Detection and analysis of large-scale Internet infrastructure outages

http://www.caida.org/funding/dals-satc/

summary

Despite the Internet's status as a critical infrastructure of our society, there is little scientific instrumentation dedicated to monitoring global Internet behavior. In particular, we have no rigorous framework for measurement, analysis, or quantifying the impact of network outages, filtering, or other abnormal connectivity dynamics on a global scale.

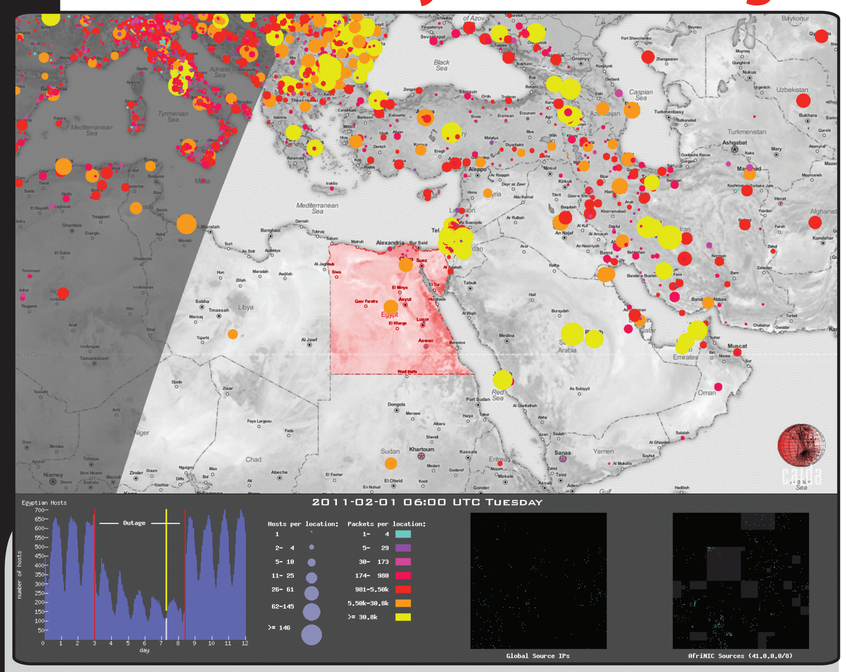
We have developed and demonstrated a methodology that can identify not only which networks have been affected by an outage, but also which techniques have been used to effect a deliberate disruption (e.g., control plane vs. data plane intervention). We have also developed metrics to quantitatively gauge the geographic and topological extent of impact of geophysical disasters on Internet infrastructure, and techniques to investigate the chronological dynamics of the outage and restoration. Our approach relies on:

- the extraction of signal from a pervasive and continuous source of malware-induced background radiation in Internet traffic (IBR);
- combining multiple types of data (active probing, passive IBR measurement, BGP routing data, and address geolocation and registry databases) to assess the scope and progression of the outage.

This project will result in an experimental operational deployment to validate and extend an empirically grounded methodology **for detection and analysis of large-scale Internet outages.** In addition to improving our understanding of how measurements yield insights into network behavior, and strengthening our ability to model large scale complex networks, use of such a system will also illuminate infrastructure vulnerabilities that derive from architectural, topological, or economic constraints, suggesting how to mitigate or eliminate these weaknesses in future Internet architecture and measurement research.

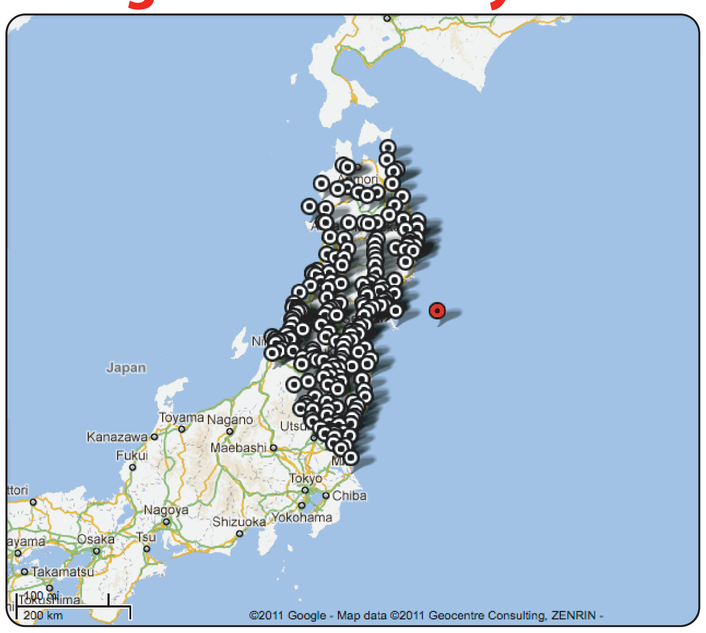
A deployed platform will detect and monitor connectivity disruption and censorship events on a planetary scale thus enabling situational awareness of the nature and causes of network outages to national decision-makers who must determine the type and extent of proper response.

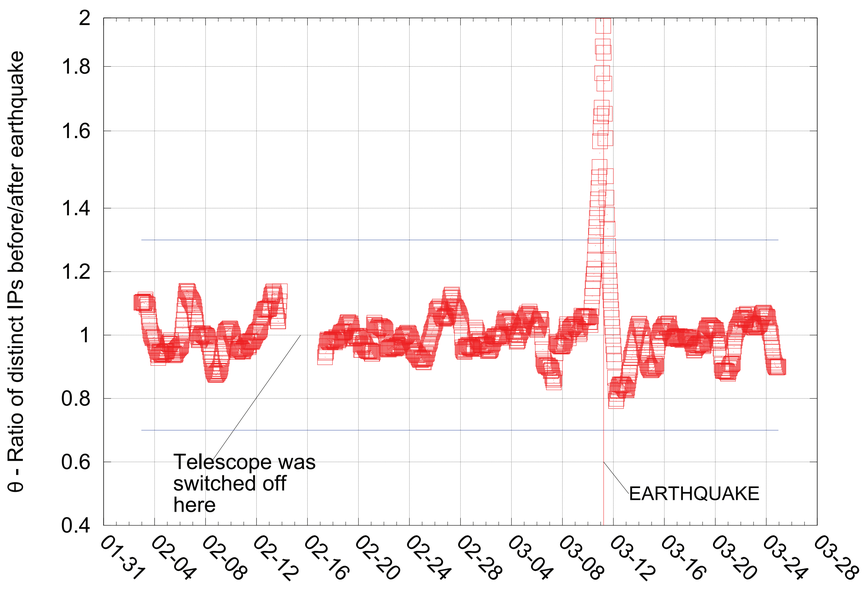
country-level outages



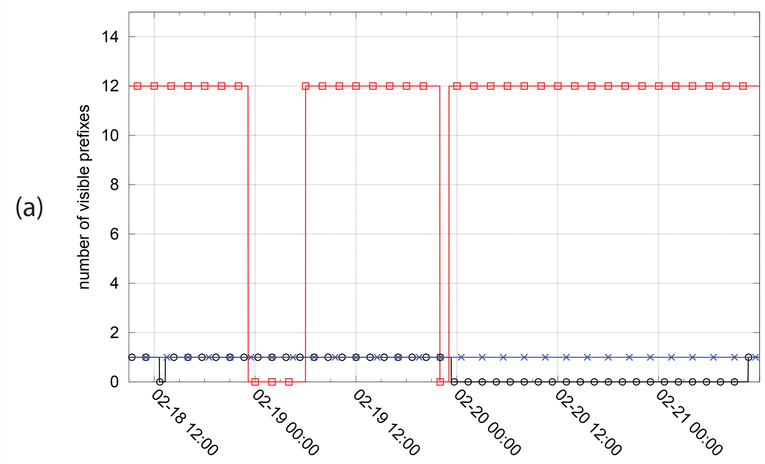
The Egyptian Internet Blackout in 2011 during the "Arab Spring": combining CAIDA's Cuttlefish and Measurement Lab's ipv4-heatmap visualization tools provides insight into the outage. A single, concise view conveys information about the geolocation of hosts sending Conficker-like packets to the UCSD network telescope over time. The map displays the geographic proximity of hosts with the dimensions of number of hosts and number of packets indicated by size and color respectively. The graph, in the bottom left, of total number of Egyptian hosts seen per hour corroborates the sudden decrease in traffic. IPv4 heat maps have the potential to reveal patterns in the address space; in this view the entire IPv4 space and the addresses delegated to AfriNIC with Egyptian IP addresses shaded in blue are represented.

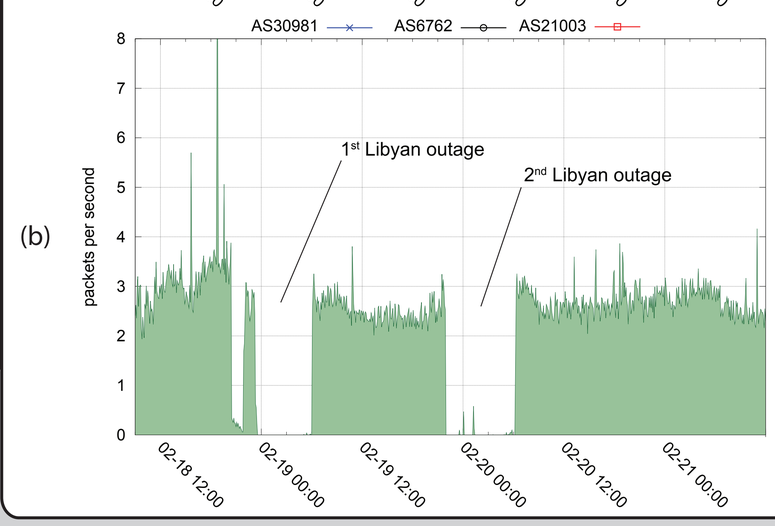
outages caused by natural disasters





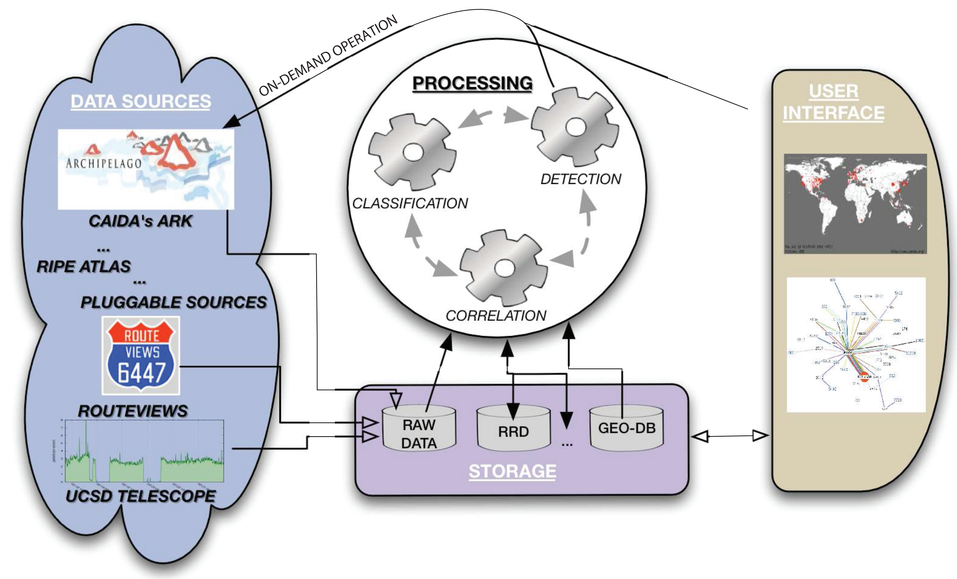
Japan Earthquake/Tsunami in 2011: visual representations and traffic metrics reveal the disruption caused by a natural disaster. Shown here is a map of the geolocated networks affected by the Tohoku earthquake and tsunami. The metric theta (graph on the right) is a ratio comparing the number of unique IP addresses seen in IBR before and after an event from a geographic region. Plotting theta for addresses within a certain radius of the earthquake's epicenter reveals that a large portion of the normally background-radiating IP addresses fall silent, i.e., lose connectivity to the global Internet.





The Libyan Outage in 2011: fusion of control plane and data plane data sets revealed insight into censorship methods used in this recent large-scale outage. Figure (a) shows the visibility of Libyan IPv4 prefixes in BGP (RouteViews and RIPE NCC RIS data) during the outage in February 2011; (b) shows the rate of unsolicited traffic to the UCSD telescope from Libya over the same period. The second outage, unlike the first, was not implemented using BGP withdrawls; the re-announcement of Libyan BGP prefixes and the absence of traffic in the data plane (observed by the darknet) is indicative of packet filtering.

architecture overview and timeline

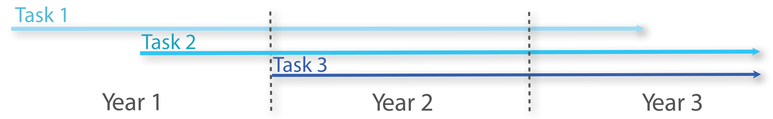


The diagram to the left illustrates the main components of the proposed system. The system will fuse pre-existing data sources collected by other Internet measurement infrastructures, creating a rich archive for subsequent processing and visualization. Processing modules will include data sanitization, classification, geolocation, AS mapping, and data aggregation. Other modules will extract and correlate statistics from these aggregated data, and store output in efficient data structures that will feed data visualization and change-point detection algorithms. A web platform will provide data visualization and a user interface through which the user can locate an event on a geographical map and track metrics, e.g., impact, calculated over time.

Task 1 : investigating and defining strategies and methodologies for how to combine multiple heterogeneous data sources to detect and characterize outage events (Years 1, 2, and 3);

Task 2 : defining (and refining) the system requirements for continuous monitoring and (near) real-time analysis of outages as they occur (will start in the second half of Year 1);

Task 3 : testing and experimental deployment of such a system (Years 2 and 3).



November 2012

team

Justin Cheng | Bradley Huffaker | Karyn Benson
Alistair King | Emile Aben | Alberto Dainotti | kc claffy

sponsored by

Funding source:
NSF CNS-1228994.

UC San Diego
SDSC SAN DIEGO SUPERCOMPUTER CENTER
caida