Society | DOI:10.1145/2330667.2330674

Marina Krakovsky

Garbage In, Info Out

Security researchers used malware to investigate large-scale Internet censorship in Egypt and Libya.

A crowd scene in Cairo's Tahrir Square during Egypt's five days of Internet outage last year; the large sign on the KFC window says "Awiz Internet" ("We want Internet").

PHOTOGRAPH BY NEBEDAAY

EARLY LAST YEAR, when antigovernment protests broke out in one oppressive regime after another, one of the casualties was Internet access as governments scrambled to stem the flow of information among the people and with the outside world. Soon after the Arab Spring, an international team of computer scientists began analyzing precisely what happened in two of the affected nations, Egypt and Libya. Their fine-grained analysis of Internet censorship in these countries, which won this year's Applied Networking Research Prize from the Internet Research Task Force, emerged in part from a surprising source of data: malware.

"We've never before seen an entire country disappear from the Internet for several days," says Alberto Dainotti, lead researcher with the Cooperative Association for Internet Data Analysis (CAIDA) at the University of California, San Diego (UCSD). The Egyptian outage was a monumental event, cutting off Internet access for 23 million users in a nation of 80 million. But news reports about the outage were incomplete and sometimes contradictory. "So the idea was to add rigorous measurement using different techniques and combine them to better understand what happened," says Dainotti.

The most novel technique was the use of the darknet, a portion of routed Internet Protocol (IP) address space on which exists little or no legitimate traffic. CAIDA researchers and others have long examined the darknet (also called a network telescope) to study Internet Background Radiation (IBR), or unsolicited one-way traffic sent to random IP addresses around the world. If a worm is scanning the Internet trying to infect other machines, for example, it sends packets in all directions—including to the UCSD network telescope, which currently collects an astounding 3.5 terabytes of IBR per month, or 1/256 of all IBR. This study, though, appears to be the first to use this Internet rubbish to make inferences about the timing and nature of Internet outages.

Unwitting Victims

Computers that generate IBR, typically unwitting victims of malware themselves, are scattered around the globe and the UCSD network telescope sees IBR from about 17 million distinct IP addresses each day. Studying IBR activity, therefore, can yield reliable clues about overall Internet activity. "The idea was that this is a signal that comes from virtually every city and country in the world," explains Dainotti, "and we can see when the signal goes up and down."

The Egyptian outage was a monumental event, cutting off Internet access for 23 million users in a nation of 80 million.

The UCSD network telescope does not have to know where to look, since it passively receives packets from all over. But tracking outages by country does require knowing where the IBR traffic is coming from. Fortunately, each packet that reaches the network telescope has a source address of the infected machine that sent it. By looking up these IP addresses in a geolocation database, the researchers were able to find packets' countries of origin. "We can discriminate between packets that come from Egypt and packets that come from Italy, let's say," Dainotti explains. Using these tools, they were able to tally how many packets per second were coming from each country during any given period. For example, as shown in their paper, the researchers were able to see a clear and sharp drop in the packet rate from Egypt starting late January 27, 2011, and a return to normal traffic levels around midday Cairo time on February 2, when the government restored Internet access. Several days later, Egypt's president, Hosni Mubarak, was forced to step down amidst continued protests and international pressure.

Dainotti and his team used the same method to analyze traffic trends from Libya following anti-Gaddafi protests there. "I was surprised we could see the signal so vividly," says K.C. Claffy, CAIDA's founder and principal investigator for the research. The darknet traffic is a mess that includes traffic that uses fake source addresses, adding significant noise to the data; however, the signal from both countries came through loud and clear because of the sheer number of infected hosts sending out IBR. Looking at real Internet traffic, as opposed to malware, might have yielded even cleaner results in theory, but privacy concerns make it illegal for service providers to share such information.

So malware proved to be an asset, but pinpointing packets by country was only the beginning. "What's interesting and new about our analysis is it gives us the opportunity to study the chronology of the event that no other source of data allows you to do," Claffy says. For example, once they had figured out which packets were coming from Egypt, they were able to map these packets to their autonomous systems of origin, such as Egyptian Telecom, the proprietary networks operating the Egyptian Stock Exchange, and the famed Library of Alexandria. The researchers could see that at the beginning of the outage, while most networks had been shut down, the stock exchange was still up and running, presumably because the government did not want to disrupt trading. Ultimately, though, even the stock exchange went down. "Things got worse in Egypt, and people were using the remaining networks to communicate with the world, so I'm imagining that the government was panicking, and they started shutting down everything," says Dainotti.

> **"Once you know the tools and methods they are using to block you, you can think about how to mitigate the effects of the shutdown," says Edward Felten.**

Mubarak's last-ditch approach to censorship in Egypt was extreme, Dainotti explains, in that the dictator completely isolated entire autonomous systems from the rest of the Internet. Dainotti's team was able to see this blunt cutoff, called BGP blocking, through information about the availability of sub-networks, which are also called prefixes, that comprise the autonomous systems. Normally, Dainotti explains, "a network advertises its own prefixes, and says to the world, 'You can reach me from here.' So when they shut down the Internet, they basically withdrew the advertisement of their routes, so suddenly these prefixes disappeared from the world and nobody knew how to reach them because the autonomous systems themselves said, 'OK, I don't have these routes anymore, I'm not advertising these prefixes anymore.'" Through their analysis, the researchers were able to show that all of Egypt went down within a 10-minute period around 10:30 Greenwich Mean Time on January 27. It is unknown how the government was able to topple connectivity so quickly, but it appears network operators complied under duress. A post-outage statement issued by Vodafone, for example, mentions "the safety of our employees," notes its lack of legal options, and says if the Egyptian authorities had exercised their technical capability to close the Vodafone net-

Milestones

Gödel Prize and Other CS Awards

The European Association for Theoretical Computer Science (EATCS), the ACM Special Interest Group on Algorithms and Computation Theory (SIGACT), and the ACM Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD) recently honored 10 leading computer scientists.

GÖDEL PRIZE

EATCS and SIGACT jointly awarded the 2012 Gödel Prize for outstanding papers in theoretical computer science to three papers: Elias Koutsoupias and Christos H. Papadimitriou, "Worst-case Equilibria"; Tim Roughgarden and Éva Tardos, "How Bad Is Selfish Routing?"; and Noam Nisan and Amir Ronen, "Algorithmic Mechanism Design."

PRESBURGER AWARD

EATCS bestowed the 2012 Presburger Award, which honors "a young scientist for outstanding contributions in theoretical computer science, documented by a published paper or a series of published papers" to Venkatesan Guruswami, associate professor, Carnegie Mellon University, and Mihai Patrascu, senior member, technical staff, AT&T Labs.

EATCS DISTINGUISHED ACHIEVEMENTS AWARD

EATCS presented the 2012 Distinguished Achievements Award to Moshe Y. Vardi, Karen Ostrum George Professor in Computational Engineering, Rice University, in acknowledgment of his "extensive and widely recognized contributions to theoretical computer science over a lifelong scientific career" and "a long service record and a strong leadership in the field."

KDD INNOVATION AWARD

SIGKDD conferred the KDD Innovation Award to Vipin Kumar, William Norris Professor, University of Minnesota, for "his technical contributions to foundational research in data mining and its applications to mining scientific and climate data."

—*Jack Rosenberger*

work, service would have taken much longer to restore, implying the authorities were willing to physically cut network cables if necessary.

The censorship in Libya, on the other hand, happened through more sophisticated means, the researchers conclude. The Gaddafi regime did use BGP blocking for the first hour of the outage, which occurred on the evening of February 19. However, the network telescope revealed restoration of service and a second outage the next evening, during which the government began testing a firewall to filter packets coming in and out of Libya. The use of packet-level blocking became apparent to the researchers through a mismatch they spotted between the BGP data and the data from the telescope. Specifically, the telescope showed a disruption in traffic that the BGP data did not. "They brought the autonomous system up again after one hour, but Libya was still blacked out," Dainotti explains. "That's what makes us think they were doing tests." Once the regime had successfully tested the firewall, they proceeded to use it for the final outage several days later.

Clearly, packet filtering is a better solution for the censor. But Libya was able to use it only because the country had the right conditions in place. For one thing, as journalists discovered after the regime was overthrown, the government possessed sophisticated Internet surveillance and control equipment, enabling them to configure and test their own firewall. It also helped that the country had only two network operators, in contrast to Egypt's 50, making it far easier for Libya to bring autonomous systems up and down.

The use of multiple data sources was one of the strengths of this research, says Edward Felten, a professor of both computer science and public affairs at Princeton University and the director of Princeton's Center for Information Technology Policy. Besides the network telescope and the BGP data, Dainotti's team gathered some data from the Archipelago Measurement Infrastructure (Ark), in which 60 machines distributed around the world actively probe different parts of the Internet. Ark measurements do not offer the same level of granularity seen through IBR traffic, but they are known to be reliable, so they validated the measurements taken by the network telescope.

The UCSD team has studied large-scale outages caused by natural disasters to learn about the impact on nearby Internet infrastructure and track how connectivity was gradually restored.

Part of the value of this kind of investigation, Felten says, is being able to see precisely how governments cut off access. "Once you know the tools and methods they are using to block you, you can think about how to mitigate the effects of the shutdown," he says.

The study also shows how the Internet is increasingly important to activists and a threat to governments, according to Steven Murdoch, a security researcher at the University of Cambridge. "Cutting connections is a crude way of controlling the Internet," he says, "but it could happen again, and so care needs to be taken in any scheme which assumes ubiquitous Internet availability. Governments can and do interfere with the Internet when it suits them." Earlier this year, for example, the Iranian government tried to quell antigovernment protests by cutting off access to email and social media, though some Iranians were able to circumvent these blocks through proxy servers and VPN connections. (Predictably, the government cracked down on such users.) Of course, ongoing Internet censorship is more widespread, affecting citizens of China, Saudi Arabia, and Vietnam, among other nations; and some regimes, such as Cuba and North Korea, are so fearful of the outside world that they maintain complete control of Internet equipment.

The researchers' methods have promising applications beyond tracking government censorship, though, and Dainotti and his team have already used similar techniques to study large-scale outages caused by natural disasters, such as the earthquakes that have wreaked havoc on Japan and New Zealand. By analyzing IBR in those regions, the researchers could see the impact of each earthquake on nearby Internet infrastructure and track how connectivity was gradually restored. Over time, this kind of research can offer clues about how a particular network topology combines with other factors to make a country vulnerable to outages.

The team's more ambitious goal is to move beyond posthoc analysis to real-time monitoring of Internet infrastructure. For example, they are looking into using visualization techniques to respond to alerts of disruptive events by zooming in on just those regions for deeper analysis. "The Internet is a piece of critical infrastructure that as a world we have little formal stewardship over," Claffy says. "When a country goes off the Net, we can't not be worried about it."

Further Reading

Crovella, M. and Krishnamurthy, B.
***Internet Measurement: Infrastructure, Traffic and Applications*, John Wiley & Sons, Chichester, England, 2006.**

Dainotti, A., Amman, R., Aben, E., and Claffy, K.C
Extracting benefit from harm: using malware pollution to analyze the impact of political and geophysical events on the internet, *ACM SIGCOMM Computer Communication Review 42*, Jan. 2012.

Dainotti, A., Squarcella, C., Aben, E., Claffy, K.C., Chiesa, M., Russo, M., and Pescapè, A.
Analysis of country-wide Internet outages caused by censorship, Internet Measurement Conference 2011, Berlin, Germany, Nov. 2–4, 2011.

Xu, X., Mao, Z. M., and Halderman, J. A.
Internet censorship in China: Where does the filtering occur? *Proceedings of 12th Passive and Active Measurement Conference*, Atlanta, GA, March 20–22, 2011.

Based in San Francisco, **Marina Krakovsky** is the co-author of *Secrets of the Moneylab: How Behavioral Economics Can Improve Your Business.*