MapKIT: Mapping Key Internet Terrain

PI Alberto Dainotti, CAIDA, UC San Diego
In collaboration with Paul Barford, University of Wisconsin Madison
CNS-1705024
http://www.caida.org/funding/satc-mapkit/

Challenge

To apply a military analogy to Internet research, the science of cybersecurity has focused heavily on weapons and tactics, but has largely neglected "terrain". Strategic points in the macroscopic Internet topology constitute key terrain in the cyberspace battlefield. Hackers, terrorists or nationstates - can disrupt, intercept or manipulate the Internet traffic of entire countries or regions by targeting structural weaknesses of the Internet topology. Despite much recent interest and a large body of research on cyber-attack vectors and mechanisms, we lack rigorous tools to reason about how the macroscopic Internet topology of a country or a region exposes its critical communication infrastructure to compromise through targeted attacks.

Part of the problem is that collecting and interpreting data about the Internet connectivity, configurations and associated vulnerabilities is challenging. Due to the massive scale and broadly distributed nature of Internet infrastructure and the scarcity of publicly available data, we must resort to complex measurement and inference methodologies that require significant effort in design, implementation, and validation.

Through a novel multi-layer mapping effort, the MapKIT project aims at identifying important components of the Internet topology of a country/region -- Autonomous Systems (ASes), Internet Exchange Points (IXPs), PoPs, colocation facilities, and physical cable systems which represent the "key terrain" in cyberspace.

Scientific Impact

The results of this project will contribute to a better understanding and modeling of key weakness of the Internet infrastructure in countries that expose them to targeted attacks by state actors and organized groups. Our results also provide data and models relevant to resiliency of critical infrastructure (e.g., preparedness to natural disasters).

Finally, our results will provide input for polictical scientis and international relations researchers to reason about "opportunity and willingness" to engage into large-scale cyber conflict

AS-Level Transit Influence

We tackle the problem of identifying the most influential transit providers in each country that may have potential to observe, manipulate or disrupt a significant fraction of the Internet traffic flowing towards that country.

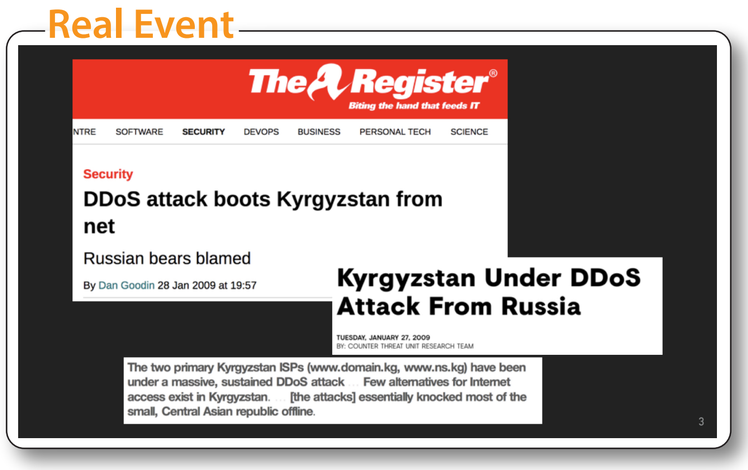
Metrics

We develop two distinct metrics. AS-Level Transit Influence (ATI) is a pairwise metric between a single transit AS and a single origin AS: intuitively it captures the fraction of each origin AS' announced addresses that the transit AS serves. Country- Level Transit Influence (CTI), on the other hand, is a measure of the same type of influence of a single transit AS on all the country's addresses.

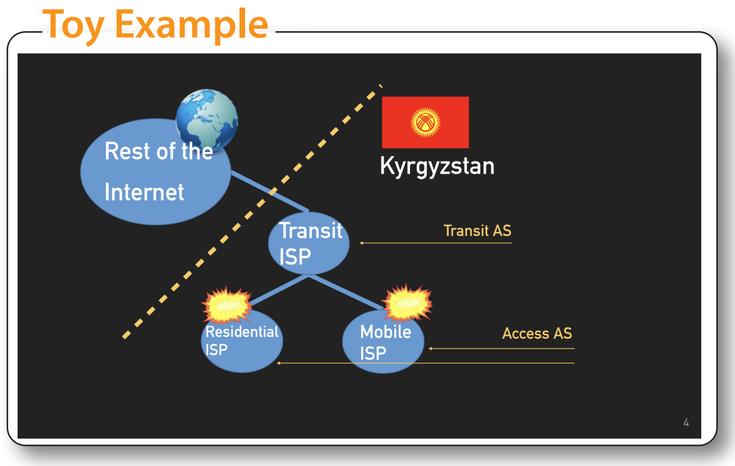
Weighting and Filtering

1) Prioritizing AS Diversity
2) Indirect Transit Filtering
3) Provider-Customer AS Filter
4) Inbound Path Filtering

Real Event



Toy Example



AS-Level Transit Influence (ATI). The *transit influence* metric is the fraction of address in announced, preferred paths (excl. backup links and less preferrred paths) towards orgin prefixes of a given AS where a particular transit AS is present. In a given country, the transit influence *ATI* (AS_t ,AS_o) $\in$ [0,1] for each transit AS_t on origin AS_o is expressed as:

$$ATI(AS_t,AS_o) = \frac{1}{A(AS_o)} \sum_{p_i \in AS_o} \sum_{m \in M} \frac{A(AS_t,p_i) \cdot w(m)}{d(AS_t,p_i) \cdot G} \quad (1)$$

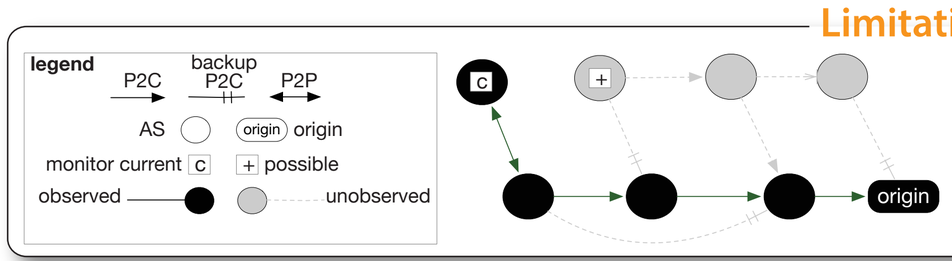
where $A(AS_t, p_i)$ is the number of addresses in prefix p_i originated by *AS* where AS_t is present in the preferred AS path, $w(m)$ is monitor m's weight among the set of BGP monitors M and G is the number of ASes hosting BGP monitors, $A(AS_o)$ is the total number of addresses originated by origin AS_o that are geolocated to the country, and $d(AS_t ,p_i)$ is the number of AS-level hops between AS_t and prefix p_i originated by AS_o.

Country-Level Transit Influence (CTI). ATI captures transit influence of a transit AS on an individual origin AS. At the country level, we want to know if a transit AS is influential on many of the country's addresses. Then, to measure the impact of an AS_t on the transit connectivity of the country as a whole, we compute an analogous metric as ATI (Eq. 1) on the prefixes of the country:

$$CTI(AS_t,C) = \frac{1}{A(C)} \sum_{p_i \in C} \sum_{m \in M} \frac{A(AS_t,p_i) \cdot w(m)}{d(AS_t,p_i) \cdot G} \quad (2)$$

Where p_i is each prefix with addresses geolocated to country *C*. Note that we do not compute transit influence where $AS_t = AS_o$; i.e., we do not count the influence of the transit AS on itself, as that is captured by the share of the country's address each AS originates.

Limitation of BGP Measurements



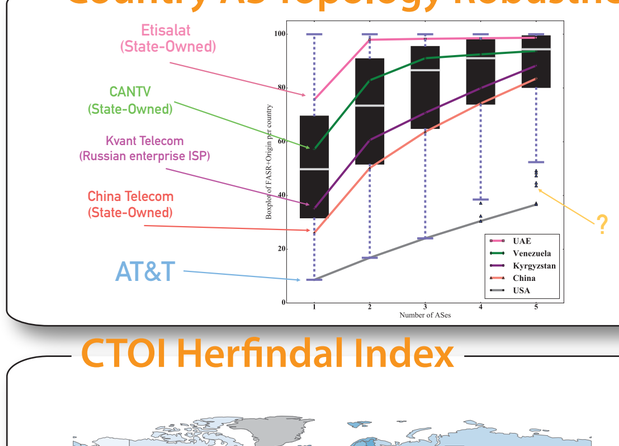
Observed topology (black ASes) is limited by the location of the BGP monitor; since we do not have a monitor in any of the neighbors of a light gray AS (e.g., in the location marked with a plus sign) the entire top AS-chain is unobserved.

Preliminary Results

Building on CTI, we investigate—at a global scale—some relevant geo-political questions pertaining to the exposure to (internal/external) control of a country's connectivity. We look into: (i) when a country's connectivity relies on a restricted set of ASes, potentially exposing it to high-profile attacks; (ii) opportunity for state control; (iii) potential foreign influence.

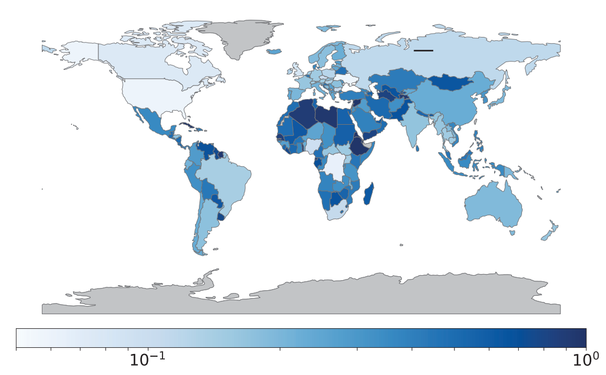
Our examination of control exposure is performed at the AS topological level and quantified in terms of potentially- affected IP addresses originated in a country. For this reason, in addition to transit influence, we also take into account the control that can be directly exerted by an AS on the address blocks that it itself originates. In order to combine these two forms of "influence" we introduce a metric called CTOI (Country Transit/OriginInfluence).

Country AS Topology Robustness



Here we explore the robustness of each country's AS topology to a specific scenario: how much of the country's address space is exposed if an attacker successfully compromises— e.g., disables the connectivity of, infiltrates or wiretaps—a small number of highly influential ASes

CTOI Herfindal Index



We use the Herfindahl index to infer concentration among a small number of ASes, an economic metric used to estimate how concentrated a market is on a firm or small set of firms, e.g, to evaluate antitrust concerns of mergers. The Herfindahl index is computed as follows: $H = \sum_{i=1}^{10} f_i^2$, where f is the CTOI of each of the ASes in each country (sorted decreasingly). The squaring of the address space fraction makes the contribution of the largest producers more dominant, and discounts the influence of small players. To intuitively understand this metric, it is useful to imagine the extremes: H = 1, one AS has CTOI over the entire address space; H > 0, the number of ASes approaches infinity, and their CTOI fraction is uniform. Lower values of the Herfindahl index signal a more robust AS topology at the country level.

Team: Alberto Dainotti, Amogh Dhamdhere, Alex Gamero-Garrido, KC Claffy, Alex C. Snoeren, Bradley Huffaker, Shuai Huao, Esteban Carisimo

As the Internet has become probably the critical infrastructure on which all other critical infrastructures depend, safety and prosperity of our society as well as international relations depend on cybersecurity. Yet the exposure of a country's Internet macroscopic infrastructure to targeted attacks with potential massive impact (e.g., in the context of cyberwarfare) is unclear. This project promises to bridge this crucial gap.

Education and Outreach

Curriculum Development
- 1Postdoctoral researcher
- 3 Graduate Students
- 7 Undergraduate Students

Presentations
- DHS S&T Cyrie Internet Infrastructructure Risk Economics Stakeholder Exchange Meeting
- CAIDA AIMS 2019 workshop
- CAIDA IMAPS 2018 workshop

Publications
- 1 Journal paper (ACM SIGCOMM CCR Jul 2019
- 2 Conference papers (TMA 2019)