COS 561 Project Proposal

Annie Edmundson, Caroline Trippel, Elba Garza

October 7, 2014

1   Motivation

It is becoming more common for ISPs to be restricted by the policies of the country in which they reside; if all ISPs in a country follow the same national policies, it could have a global impact on Internet traffic. This is particularly evident for countries that are known for their national policies, such as wiretapping or censorship. Because there is incomplete and incorrect data about Autonomous Systems (ASes), traffic paths, and geo-location information, it is difficult to address the questions: Which countries does traffic originating from a certain country traverse? How many countries rely on China (known for censorship) to transmit their traffic? How many countries rely on the United States (known for wiretapping) to transmit their traffic? It is likely that the answers to these questions would change based on when they were asked.

This motivates a longitudinal study of how central a country is, in terms of transmitting Internet traffic. Not only is it important to measure how central a country is at the present time, but also how central they were in the past. This raises more questions: Which countries transmit more traffic for other countries now than they did in the past? Which countries transmit less traffic for other countries now than they did in the past? Why did their centrality change?

2   Goals

We have three primary goals for this project:

1. Conduct a longitudinal study of country centrality to measure how it has evolved over time.
2. Compare different techniques for determining country centrality.
3. Verify our results or a subset of our results.

3   Methodology

In order to meet the goals outlined above, we have planned a methodology that includes: (1) gathering data, (2) determining country paths, (3) measuring centrality, and (4) verification.

3.1   Gathering Data

We will gather data from multiple public databases.

- **iPlane.** iPlane is a project out of the University of Washington (`http://iplane.cs.ucr.edu/data/data.html`) that archives traceroutes.

- **The Flattening Internet Topology.** This dataset (`http://ita.ee.lbl.gov/html/contrib/gill-PAM08.html`) consists of traceroute data collected in 2007.
- **Team Cymru.** This project (`http://www.team-cymru.org/`) keeps track of registry allocated prefixes and the corresponding country code and AS mappings.
- **RouteViews.** The RouteViews Project (`http://www.routeviews.org/`) provides archived BGP data, including AS paths.
- **traceroute.org** This service (`http://www.traceroute.org/`) can provide a traceroute from 73 different countries to the client's location.

3.2   Determining Country Paths

There are two different methods for determining country paths: (1) mapping IP to AS and then AS to country, or (2) mapping IP to country. Here we give an overview of both:

1. The first step is to map all prefix sources/destinations to an AS; in order to do this, we use the algorithm in [1], which is a simplified model of the BGP protocol. The next step is to map traceroutes to AS and country paths. We will query Team Cymru (`http://www.team-cymru.org/`) for each IP address in a traceroute to get country and AS level information. The last step is to map an AS path to a country path; we will use the algorithm presented in [1].
2. This technique is only slightly different from the first technique. We will follow the same first step of mapping all prefix sources/destinations to an AS. The second step differs; instead of querying Team Cymru, we will geo-location services (`http://www.iplocation.net/`) to determine the location of each IP address. We will use Team Cymru just for AS level information. The last step of mapping an AS path to a country path remains the same as the first technique.

3.3   Measuring Centrality

We will use the metrics in [CITE] to determine Country Centrality, as well as Strong Country Centrality. We will make a slight modification for better accuracy: weight countries by demographic information, such as population, to account for small countries (particularly in Europe).

3.4   Verification

We would like some form of verification of our centrality measurements. We plan to use `www.traceroute.org` to use new traceroutes from other countries and compare the countries they traverse to our measurements.

4   Caveats

There are many limitations and assumptions involved in this work. Many datasets are incomplete, and some do not have completely accurate information. Our methods of inferring country level paths propogate the incompleteness and inaccuracy. Additionally, we have no way to verify our measurements of past centrality.

References

[1] Josh Karlin, Stephanie Forest, and Jennifer Rexford. "Nation-State Routing: Censorship, Wiretapping, and BGP." *arXiv preprint arXiv:0903.3218*, 2009.