The Resilience of the Internet to Colluding Country Induced Connectivity Disruptions

Peter Mell
National Institute of Standards and Technology
peter.mell@nist.gov

Richard Harang
U.S. Army Research Laboratory
ICF International
richard.e.harang.civ@mail.mil

Assane Gueye
University of Maryland
agueye@umd.edu

***Abstract*— We show that the strength of Internet-based network interconnectivity of countries is increasing over time. We then evaluate bounds on the extent to which a group of colluding countries can disrupt this connectivity. We evaluate the degree to which a group of countries can disconnect two other countries, isolate a set of countries from the Internet, or even break the Internet up into non-communicative clusters. To do this, we create an interconnectivity map of the worldwide Internet routing infrastructure at a country level of abstraction. We then examine how groups of countries may use their pieces of routing infrastructure to filter out the traffic of other countries (or to block entire routes). Overall, bounds analysis indicates that the ability of countries to perform such disruptions to connectivity has diminished significantly from 2008 to 2013. However, we show that the majority of the gains in robustness go to countries that had already displayed significant robustness to the types of attacks that we consider. The countries that displayed higher initial vulnerability to such attacks did not become significantly more robust over the time period of analysis.**

Keywords*— *Internet, resilience, connectivity, autonomous systems, security, countries

I. Introduction

The Internet was designed to be robust to disruption due to the failure of specific networks or routers [1]. However, events have empirically demonstrated that despite this design, single points of failure (e.g., the digging up of a fiber optic cable, or the cutting of a submarine line by a ship anchor) are capable of disconnecting entire countries from the rest of the internet [2] [3] [4]. In this study, we examine the potential of deliberate disruption or filtering enacted on inter-country network routes. In particular, we examine to what extent a group of colluding countries can disrupt Internet connectivity for other countries in several ways: by disconnecting two other countries, isolating a set of countries from the Internet, or even breaking the Internet up into small non-communicative clusters.

To do this, we create an interconnectivity map of the worldwide Internet routing infrastructure, aggregated at the level of individual countries. We then apply graph algorithms to determine bounds on the degree to which groups of colluding countries of varying sizes can adversely affect the connectivity of other countries. We evaluate pairwise connectivity, node partitioning while minimizing the maximal cluster size, and node partitioning while maximizing the number of clusters. We perform these experiments on graphs of Internet interconnectivity, at the country level of abstraction, covering the years 2008 to 2013.

The data indicates that the overall ability of countries to disrupt connectivity in the ways we examine has consistently decreased from 2008 to 2013. However, this increase in robustness appears to be limited to sets of countries that already exhibited high degrees of robustness to such disruptions. The remaining countries remained relatively vulnerable to connectivity disruptions by groups of colluding countries. This “rich get richer” scenario has been observed in generative models of scale-free graphs, such as interconnectivity of the World Wide Web [5], where entities with the highest number of links receive a higher share of new links. We hypothesize that countries with fewer independent connections to the Internet find it difficult to obtain additional links, and so remain vulnerable to connectivity disruptions by groups of colluding countries (typically the more resilient highly connected countries).

Due to the complexity of the worldwide Internet and limitations in the measurement infrastructure, our interconnectivity map is a proper subset of actual worldwide connectivity. We also do not include policy-based routing restrictions (the ‘valley-free’ condition of [6]), which are known to limit connectivity as well [7] [8] [9]. For that reason, our results provide bounds to the damage that can be done, as opposed to exact measurements. Because of these same data limitations, we are unable to perform a precise comparative study of countries, and so we focus on overall trends rather than on any specific countries.

This work constitutes the first study to propose, quantify, and measure this important class of Internet threats. To our knowledge, this is also the first study to construct a country to country connectivity graph of the Internet to study security. Furthermore, we provide a defensible methodology to provide trending analyses on this graph in the presence of only partial data. Indeed, at the beginning of our study we thought that the data limitations would prevent us from obtaining rigorous results and an important contribution of this work is to show how to work through those data limitations.

SENT ’15, 8 February 2015, San Diego, CA, USA
ISBN 1-891562-39-8
http://dx.doi.org/10.14722/sent.2015.23007

The rest of the paper is organized as follows. Section 2 discusses the background on the data set we used to model inter-country Internet topology and limitations surrounding the use of that data. Section 3 discusses our actual data collection activities and provides general statistics on the resulting country connectivity graphs. Section 4 describes our experiments and section 5 provides our results. Section 6 discusses related work. Section 7 concludes and discusses future work.

II. Background

The Cooperative Association for Internet Data Analysis (CAIDA) has a worldwide monitoring network that provides an approximate topological map of the Internet at the Internet Protocol (IP) layer [10]. It then uses the RouteViews [11] Border Gateway Protocol (BGP) data to collapse the IP topology into a map of autonomous systems (ASs), approximately the set of Internet Service Providers [10]. More formally, an AS is a set of routers under common management where the group of routers presents a unified routing policy (to other ASs and to a set of network prefixes for which it provides Internet access) [12]. Individual ASs are the basic units of routing policy, and together collectively form the routing infrastructure of the Internet. CAIDA maps each AS to the country in which that AS is registered. We use this to condense the global AS map into a network of countries, and then evaluate inter-country connectivity (the handling of multinational ASs is discussed in section 3).

A. Archipelago Infrastructure

CAIDA continuously updates its IP level topological map through the employment of its Archipelago (Ark) measurement infrastructure. As of 2014-05-09, Ark had 94 monitors distributed worldwide, separated into three teams. Every 2 to 3 days, each team uses a traceroute-like procedure to probe a random IP address within each /24 subnet in the IPv4 address space. This yields a list of routers connecting the monitor to the target IP. The monitor to target IP mapping varies randomly so that, over time, each subnet is accessed from many different parts of the world, revealing the primary pathways through the Internet.

As shown in Fig. 1, the monitors cover every continent with the exception of Antarctica, although the majority of monitors are placed in North America, Europe, and Southeast Asia. Monitors are sparse in Africa (5) and South America (4). There are no monitors in the Middle East, Eastern Europe, or in Russia. The dearth of monitors in a particular geographical area does not prohibit probing IPs in that region, however it does present limitations in mapping connectivity between regional ASs, as we discuss below.

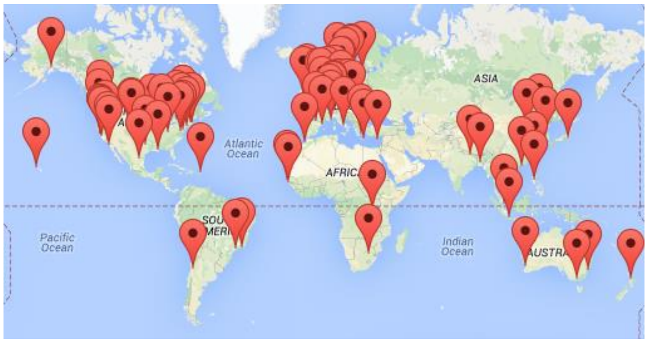

Fig. 1. Physical Location of Ark Monitors (2014-05-12) (map provided by [5])

B. Data Limitations

The Ark infrastructure is only capable of discovering preferred paths to and from ASs containing a monitor. Routes between ASs that do not lie on a preferred route from a monitor to a target AS will not be discovered. Another limitation is that we must drop ASs that cannot be mapped to a particular country (0.11 % of the ASs). The Ark infrastructure also identifies routes between ASs as either direct or indirect. With a direct route, Ark sees an IP in one AS directly communicate with an IP in a second AS. With indirect routes, the two ASs are separated by one or more IPs for which an AS could not be identified (either the AS was not registered for the IP or the IP was non-responding). We omit indirect routes from our graphs since we do not know to which country the intermediate ASs belong.

These three limitations cause us to ignore many potential routes that could contribute to the strength of connectivity of the Internet (often localized to specific parts of the globe). Despite this, we can still perform rigorous and defensible experimentation by focusing on calculating bounds. In particular, we focus on calculating the maximum disruption a group of countries can cause to global Internet connectivity or to connectivity between arbitrary pairs of countries, given our incomplete measurements of the total connectivity. Future improvements to Ark relative to these limitations will enable more accurate measurements that will allow us to sharpen these bounds.

Lastly, the Ark monitors are biased towards detecting routes with high capacity since these are likely to be the preferred routes announced by ASs. However, we cannot measure route capacity and are thus limited to evaluating the basic connectivity between countries.

III. Data Collection

We downloaded all CAIDA data files on AS connectivity from 2008 to 2013 inclusive (2855 data files totaling 849 MB). We created a graph for each of these 6 years using the NetworkX graph library version 1.8.1 [14] and Python version 2.7.6. As discussed above, we removed AS nodes for indirect links and where country information was not available (as of 2014-05-16). We also model all direct links as bidirectional since, for our experiments, we are concerned about the capability to transmit data as opposed to currently policy-based directionality of traffic flow. The number of ASs grew from 28821 in 2008 to 44390 in 2013 with the number of edges growing from 111487 to 213883. There was thus a 54 % growth of the number of ASs and the 92 % increase in the number of observed edges from 2008 to 2013.

We then used the AS to country mappings to aggregate the AS nodes into country nodes where edges represent inter-country connectivity. Table 1 shows the number of countries and edges in each year of data. The number of countries participating in the global Internet routing infrastructure increased 8.7 % during the timeframe we examine, while the visible inter-country edges increased 54.5 %. Note that the current upper bound on the possible number of visible countries is 249 as defined by the "officially assigned" International Standards Organization (ISO) country codes [15].

TABLE 1: NUMBER OF COUNTRIES AND EDGES PER MONITORED YEAR

Year	Number of Countries	Number of Edges	Number of Monitors
2008	206	2235	33
2009	211	2343	42
2010	218	2644	54
2011	219	2925	59
2012	221	3138	65
2013	224	3452	89

The graphs from which the information in Table 1 is derived are useful for evaluating bounds on the resilience of a particular year. However, the addition of new monitors with new vantage points, as well as the retiring of old monitors, means that the visible portions of the routing graph vary significantly from year to year, independent of changes in the underlying routing graph itself. This makes it difficult to reliably compare the resilience between different years. To enable year to year comparisons, we restrict the discovered routes to those visible to a set of 24 monitors that were active in all 6 years of our evaluation. Table 2 shows the number of countries and edges in each year of this restricted data set.

TABLE 2: NUMBER OF COUNTRIES AND EDGES PER MONITORED YEAR FROM A PERSISTENT SET OF MONITORS

Year	Number of Countries	Number of Edges	Number of Monitors	% Excluded Monitors
2008	206	2149	24	27 %
2009	210	2243	24	43 %
2010	218	2370	24	56 %
2011	219	2513	24	59 %
2012	221	2731	24	63 %
2013	223	2829	24	73 %

Unfortunately, this approach is biased against the latter years, making them appear less resilient. The problem is that the percentage of excluded monitors increases from 27 % in 2008 to 73 % in 2013 and these exclusions are tightly correlated to the amount of data that is discarded. Since every IPv4 subnet is scanned 3 times every 2 to 3 days regardless of the number of monitors, this means that in the latter years we are discarding a larger and larger percentage of these periodic scans (reducing our chances of finding new routes). As we will show, despite this limitation in showing the full resiliency of the latter years, the data still suggests the increasing resiliency of the Internet over time (although this limitation in our data will cause us to underestimate the strength of the increase).

A visualization of country connectivity for 2013 is shown in Fig. 2. The nodes are sized proportional to their degree. The 140 nodes with degree greater than 10 are shown in white (mostly in the center). The 84 nodes of degree 10 or less are shown in red on the periphery. Notice how the red nodes primarily connect to the white nodes. Of the 3452 edges, only 14 edges (0.41 %) connect two red nodes. This suggests that the graph has a degree of negative assortativity, as is common in many communications networks [16]. Connectivity-poor nations tend to be connected to connectivity-rich nations, and only rarely to other low-degree countries. Notice also the large number of high degree nodes in the middle. While the connections are not visible in the graph, these high degree nodes are densely mutually connected, forming an extremely resilient core for Internet communications.

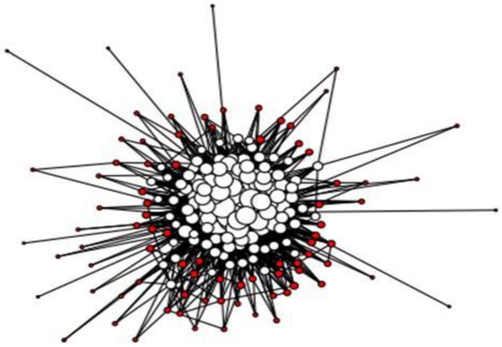

Fig. 2. Country Connectivity for 2013 using all monitors

One aspect of the data that deserves special consideration is the existence of multinational ASs (MOAs). MOAs are ASs that have points of presence (PoPs) within multiple countries. Some MOAs have PoPs in many countries on multiple continents. For our work, we map MOAs to their home country (i.e., country of registration). Under this model, a MOAs is required to implement Internet filtering laws or government directives from their home country regardless of the physical location of the routers they own. This is similar to the U.S. export control system [17] where national security sensitive technology is not allowed to be sent to a target list of countries regardless of the physical location of a branch of a company. The rationale for our approach comes from the legal literature: "the home country may also have laws that attempt to regulate business activities of the company that are conducted outside the home country" [18]. Also see [19] [20] for similar analyses. The work of [9] briefly examined the assumption that 'national incumbent telecom operators will follow orders from their respective governments,' and found no 'behavior that suggests otherwise.' If for a particular country, the strength of control over a MOA is less than represented by our model, this presents another limitation analogous to the previously stated data limitations that only lower our upper bounds (or raise our lower bounds) and thus do not nullify our results.

IV. Experiment Description

In our experiments, we evaluate 3 Internet connectivity security questions pertaining to our country level abstraction:

1. **Cutting Pairwise Communications**: What is the minimal number of colluding countries required to prevent two other countries from communicating?
2. **Country Isolation**: What is the maximal number of countries that can be cut off from the Internet by a group of colluding countries?
3. **Non-Communicative Clusters**: Into how many non-communicative clusters can a group of colluding countries divide the Internet?

These questions are modeled as countries completely cutting off routes to other countries. While this certainly can be done, we use this approach to model a large class of attacks (possibly as yet undiscovered) whereby selective traffic along country to country routes may be maliciously handled.

We evaluate each of these questions against the connectivity graphs covering the years 2008 to 2013. We use all the monitor data (i.e., Table 1) to create lower bounds for question 1 and upper bounds for questions 2 and 3. Note that these bounds are not comparable year to year because of the shift in monitor placements over time (discussed in section 3). To enable year to year comparison, we then use the persistent monitor data (i.e., Table 2) to provide us resiliency trends for the 3 questions using a static set of 24 monitors.

For the analysis of our experiment, we introduce the following notations. We assume that the network is given as an un-weighted undirected graph $G=(V,E)$, where V is the set of nodes and E is the set of edges, with $|V|=n$ and $|E|=m$. We denote by S the set of colluding countries (i.e., nodes of the graph) with its cardinality given by $|S|=k$. We let $G\setminus S$ designate the graph of the network when nodes in S are removed, along with their associated edges (i.e., the colluding countries stop forwarding packets). With the nodes S removed, the graph $G\setminus S$ may be disconnected into a set of clusters unable to communicate with each other (i.e., disjoint connected components). We let $\mathcal{C}(G\setminus S)$ be the set of connected components, with its cardinality denoted by $|\mathcal{C}(G\setminus S)|$. Finally, we let $C_{\max}(S)$ be the component in $G\setminus S$ with the maximum number of nodes. Notice that in this paper, we are only interested in the size of the maximal cluster $|C_{\max}(S)|$. For notational convenience, we drop the dependence on S and only use $C_{\max}$ and $|C_{\max}|$ in the rest of this paper when there is no chance of confusion.

Note that these experiments are scoped to model a set of countries passing laws or directives relative to Internet filtering that pertain to their own companies (wherever the physical routers may reside). Out of scope is modeling a country exerting control over routers physically in its geographic boundary that are owned by a company from another country (this could be considered in future work). Also out of scope are actions that countries might take in response to Internet filtering laws enacted by another country (e.g., nationalizing routers or peering points within their geographic boundaries).

A. *Cutting Pairwise Communications*

The communications of two countries, *s* and *t*, are considered prevented or cut if the ASs registered to s cease having connectivity the ASs registered to t. To evaluate the cutting of pairwise communications (our first security question), we iterate over all pairs of countries (*s*,*t*) that are not adjacent and determine the minimum, maximum, and mean node connectivity. The node connectivity calculation determines the minimum number of nodes required to disconnect the graph such that *s* and *t* end up in separate components. We used a NetworkX implementation that is based on the Ford and Fulkerson flow algorithm [21].

In our modeling of the problem, we remove the nodes corresponding to the ASs of particular countries in order to separate *s* and *t*. The removed nodes provide the minimal set of countries that would need to require their ASs to filter out traffic between *s* and *t*. Thus, overall Internet connectivity and AS functionality would remain, with *s* and *t* suffering from a lack of connectivity.

B. *Country Isolation*

A country is considered cut off from the Internet or isolated if the ASs registered to that particular country cease having connectivity to the largest remaining connected component of the global Internet. In the country isolation problem (our second question) we would like to know the maximal number of countries that can be cut off from the Internet (i.e., the largest remaining connected component) by a group of k colluding countries. Recall that we use S (with $|S| = k$) to designate a generic group of colluding countries and $|C_{max}|$ for the size of the maximum connected component of the graph of the network once the nodes in S are removed. With these notations, the country isolation problem can be cast as the following combinatorial optimization.

$$\max_{S\subseteq V,|S|\leq k}(n-k-|C_{max}|), \qquad (1)$$

$n=|V|$ is the total number of nodes in the network.

For a given network and a fixed number of colluding countries k, this is equivalent to minimizing the size of the maximal cluster.

$$\underset{S\subseteq V,|S|\leq k}{min}(|C_{max}|). \qquad (2)$$

For fixed k, this is obtained when all components are of as equal size as possible. This problem is known to be in general NP-hard [22] [23] [24].

To work around this, we resort to 5 different heuristic algorithms (described in section 4.4). For each value of the number of colluding countries k, we run each of the algorithms and choose the best result as our approximation for the solution of the optimization.

To evaluate country isolation, we iteratively increase the size k of the set of colluding countries and, at each iteration, use our algorithms to choose a specific set of country nodes that (approximately) minimize the maximal cluster size. We then calculate the number of isolated countries as equal to the total number of countries minus the number of colluding countries minus the size of maximal cluster (i.e., $n-k-|C_{max}|$).

C. Non-Communicative Clusters

The Internet is considered to be broken into non-communicative pieces if the ASs registered to a group of countries are connected while being disconnected from the rest of the Internet. In our country graph, there will be multiple isolated clusters (possibly consisting of just single nodes) after removing the colluding nodes. In this non-communicative clusters problem, the goal of the colluding countries is to maximize the number of non-communicative clusters. This also can be modelled as the following combinatorial problem

$$\max_{S \subseteq V, |S| \leq k} (|\mathcal{C}(G \setminus S)|). \qquad (3)$$

This is similar to country isolation except that instead of optimizing on the number of isolated countries, the algorithms must optimize on the number of isolated clusters. We use the same set of algorithms to compute an approximate solution. As in the previous subsection, to evaluate non-communicative clusters (our third security question), we iteratively increase the size *k* of the set of colluding countries and, at each iteration, we use our algorithms to choose a specific set of country nodes that best approximate the maximum above.

D. Heuristics Algorithms

As stated earlier, solving the country isolation problem (or the non-communicative clusters problem) is in general NP-Hard. While many approximation algorithms exist for graph cut problems where edges are to be cut (e.g., see [25] for a survey), there appears to be much less work on vertex-based cuts (although see [26]). In this paper, we combine five heuristic algorithms to find approximate solutions to the graph partition problems we describe when nodes rather than edges are to be removed from the graph to create the cuts. The algorithms are all iterative and are described below.

1. Iterative removal of maximal degree node (DEG): This approach is motivated by the observation that scale-free networks, such as the Internet, are very sensitive to attacks that target nodes with largest degrees [27]. In our DEG algorithm, we iteratively remove the node with the maximum degree in the remaining graph as well as all edges that are incident to it. After each iteration, we re-compute the degree of all nodes and recall the routine until *k* nodes are removed.

2. Iterative greedy removal (GRD): In this approach, as in DEG, nodes are removed one-by-one. However, instead of removing the node with largest degree, at each step, we remove the node that minimizes the size of the current largest cluster (or maximize the number of components for the non-communicative clusters problem). After each iteration, we update the graph and iterate until *k* nodes are removed.

3. Iterative Minimal Separator (IMS): For small networks of hundreds of nodes (like the ones considered in this paper), there exists an efficient algorithm that enumerates all minimal vertex separators [26]. A vertex separator of a graph is a set of nodes whose removal separates the graph into at least two connected components. A minimal separator is one that is not a proper subset of any other separator. In our IMS heuristic, we iteratively use the minimal separator enumeration algorithm proposed in [26]. In each run, we enumerate the minimal separators of the current largest component and choose the minimal separator that optimizes a given criterion (e.g., minimize the size of the largest cluster for country isolation, or maximize the number of clusters for non-communicative clusters). We remove the nodes in the chosen separator and iterate until *k* nodes are removed.

4. Iterative Vertex Bisection I (IVB-I): In this approach, we iteratively apply the bisection algorithm presented in [28]. In each run, the algorithm in [28] is applied to the largest cluster at hand to find a vertex separator that produces two clusters of roughly equal size. However, sometimes the algorithm returns more than two components when one of the clusters (after removing the vertex separator) is not connected. We remove the nodes in the vertex separator and repeat this procedure until *k* nodes are removed. The main motivation in using this algorithm (and the next one) is that a solution to the optimization in equation (2) will lead to connected components that are roughly of the same sizes and thus minimize the size of the largest cluster.

5. Iterative Vertex Bisection II (IVB-II): We designed a new approach to iteratively bisect the largest remaining cluster until all *k* nodes are removed. For each bisection attempt, we randomly choose pairs of nodes and iteratively grow pairs of non-overlapping trees. To grow trees, we maintain a first-in-first-out (FIFO) queue of nodes to be processed (that are already in the tree) for each tree. When processing a node, we choose a random edge that has never been used and that is incident with a new node that is in neither tree. The new node and the processed node are both added to the end of the queue. We take the pair of trees that best optimizes the objective and then compare them against the output of the DEG algorithm (limited to the number of nodes removed in the tree bisection), with the best result being chosen.

Analyzing the performance of the different algorithms is beyond the scope of this paper. However, we do note that no single algorithm empirically dominated the others, and that all five approaches were capable of producing high-quality partitions for some combination of graph and quality function. For the networks considered in this paper, we have observed that IVB-I and GRD provided the best results most often, however they did not do so uniformly. While not immediately apparent, DEG can be implemented in O(*n*+*m*) linear time, by using an array of dictionaries, and is reasonably effective overall.

For our empirical results, each data point is analyzed by all 5 algorithms but only the result that best optimizes the relevant security question is kept. Thus, each presented curve in section 5 is made up of answers from all 5 algorithms, representing our best available approximation of the true NP-Hard answer.

V. Bounding Results

Our experiments indicate that the robustness of the Internet as a whole is increasing over time with respect to connectivity disruptions by a group of colluding countries. However, these gains tended to concentrate in countries already robust to such attacks. The group of less robust, less connected countries had results that remained fairly constant over the time period of investigation. Despite the increase in overall robustness, a small group of densely connected colluding countries remained capable of doing significant damage, even as late as 2013.

As discussed previously, all of these results show bounds on the worst case damage that could be inflicted. Since this applies to all of our results, we do not repeat this when giving each result. However, we emphasize that any conclusions taken from the data must take this into account. A reduction in the upper bound over time may represent a real improvement in the robustness of the Internet, or it may represent a more accurate measure of a stable quantity that results in a reduction of the worst case upper bound. As we keep the measurement infrastructure constant from year to year when using the 'persistent monitor' data, we then can interpret changes to a measured bound as being more likely correlated to actual changes in the security posture.

A. *Cutting Pairwise Communications*

The data shows that it is feasible for groups of colluding countries to block communications between pairs of countries, although this is becoming more difficult to implement over time. For all years and both data sets (those from Table 1 and Table 2), the minimum pairwise connectivity was 1. This indicates that given the routes visible through the CAIDA dataset, there always exists some country that by itself can disconnect some other pair of countries. The maximum connectivity grew steadily from 53 in 2008 to 68 in 2013 indicating an increasing connectivity strength of the most well connected countries. This means that there was a pair of non-adjacent countries that required collusion between a minimum of 68 countries in order to disconnect them.

We evaluate the mean connectivity in Fig. 3. On average, it took 9.41 countries to disconnect a pair of countries in 2013 (using all the data). Using the persistent monitor data, we see a clear increase in the mean connectivity over time showing the increasing resilience of the Internet to such collusion attacks.

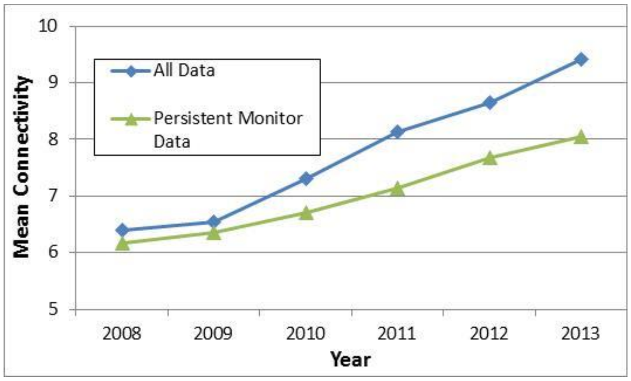

Fig. 3. Mean Number of Collaborating Countries Needed to Disconnect a Pair of Countries

B. *Country Isolation*

Fig. 4 shows the fraction of isolated nodes as a function of the number of colluding countries for the period 2008-2013.

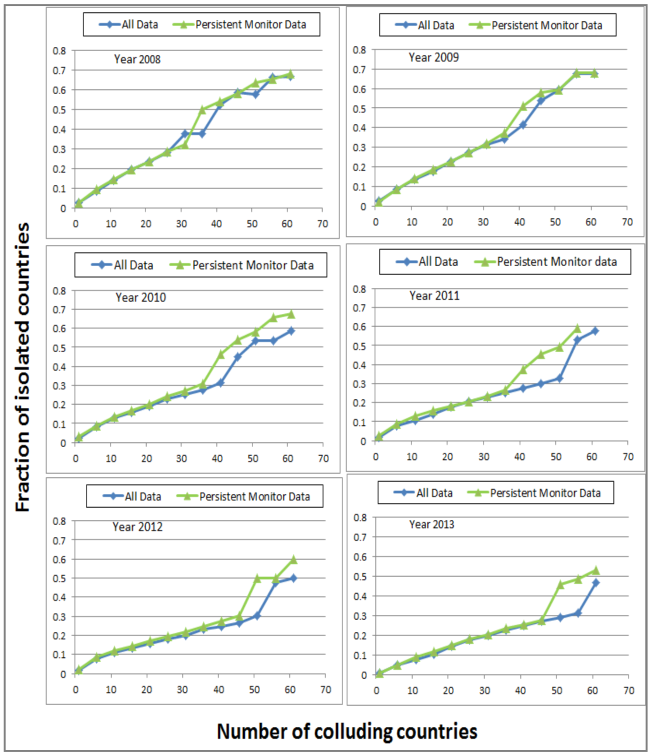

Fig. 4. Fraction of Isolated Nodes as a Function of the Number of Colluding Countries.

We have plotted the curves for both data sets (from Table 1 and Table 2). We observe a linear increase in the number of isolated countries as the number of colluding countries increases from 1 to around 60. This is the case for both data sets for all years. Notice that, as expected, the fraction of isolated nodes is slightly larger for the persistent monitor data due to less visibility of network links. The increasing resilience of the Internet can be seen by observing the slopes of the curves, which decrease over the years.

Another way to see this improvement in robustness is to ask: how many colluding countries does it take to isolate a fraction of x % of the Internet? Fig. 5 shows that over time, more countries need to collude in order to cut off the same fraction of countries from the Internet. However, these increases are for larger percentages of the countries. The 'cost' of isolating just 10 % of the countries in the Internet is small and relatively stable. This indicates that the Internet is still sensitive to colluding attacks, and suggests that a fringe set of countries are not receiving the benefits of increased robustness.

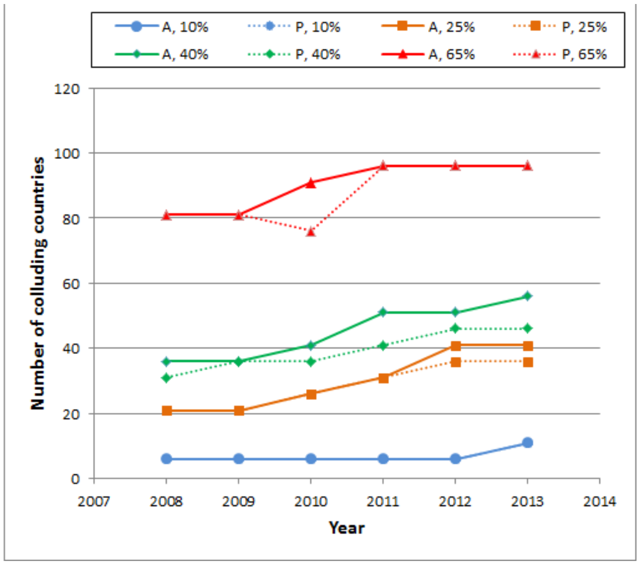

Fig. 5. Number of Colluding Countries Needed to Isolate a Fraction x % of the Nodes as a Function of the Year. "A, 10 %" in the Legend, means that the plot is done using the entire data set, and the fraction of isolated nodes is 10 %; "P" indicates that only the persistent monitor data is used.

C. *Non-Communicative Clusters*

Fig. 6 shows the number of non-communicative clusters as a function of the number of colluding countries. We observe a linear increase in the number of clusters as more countries collude. However, the rate of increase diminishes over time indicating that the Internet has become more resilient since 2008. The number of clusters that can be created by a particular sized group of countries has approximately halved from 2008 to 2013.

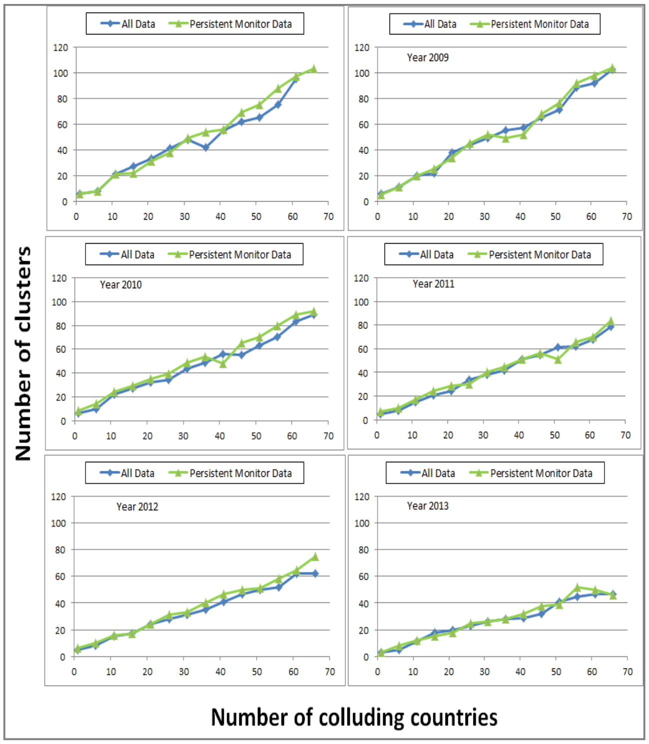

Fig. 6. Number of Clusters as a Function of the Number of Colluding Countries.

This is confirmed with the plots of Fig. 7 which shows the number of colluding countries needed to separate the network in differing numbers of connected components. With both data sets, the trend is that, over time, more countries are required to collude in order to separate the Internet into some predetermined number of non-communicative clusters for approximately 25 clusters or more. For smaller isolated clusters, the number of required countries remains relatively constant.

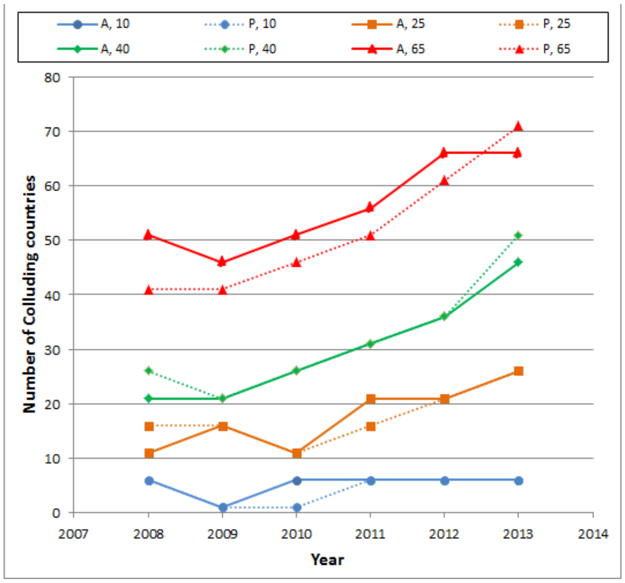

Fig. 7. Number of Colluding Countries Needed to Divide the Network into x Clusters as a Function of the Year. "A, 10" in the Legend, means that the plot is done using the entire data set and the number of clusters is 10; "P" indicates that only the persistent monitor data is used.

D. *Overall Analysis*

The Internet as a whole is becoming more resilient to colluding country induced connectivity disruptions. This can be seen as the lower bound on both the mean and maximal connectivity between countries increased from 2008 to 2013 (as did the percentage of 'invulnerable pairs'). Also, the number of countries needed to either isolate a significant fraction of countries or disconnect a significant number of clusters has generally increased substantially throughout the same time period.

However, these observed increases in resilience were only observed when isolating 25 % or more of the countries and 25 or more clusters from the core of the Internet. Some of the metrics saw almost no change. The minimum pairwise connectivity metric stayed at 1 for all years in the test set. This means that the lower bound on the number of countries required to isolate two countries never rose above 1. Also, for isolating 10 % of the countries from the Internet, the lower bound on the number of necessary colluding countries stayed constant for each year with a slight uptick in 2013. Likewise in the non-communicative cluster analysis, the lower bound on separating 10 clusters from the Internet required the same number of colluding countries every year (with a small dip in 2009 and 2010).

A point of commonality among these three findings is that they involve the smallest grouping of target countries in our experiment, where those countries were the least connected ones in the country-based AS graph. This suggests that, for these most vulnerable countries (that are most easily cut off from the Internet using the fewest colluding countries), their robustness has not noticeably increased over the time period of experimentation. This appears to be due to the fact that the majority of new edges observed in the country AS graph had at least one endpoint on a country that was already highly connected. This "rich get richer" phenomenon has been observed in other communications networks [5], and suggests that countries that cannot offer significant connectivity to potential partners will have difficulty obtaining sufficient communicating partners to ensure the robustness of their connectivity to the rest of the Internet.

Fig. 8 shows the average degree of the isolated nodes per year when some specific percentage of the countries in the Internet are being isolated. Note that, when just 10 % of the countries are isolated, the average degree of the isolated countries is less than 2. As the size of the set of isolated countries increases, the average degree of the isolated nodes also increases. In addition, while the degree of the isolated nodes for 10 % isolation appears relatively constant with respect to year, the degree increases sharply for later years when examining the isolation of 65 % of the nodes in the graph, supporting the hypotheses that the majority of new links are being formed to the most densely and robustly connected nodes.

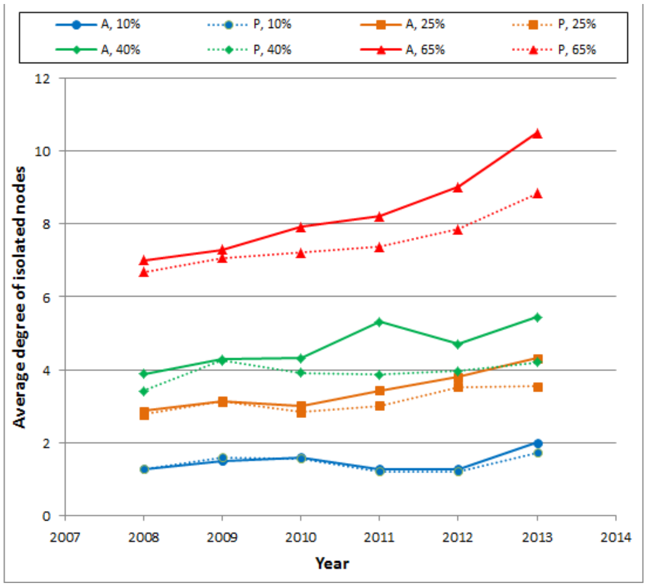

Fig. 8. Average Degree of Isolated Nodes per Year. "A, 10 %" in the Legend, means that the plot is done using the entire data set, and the fraction of isolated nodes is 10 %; "P" indicates that only the persistent monitor data is used.

Since this is an analysis of observed bounds, our results provide strong indications (given our consistent and static method of data collection and analysis). However, to further strengthen our argument, we can greatly bias the results so that the latter years should show a much greater resilience when attempting to isolate a small percentage of countries. We can do this by evaluating the 'persistent monitor' data from 2008 compared to the 'all data' data set from 2013. When we do this, we bias the year 2013 with 271 % more unique monitors than in the year 2008. This larger set of monitors covers 90 % more countries, including countries in under-represented portions of the globe with respect to monitor distribution (e.g., Bangladesh, The Gambia, Iceland, Indonesia, Mauritius, Mexico, Nepal, and Senegal). Given that these monitors are actively probing (as opposed to passively monitoring), we should then discover many more routes leading to results of greater resilience. However, our overall conclusions remain the same as can be seen from inspection of Fig. 5 and Fig. 7. It requires an increasingly large group of colluding countries to cut off large sets of nodes while the cost to cut off a small set remains relatively static. Countries with poor connectivity remain susceptible to connectivity attacks from groups of colluding countries even when intentionally biasing the data to promote the chance of obtaining the opposite result.

VI. Related Work

We are unaware of any other study with the same focus as ours: evaluation of the damage to Internet connectivity achievable by groups of colluding countries. However, there are related studies that cover individual countries cutting themselves off from the Internet, accidental and deliberate failures of groups of ASs, cascading failures, and the effect of routing policy on Internet robustness.

The work of [29] and [30] focuses on national actors disrupting their own country's connectivity to the Internet. In [29], historical incidents from 2011 in two countries were evaluated for the factors that made it comparatively easy for the countries to isolate themselves. This work was extended by [30], where they examined the number of internationally-facing ASs for each country as a measure of the robustness of that country to a similar self-isolation attack (i.e., how many domestic ASs would a government need to 'notify' in order to cut off country connectivity). They conclude that 133 countries could easily isolate themselves in such a fashion (having fewer than 10 internationally facing ASs).

The work of [31] provides a theoretical analysis of the robustness of the Internet to random failures, focusing on the giant component of power law graphs. They report that the Internet is extremely robust to random node removal, being able to retain a giant component even with random deletion of up to 99 % of the nodes. While not their primary focus, the work of [32] finds that the highly skewed degree distribution for individual nodes within ASs can result in the loss of a large number of links for removal of a single node. This supports the findings of [33] where they investigate the ability to maliciously partition the Internet through causing intentional AS failures. This work is most similar to ours, but focuses on the intentional failure of sets of ASs irrespective of their physical location or relationship to nation states. They also only examine the conventional graph partition problem, in which the focus is on generating a small number of partitions of roughly equal size or weight. They conclude that anywhere from 200 to 1500 nodes are required to find such a partition, depending on the desired properties.

While our work focused on complete disruptions to connectivity, the work of [13] evaluates how partial disruptions can shift traffic flow, causing cascading failures. They suggest that the existence of highly-connected nodes may exacerbate the problem of link overload.

Lastly, several studies evaluate the effect of routing policy (specifically the ‘valley-free’ restriction [6] on paths of directed links created by customer-provider relationships) on the Internet’s resilience to failures. In [8], a simulation-based model shows that policy-based routing can significantly exacerbate the impact of regional failures, in many cases leading to complete loss of logical connectivity, despite the presence of physical connectivity. Similarly, the work of [9] demonstrates that under policy-based routing, removal of 25 ASs from the graph is sufficient to reduce the size of the largest connected component to under half the size of the original AS graph they construct. However, these studies assume that routing policy will not change during times of severe network stress. In [7], they show that a relaxation of routing policy during emergencies will significantly increase Internet resilience (this is the approach taken by our work where we model all physical links as bidirectional). They examine several different failure models (depeering, link teardown, and regional failures) and they demonstrate that relaxation of policy-based routing allows for recovery of up to 80 % of communicating pairs in a network under Tier-1 depeering.

VII. Conclusion and Future Work

The apparent robustness of Internet networking hides underlying weakness. In this work, we have revealed a class of such weaknesses in the form of colluding countries deliberately filtering out other countries. We analyzed a group of countries disconnecting two other countries, isolating a set of countries from the Internet, and breaking the Internet into non-communicative clusters. We find that despite the potential for these attacks, the Internet as a whole has become increasingly resilient over the period of examination from 2008 to 2013. However, the gains in robustness and resilience were primarily concentrated in well-connected countries, which form an extremely resilient core. We found that the less connected countries formed new links largely with well-connected countries, and not to each other, maintaining the centrality of the well-connected countries in the paths between the less connected countries on the fringe. Because of this, the resilience of these less connected countries did not increase significantly over the time period studied. The result is that a small set of countries is able to isolate significant portions of the Internet or to divide it up into clusters. Individual well connected countries are often able to unilaterally isolate network dependent neighbors. These weaknesses could be addressed through a focus on establishing links between poorly connected countries. This would move the poorly connected countries away from dependence on the infrastructure of more highly connected countries.

Future work could evaluate a country or group of countries taking physical control of AS PoPs within their boundaries, ignoring ownership of the AS. A similar analysis could be done on the ability of such colluding countries to cause connectivity disruptions among target countries.

Acknowledgements

The authors would like to thank Dr. Richard La of the University of Maryland for his advice. This research was sponsored by the U.S. Army Research Laboratory (ARL) and the National Institute of Standards and Technology (NIST). It was partially accomplished under Army Contract Number W911QX-07-F-0023 and NIST Cooperative Agreement No.70NANB13H012 with the University of Maryland. The views and conclusions contained in this document are those of the authors, and should not be interpreted as representing the official policies, either expressed or implied, of the ARL, NIST, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes, notwithstanding any copyright notation hereon.

References

[1] W. Willinger and J. Doyle, "Robustness and the Internet: Design and evolution," 1 3 2002. [Online]. Available: http://netlab.caltech.edu/publications/JDoylepart1_vers42002.pdf. [Accessed 9 5 2014].

[2] A. Chang, "Undersea cables are actually more vulnerable than you might think," 3 4 2013. [Online]. Available: http://www.wired.co.uk/news/archive/2013-04/3/vulnerable-undersea-cables. [Accessed 9 5 2014].

[3] O. Malik, "Undersea cable cut near Egypt slows down Internet in Africa, Middle East, and South Asia," 27 3 2013. [Online]. Available: https://gigaom.com/2013/03/27/undersea-cable-cut-near-egypt-slows-down-internet-in-africa-middle-east-south-asia/. [Accessed 9 5 2014].

[4] T. Parfitt, "Georgian woman cuts off web access to whole of Armenia," 6 4 2011. [Online]. Available: http://www.theguardian.com/world/2011/apr/06/georgian-woman-cuts-web-access. [Accessed 9 5 2014].

[5] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science,* vol. 286, no. 5439, pp. 509-512, 1999.

[6] L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Transactions on Networking* , vol. 9, no. 6, pp. 733-745, 2001.

[7] C. Hu, K. Chen, Y. Chen and B. Liu, "Evaluating potential routing diversity for internet failure recovery," in *INFOCOM*, 2010.

[8] J. Wu, Y. Zhang, Z. M. Mao and K. G. Shin, "Internet routing resilience to failures: analysis and implications," in *ACM CoNEXT*, 2007.

[9] D. Dolev, S. Jamin, O. O. Mokryn and Y. Shavitt, "Internet resiliency to attacks and failures under BGP policy routing," *Computer Networks,* vol. 50, no. 16, pp. 3183-3196, 2006.

[10] "About CAIDA," [Online]. Available: http://www.caida.org/home/about/. [Accessed 9 5 2014].

[11] "University of Oregon Route Views Project," [Online]. Available: http://www.routeviews.org/. [Accessed 9 5 2014].

[12] J. Hawkinson, "Guidelines for creation, selection, and registration of an Autonomous System (AS)," 3 1996. [Online]. Available: http://tools.ietf.org/html/rfc1930. [Accessed 9 5 2014].

[13] C. Guo, L. Wang, F. Zhou, L. Huang and Z. Peng, "Analysis on the “robust yet fragile” nature of Internet: load, capacity and the cascading failure avalanche effect," in *The 9th International Conference for Young Computer Scientists*, 2008.

[14] "NetworkX," [Online]. Available: http://networkx.github.io/.

[15] International Standards Organization, "Online Browsing Platform," [Online]. Available: https://www.iso.org/obp/ui/. [Accessed 10 7 2014].

[16] M. E. Newman, "Assortative mixing in networks," *Physical review letters* , vol. 20, 2002.

[17] U.S. Department of State, "Overview of U.S. Export Control System," [Online]. Available: http://www.state.gov/strategictrade/ overview/. [Accessed 13 5 2014].

[18] C. Bagley, "Managers and the Legal Environment: Strategies for the 21st Century," Cengage Learning, 2012, pp. 832-833.

[19] J. O'Toole, "Good Business," New York, NY, Routledge, 2010, pp. 159-162.

[20] C. McPillips, "International Legal Advisory," Kaufman & Canoles, 2003. [Online]. Available: http://www.kaufmanandcanoles.com/documents/misc/international1.pdf. [Accessed 13 5 2014].

[21] A.-H. Esfahanian, "Connectivity Algorithms," [Online]. Available: http://www.cse.msu.edu/~cse835/Papers/Graph_connectivity_revised.pdf. [Accessed 12 5 2014].

[22] T. Bui and C. Jones, "Finding good approximate vertex and edge partitions is np-hard," *Information Processing Letters,* vol. 42, no. 3, pp. 153-159, 1992.

[23] J. Fukuyama, "Np-completeness of the planar separator problems," *Journal of Graph Algorithms and Applications,* vol. 4, pp. 317-328, 2006.

[24] D. Marx, "Parameterized graph separation problems," *Theoretical Computer Science* , vol. 351, no. 3, pp. 394-406, 2006.

[25] A. n Buluç, H. Meyerhenke, I. Safro, P. Sanders and C. Schulz, "Recent advances in graph partitioning," 2013.

[26] T. K. a. D. Kratsch, "Listing All Minimal Separators of a Graph," *SIAM Journal on Computing,* vol. 27, pp. 605--613, 1998.

[27] E. K. A. D. B. H. S. Cohen Reuven, "Breakdown of the Internet under Intentional Attack," *Physical Review Letters,* vol. 86, no. 16, pp. 3682--3685, 2001.

[28] U. a. B. J.-K. Hao, "Breakout Local Search for the Vertex Separator Problem," in *23th Intl. Joint Conference on Artificial Intelligence (IJCAI-13)*, Beijing, 2013.

[29] A. Dainotti, C. Squarcella, E. Aben, K. C. Claffy, M. Chiesa, M. Russo and A. Pescapé, "Analysis of country-wide internet outages caused by censorship," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement*, 2011.

[30] J. Cowie, "Could It Happen In Your Country?," 30 11 2012. [Online]. Available: http://www.renesys.com/2012/11/could-it-happen-in-your-countr/. [Accessed 10 6 2014].

[31] R. Cohen, K. Erez, D. Ben-Avraham and S. Havlin, "Resilience of the Internet to random breakdowns," *Physical review letters,* vol. 85, no. 21, 2000.

[32] L. Subramanian, V. N. Padmanabhan and R. H. Katz, "Geographic Properties of Internet Routing," in *USENIX Annual Technical Conference*, 2002.

[33] M. Wachs, C. Grothoff and R. Thurimella., "Partitioning the Internet," in *7th International Conference In Risk and Security of Internet and Systems (CRiSIS)*, 2012.