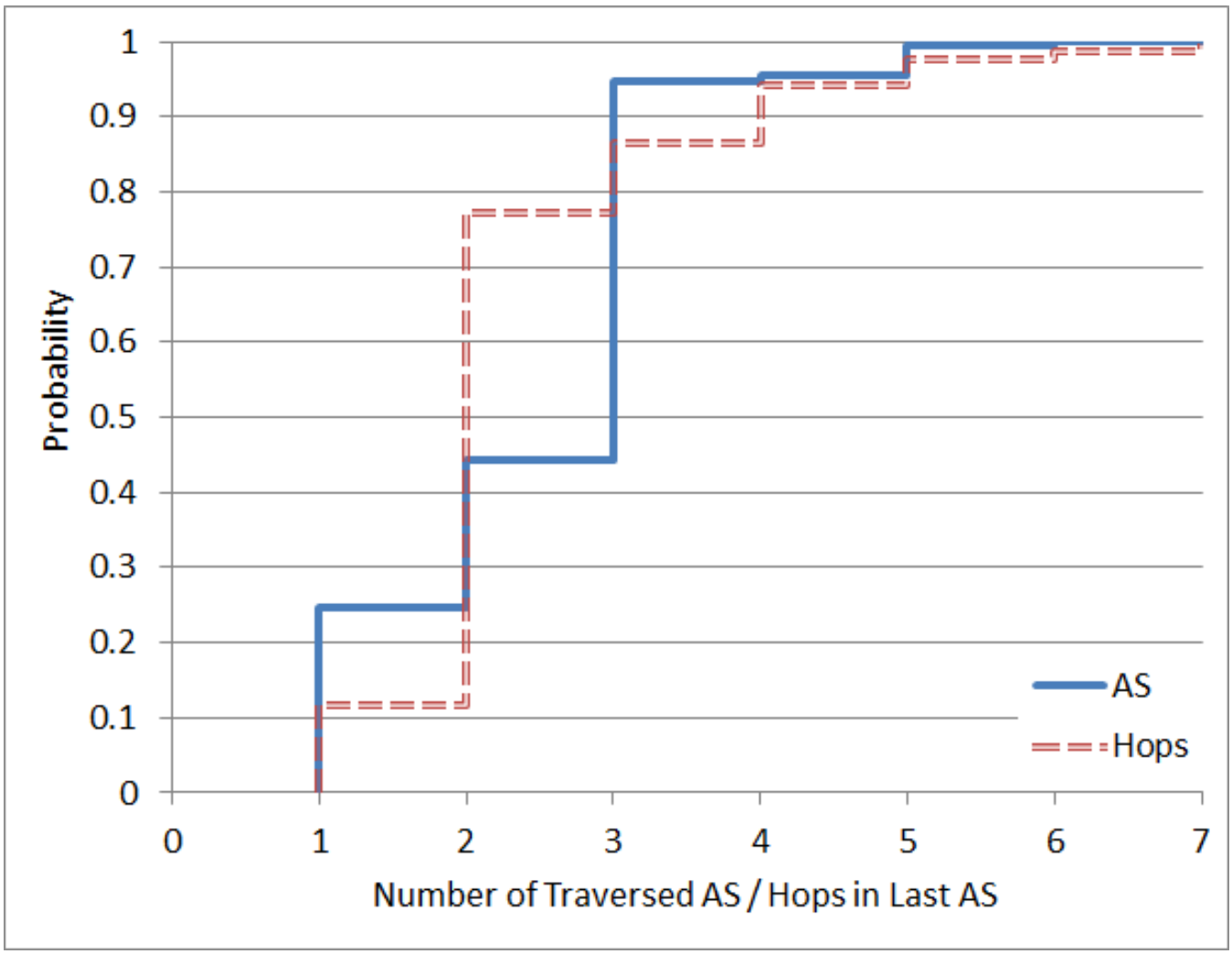

Figure 4: CDF of number of ASes in a Traceroutes, and the number of IP hops in the last AS.

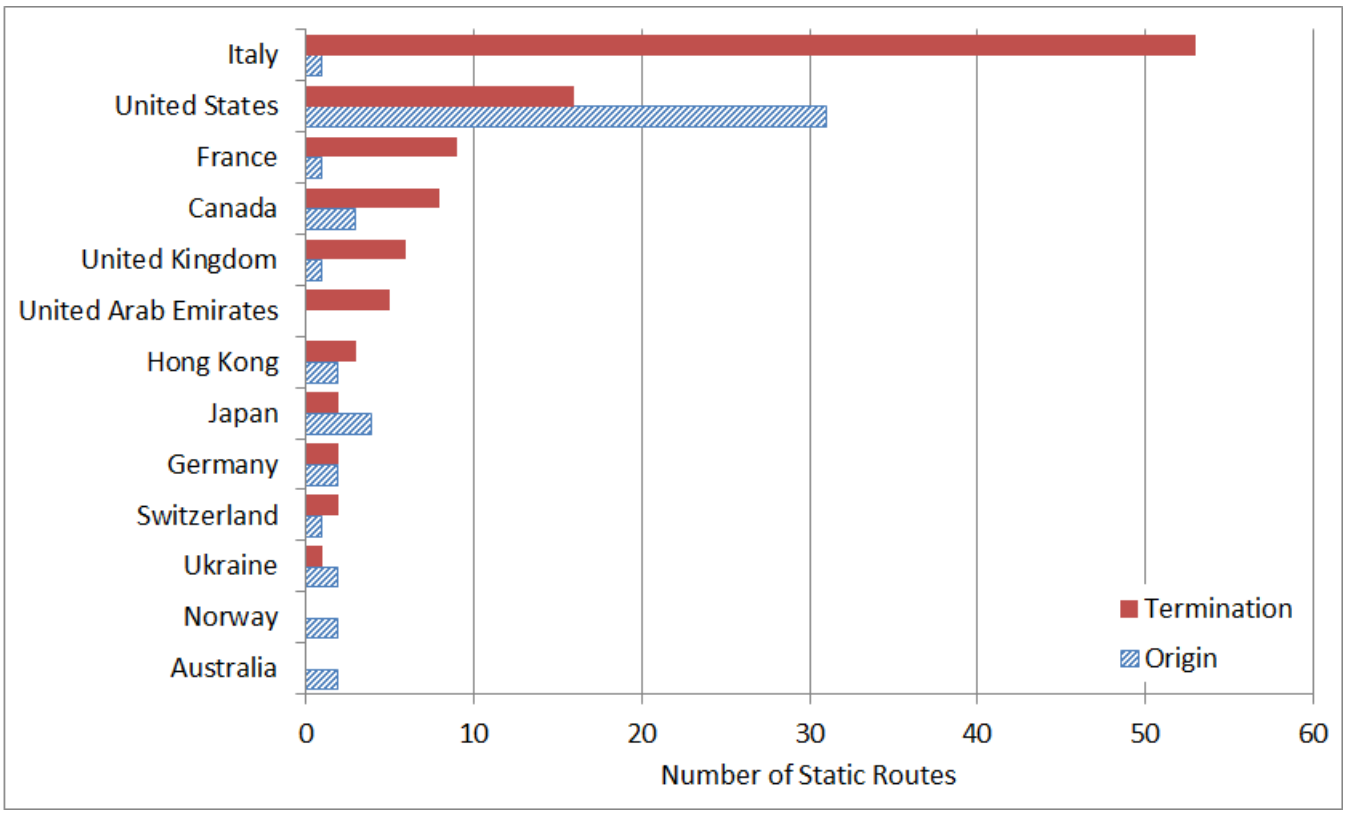

Figure 5: Static Routing: The Number of Originating and Terminating ASes in Key Countries

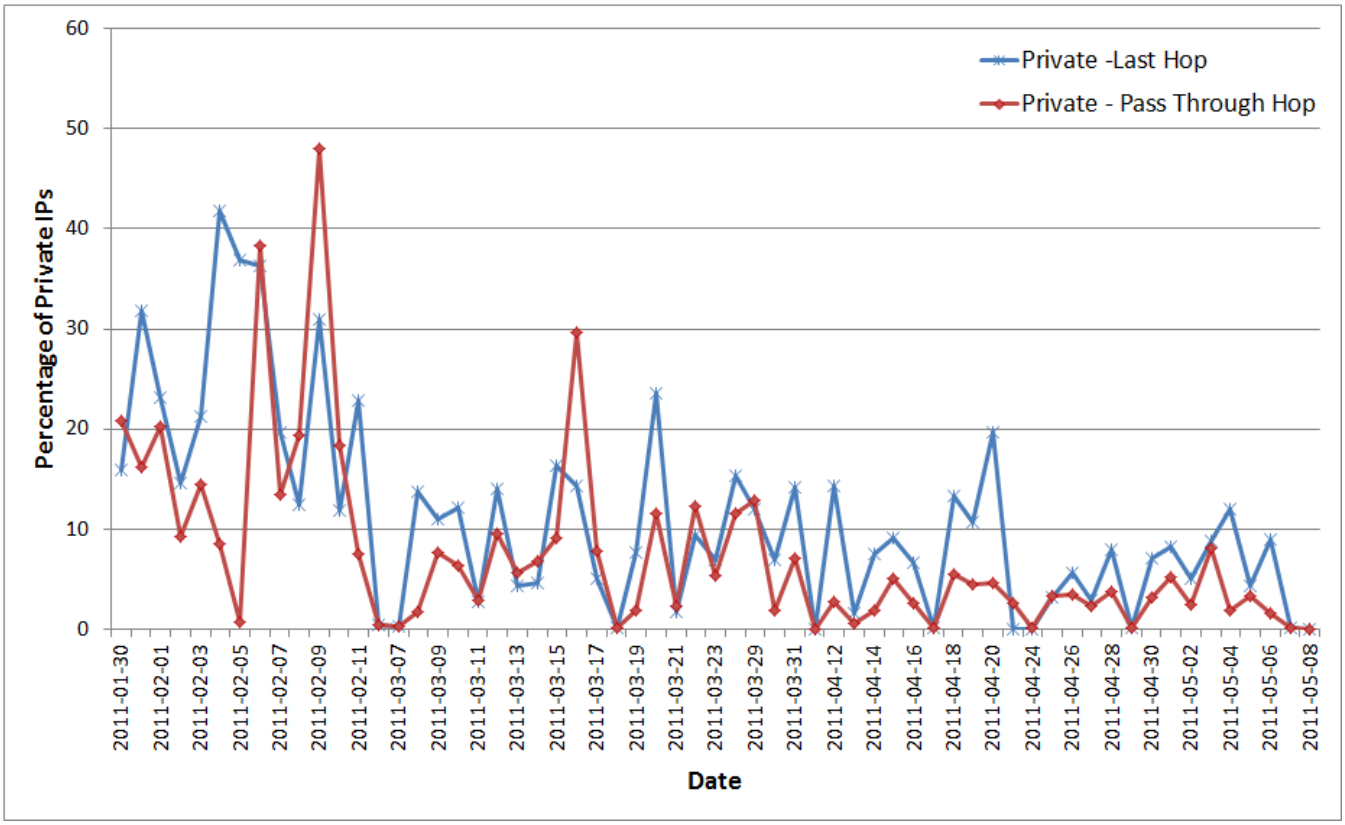

Figure 6: Percentage of Private IP Addresses by Date - Syria