Internet Routing Resilience to Failures: Analysis and Implications

Jian Wu Ying Zhang Z. Morley Mao Kang G. Shin
University of Michigan

ABSTRACT

Internet interdomain routing is policy-driven, and thus physical connectivity does not imply reachability. On average, routing on today's Internet works quite well, ensuring reachability for most networks and achieving reasonable performance across most paths. However, there is a serious lack of understanding of Internet routing resilience to significant but realistic failures such as those caused by the 911 event, the 2003 Northeast blackout, and the recent Taiwan earthquake in December 2006. In this paper, we systematically analyze how the current Internet routing system reacts to various types of failures by developing a realistic failure model, and then pinpoint reliability bottlenecks of the Internet. For validity of our simulation results, we generate topology graphs by addressing concerns over the incompleteness of topology and the inaccuracy of inferred AS relationships. By focusing on the impact of structural and policy properties, our analysis provides guidelines for future Internet design. The simulation tool we provide for analyzing routing resilience is also efficient to scale to Internet-size topologies.

1. INTRODUCTION

Given our growing dependence on the Internet for important and time-critical applications such as financial transactions and business operations, there is a strong need for high availability and good performance at all times for most network paths on the Internet. To provide such assurance, Internet routing plays a critical role, as its main function is to identify network paths with sufficient resources between any two network prefixes. However, it is well-known that interdomain routing on today's Internet is *policy-driven* to satisfy commercial agreements. Policy restrictions prevent the routing system from full exploitation of the underlying topology, as physical connectivity does not imply reachability. It is unknown how such restrictions affect the failure-resilience of the Internet routing system.

On average, routing on today's Internet works well, ensuring reachability for most networks and achieving reasonable performance over most paths. However, there is a serious lack of understanding of Internet routing resilience to significant but realistic failures such as those caused by the 911 event [1] and the Taiwan Earthquake in December 2006 [2]. For instance, for several ten minutes to hours after this earthquake many Asian sites of U.S. companies cannot communicate with their headquarters or data centers in North America, preventing important business operations. A particularly noteworthy observation is that due to the North-America-centric placement of most top-level DNS domain servers for *.COM* domain, some Asian Web users cannot reach even regional servers due to the inability to contact authoritative DNS servers.

In this paper, we systematically analyze how the current Internet routing system reacts to various types of failures by establishing a realistic failure model, and then pinpoint reliability bottlenecks of the Internet. To achieve this, we first construct a topology graph which accurately captures the AS-level structure of today's Internet. Techniques are designed to address issues of topology completeness and relationship accuracy. Then we develop a generic failure model that captures the effect (not the cause) of most common failures affecting routing at the interdomain level. Note that such failures can also result from attacks instead of natural disaster. We attempt to identify critical links whose failures can cause large and severe impact on the Internet. They are effectively Achilles' heels of the Internet.

We develop a simulation tool to perform such what-if failure analysis to study routing resilience which is efficient to scale to Internet-size topologies. We focus on fundamental structural and policy properties that influence network resilience to failures. We attempt to draw conclusions independent of inaccuracies in relationship inference and topology construction by focusing on the underlying properties of networks that affect network-resilience properties. For example, there are a limited number of trans-oceanic links, which can easily become reliability bottlenecks. By focus-

*CoNEXT'*07, December 10-13, 2007, New York, NY, U.S.A.

ing on the impact of structural and policy properties, our analysis provides guidelines for future Internet design.

We summarize our main results of analyzing routing resilience to failures. (i) Tier-1 depeering, despite its infrequent occurrence, disrupts most of the reachability, *i.e.,* 94%, between the single-homed customer ASes of the affected Tier-1 ASes. (ii) Most of the reachability damage in today's Internet is caused by failures of the *critical access links*, which are traversed by all possible paths from the affected AS(es) to the rest of the Internet. We found out that 32% of the ASes are vulnerable to this type of the failure, most of which we believe is due to the nature of single-homing. Today's Internet might not be as resilient as we thought. (iii) BGP policy limits the ASes' option in selecting paths to reach other ASes, an additional 255 (6%) non-stub ASes can be disrupted by a single link failure even though the physical connectivity might be available to bypass the failure. (iv) Traffic is not evenly re-distributed during the failure and results indicate that more than 80% of the traffic over the failed link can be shifted to another link. (v) Adding extra links into the graph and perturbing relationship on certain links slightly improves the resilience of the network. The fundamental conclusion drawn above, nevertheless, stays the same.

Given our simulation-based failure analysis, we make the following observations to help enhance routing resilience: (i) We need extra resources (e.g., multi-homing) to be deployed around the weak points of the network. Approaches like sharing resources among neighboring ASes [3] can also be used. (ii) Based on the observation that policy further restricts path selection, other techniques to better utilize physical resources can also improve the resilience during failures, *e.g.,* selectively relaxing BGP policy restrictions. (iii) From our earthquake study, we learn that for some cases, even though reachability might not be affected, the performance will be severely degraded. (iv) Regional failures such as 911 has more global impact due to long-haul links connecting to remote regions.

To our best knowledge, this is the first detailed study of the impact of significant but realistic failures on the Internet, using both reachability and increase in traffic paths along links which reflect the impact on application performance. Our study reveals the vulnerability of the Internet routing through detailed data analysis of existing well-known failure events to provide insights into the derivation of solutions. The critical links identified by our simulation analysis tool can benefit the design of both short-term mitigation responses as well as other long-term improvements.

The paper is organized as follows. Section 2 introduces our overall methodology. The failure models used in our study as well as the corresponding empirical events are described in Section 3. The detailed resilience analysis under different types of failures using our simulation tool are discussed in Section 4. Finally, we discuss the related work and conclude the paper.

2. ANALYSIS METHODOLOGY

We describe our methodology for failure resilience analysis. It consists of three main components: (i) building the AS-level topology, (ii) inferring AS routing policies, and (iii) conducting failure analysis. Unlike previous studies, we carefully perturb relevant parameters.

2.1 Topology construction

We use publicly available BGP data from a large number of vantage points in the form of routing table snapshots as well as routing updates to construct an *AS-level network topology*. Combining routing updates with tables improves the completeness of the topology by including potential backup paths revealed only during transient routing convergence. However, history data may also introduce inaccuracies in the network topology caused by AS links that are no longer valid. We would like to obtain a topology graph that is as complete as possible to avoid underestimating routing resilience of today's Internet. By including network paths obtained from history data that may no longer exist, we may nevertheless overestimate its failure resilience.

To balance between the completeness and accuracy of network topology, we use 2 months of routing data from RouteViews [4], RIPE [5], public route servers [6] as well as a large content distribution network from March to April, 2007. The measurement data were collected from vantage points located in a total of 483 different ASes. To reduce the size of the network graph and speed up our analysis, we prune the graph by eliminating *stub* AS nodes [7], defined to be customer ASes that do not provide transit service to any other AS. These can be easily identified from routing data as ASes that appear only as the last-hop ASes but never as intermediate ASes in the AS paths. As a result, we could eliminate 63% of the links and 83% of the nodes. For the analysis of routing resilience to failures, we can restore such information by tracking at each AS node in the remaining graph the number of stub customer nodes it connects to including information regarding whether they are single-homed or multi-homed to other ISPs.

2.2 Topology completeness: missing AS links

The BGP data collected from a limited number of vantage points, such as RouteViews and RIPE, cannot locate all of the links in today's Internet [8, 9]. Certain links, especially peer-to-peer links in the edge of the Internet, only appear in the BGP paths between their associated ASes, therefore cannot be captured unless we place vantage points in these ASes. In our analysis, we address the incompleteness of topology by adding additional AS links which have been confirmed by other studies. In particular, we choose the data set provided by the latest link discovery study by He *et al.* [9] at UC Riverside, which we call graph *UCR*, and add their newly-found links missing in our topology data.

According to He's study [9], graph UCR is generated based on the data set collected in May 2005. Despite the time

Graph	# of nodes	# of links	# of peer-peer links	# of cust.-prov. links	# of sibling links
CAIDA	4342	14815	3558 (24.0%)	11168 (75.4%)	89 (0.1%)
SARK	4430	25485	3801 (14.9%)	21684 (85.1%)	0 (0.0%)
Gao	4427	26070	11446 (43.9%)	14343 (55.0%)	281 (1.1%)
UCR	3794	23913	14293 (59.8%)	9421 (39.4%)	199 (0.1%)

Table 1: Statistics of topologies generated by different algorithms

difference, we believe most of the links in the old data set still exist today. Table 1 presents the basic statistics of graph UCR and 3 other different graphs we generate based on different relationship algorithms, described in Section 2.3. Graph CAIDA is directly downloaded from [10] due to the lack of access to the source code of the study [11], Graph SARK and graph Gao are computed based on [7] and [12], respectively, from our collected raw dataset [1]. We discuss these graphs in details in Section 2.3. In comparison, graph UCR, slightly smaller than graph SARK and Gao due to its older raw dataset, nevertheless has a higher percentage of peer-peer links most of which were discovered by their proposed techniques. A further comparison of graph UCR with graph Gao shows that 10876 of the 23913 (45.5%) links in the former are missing in the latter. 10847 (99.7%) of these missing links are associated with existing nodes in the latter, indicating that they might be captured if other graph construction techniques (e.g., traceroute in [9]) are used. In Section 4, we evaluate how the addition of these missing links affects the overall resilience of the Internet.

2.3 AS routing policy inference

It is well-known that there are three basic AS relationships [13]: customer-to-provider, peer-to-peer, and sibling relationships. We need to label each link in the topology graph with relationship information required to infer valid, policy-compliant AS paths [14]. Thus, accurate AS relationships are critical to our analysis. Most previous studies on inferring AS relationships [7, 11, 12, 13, 15] are based on heuristics which might not always hold on the real Internet, and therefore, may produce incorrect relationships that directly affect our analysis. For example, a simple test on graphs annotated with AS relationships generated from CAIDA's work [11] reveals the presence of AS routing policy loops.

Although constructing a topology graph matching exactly the current Internet is impossible due to proprietary relationship information, we attempt to create one with maximum accuracy and understand the effect of network topology on routing resilience. A recent study [16] shows that the latest Gao's algorithm [12, 13] and CAIDA algorithm [11] present better accuracy in satisfying "valley-free" [13] policy rule for more AS paths. As such, we first generate a graph using Gao's algorithm with a set of 9 well-known Tier-1 ASes (AS 174, 209, 701, 1239, 2914, 3356, 3549, 3561, 7018) as its initial input. Then we compare the computed graph with

[1]Heuristics adopted by the different algorithms do not have definitive relationship inference for certain links in the graph, which results in the little discrepancy between SARK and Gao in Table 1.

Property	Value
# of AS nodes	4427
# of Tier-1 AS nodes	22 (0.5%)
# of Tier-2 AS nodes	2307 (52.1%)
# of Tier-3 AS nodes	1839 (41.5%)
# of Tier-4 AS nodes	254 (5.7%)
# of Tier-5 AS nodes	5 (0.1%)
# of AS links	26070
# of customer-provider links	14343 (55.0%)
# of peer-peer links	11446 (43.9%)
# of sibling links	281 (1.1%)

Table 2: Basic statistics of constructed topology

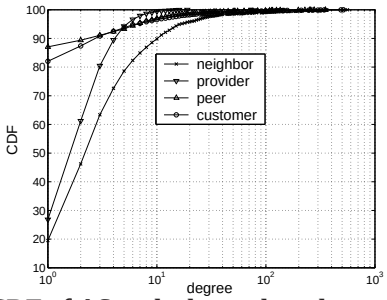

Figure 1: CDF of AS node degree based on relationships

graph CAIDA downloaded from [10]. We take the set of AS relationships agreed on by both graphs, which we believe are most likely correct, as the new initial input to re-run Gao's algorithm to produce the graph for our analysis. To ensure valid analysis of the constructed graph, we perform several consistency checks as described below.

- **Connectivity check**: The original topology graph needs to ensure that all AS node pairs have a valid policy path.
- **Tier-1 ISP validity check**: A Tier-1 ISP by definition does not have any providers, nor should their siblings. A Tier-1 ISP's sibling cannot be sibling of another Tier-1 ISP.
- **Path policy consistency check**: There should not be any valid AS path containing policy loops, *e.g.*, a path going from a customer to its provider and eventually returning to the customer serving as the previous hop's *provider*.

Table 2 describes the basic statistics of our constructed topology. We classify the nodes into 5 tiers as follows. We start with the 9 well-known ISPs and classify them and their siblings as Tier-1. Tier-1's immediate customers are then classified as Tier-2. We also ensure all non-Tier-1 providers of these nodes are included in Tier-2. We repeat the same process with the subsequent tiers until all of the nodes are

Previous link	Current link	Next link
↗	↗	↗, ⟷, ↘
↗	⟷	↘
↗, ⟷, ↘	↘	↘

Table 3: Relationship combinations of 3 consecutive links (↗: customer-to-provider link, ⟷: peer-to-peer link, ↘: provider-to-customer link)

categorized. As we can see, most of the nodes, after the removal of stub AS nodes, are in Tier-2 or Tier-3. Figure 1 also illustrates the node degree distribution of the graph. As expected, most networks have only a few providers. About 20% of the networks have at least one peer, which are typically equal-sized networks.

In reality, AS relationships can be much more complicated including per-prefix-based arrangements or combined relationships of transit or provider with customer services [17]. We argue that our simplified approach to constructing the AS-level topology with policy annotations is sufficient for failure analysis, as majority of the prefixes between AS pairs follow one type of policy arrangement. However, we do take care of special exceptions. For example, both Cogent (AS174) and Sprint (AS1239) are well recognized as Tier-1 ISPs, but they do not peer directly as evidenced by lack of AS paths containing links connecting them directly. In reality, Verio (AS2914) provides a transit between their customers. We deal with this case explicitly when computing AS paths.

2.4 AS relationship perturbation

As described earlier, no relationship inference algorithm is able to produce a set of AS relationships that exactly matches the actual ones. As a matter of fact, different algorithms could produce vastly different relationship inferences. As shown in Table 1, graph SARK has much fewer peer-peer links than graph Gao even though both graphs are computed from the same raw BGP dataset. To justify our evaluation of the Internet resilience, which relies on an accurate AS relationship, we propose a technique to perturb the relationship of certain links to understand the effect of AS relationship distributions on routing resilience.

Each link can be a "peer-peer", "customer-provider" or "provider-customer" link. Here we do not consider perturbation on a sibling link because of its rarity. As such, we have for each link 9 possible combinations of relationship tweaks, based on its relationship before and after the change. First, we discuss how each tweak affects the resilience. Table 3 presents all possible combinations of any three consecutive links in a policy-complaint AS path from the perspective of the second link (i.e., the link in the middle). Obviously, a peer-peer link is most restricted in finding paths as its previous link has to be a customer-provider link and its next link has to be a provider-customer link. In contrast, a customer-provider or provider-customer link has more options. Changing a peer-peer relationship to a customer-provider or provider-customer relationship thus provides the corresponding link more flexibility in choosing paths, and the overall network resilience is enhanced.

	p-p in SARK	p-c in SARK	c-p in SARK
p-p in Gao	2061	4847	3742
p-c in Gao	1011	9061	359
c-p in Gao	582	296	2723

Table 4: Relationship comparison (Gao, SARK)

In each of our relationship perturbation, we change the relationship of a number of links. To prevent the tweak of one link from offsetting the tweak of another link, we have to ensure that the tweaks of all of the links are consistent. That is, all of the links involved changing relationships from peer-peer to customer-provider/provider-customer or vice versa. In our current analysis, we only focus on relationship changes between peer-peer and customer-provider/provider-customer. The perturbation between a customer-provider link and a provider-customer link is less realistic and we thus leave it as future work.

Table 4 illustrates the comparison between graph Gao and graph SARK. The discrepancies provide candidates for perturbation. Each field indicates the number of links that satisfy the relationship combination. For example, there are 2061 links identified as peer-peer in both graphs and 4847 links identified as peer-peer in graph Gao but as provider-customer in SARK. As shown, there are altogether 8589 peer-peer links in Gao which are customer-provider or provider-customer links in SARK. This set of links is our main focus for the relationship perturbation analysis in Section 4. Note that each relationship tweak can only be applied when it does not violate any valley-free rule – the change will not invalidate any AS paths containing the link.

2.5 What-if failure analysis

Given the inferred AS relationships, we developed an efficient algorithm to construct valid AS-level policy paths between arbitrary AS node pairs. We modify the state-of-the-art algorithm [14] to ensure that the common practice of preference ordering is enforced by preferring customer routes to peer routes and peer routes to provider routes [18]. Figure 2 presents the pseudo-code of the algorithm with running time complexity of $O(|V|^3)$. Links in the AS graph are classified as one of the following categories: customer-to-provider link (UP link), provider-to-customer link (DOWN link), and peer link (FLAT link). Accordingly, a path which only follows UP links is called an *uphill* path. Any AS path conforming to BGP policy is of the form of an optional uphill path, followed by zero or one FLAT link, and an optional downhill path. The algorithm starts with the computation of the shortest uphill/downhill paths for all node pairs. Then, it selects from all possible path combinations the shortest path with the preference ordering applied.

Our algorithm is efficient, as we impose an ordering to compute a given AS's provider's routes first both for eliminating unnecessary computation and ensuring consistent routes. Our simulator [19] supports a variety of what-if

1. Compute shortest *uphill* paths for all (src, dst) pairs.
   $Dist_{src,dst}$ is the distance of the shortest uphill path
   $Uphill_{src,dst}$ is the shortest uphill path
2. Compute the shortest policy path from src to dst
   function **shortest_path**(src, dst, $D_{src,dst}$, $P_{src,dst}$)
   # returns $D_{src,dst}$, the length of the shortest path,
   # and $P_{src,dst}$, the shortest path
   if $Dist_{dst,src} < \infty$ # choose customer's path
   $D_{src,dst} = Dist_{dst,src}$;
   $P_{src,dst} = Reverse(Uphill_{dst,src})$;
   else # choose peer's path
   $D_{src,dst} = \min_p\{Dist_{dst,p} + 1\}$;
   where p is a peer of src
   if $D_{src,dst} < \infty$
   $P_{src,dst} = (src, p) + Reverse(Uphill_{dst,p})$;
   else # choose provider's path
   foreach src's provider m
   shortest_path(m, dst, $D_{m,dst}$, $P_{m,dst}$);
   $D_{src,dst} = \min_m\{D_{m,dst} + 1\}$;
   $P_{src,dst} = (src, m) + P_{m,dst}$;

Figure 2: Algorithm to compute shortest policy paths for all src-dest pairs

analyses by deleting links, partitioning an AS node to simulate the various types of failures described in Section 3. The simulation tool is designed to be efficient in computing AS paths: all AS-node pairs' policy paths can be computed within 7 minutes with 100 MB memory requirement on a desktop PC with an Intel Pentium 3GHz processor.

3. FAILURE MODEL

Although the Internet has built-in failure recovery mechanisms through rerouting, there are several real incidents of serious connectivity problems during natural disasters, power outage, misconfigurations, and even intentional attacks [20] against the infrastructure. In Table 5, we introduce a failure model capturing the *effect* of network disruption at the global Internet level based on empirical evidence.

As shown in Table 5, we categorize the failure scenarios based on the *impact scale*, which we measure by the number of *logical* links affected by the failure. Here, a *logical* link is defined as the peering connection between an AS pair. A logical link might involve several physical links, *e.g.,* two large ISPs peer at multiple geographical locations. We do not explicitly model physical links due to a lack of physical topology information. Based on the number of impacted logical links, we classify failures into three types: no logical link failure, single logical link failure, and multiple logical link failures.

No logical link failure: For reliability and performance reasons, ASes might have more than one single physical link to connect to each other. In particular, if the peering is present at geographically diversified locations, it is be very difficult to completely break the connection between these two ASes. We usually observe the following two types of failures.

- Partial peering teardown: As reported in [21], session reset, due to hardware/software malfunction or maintenance operations, is one of the most frequent routing events in the network. Unless all peering sessions between an AS pair have reset, the two ASes can still maintain their reachability even though traffic performance might be degraded.

- AS partition: Certain physical link failures, occurring inside a single AS, do not cause any damage to its connection to its neighboring ASes. The most severe condition is that the failure breaks the AS into two or more isolated regions, and the networks in different regions can no longer reach each other. We call this type of failure "*AS partition*", as evidenced by a recent event in Sprint backbone [22].

Single logical link failure: A logical link failure indicates the loss of direct connection between the pair of ASes associated with the link. Based on the types of the failed link, we further categorize it into the following two sub-classes.

- Depeering: Depeering occurs when the failure disables the peer-peer link between a pair of ASes. In today's Internet, the largest ISPs (*i.e.,* Tier-1 ASes) establish peer-peer relationships to distribute traffic for their respective customer networks. To gain extra connectivity without increasing financial burden, low-tier ASes also peer with each other. Depeering over a Tier-1 peer-to-peer link can cause significant impact on the Internet as it disrupts the communication between their respective customers and is mostly intentional as evidenced by recent contractual disputes between Cogent and Level3 [23]. In contrast, in the case of lower tier depeering, which is possibly caused by physical damage, misconfiguration, or even intentional link termination, reachability can still be maintained through other provider links with possible performance degradation.

- Teardown of access links: Most networks connect to their providers through the access (*i.e.,*customer-provider) links to reach the rest of the Internet. A failure on such access links can severely disrupt the customer's reachability. This type of failure might be one of the most common link failures, as evidenced by the frequent reports in NANOG [24].

Multiple logical link failures: This type breaks multiple logical links, thus causing much more severe impact.

- AS failures: one particular scenario, we denote as "AS failure", occurs when all the logical links between an AS and its neighbors fail, indicating that the corresponding AS is unable to originate or forward any traffic. This can be caused by hardware malfunction or misconfiguration inside the failed AS. For instance, UUNet backbone problems [25], despite its undisclosed causes, resulted in significant network outages.

- Regional failures: are often caused by natural disaster and intentional attacks, resulting in multiple logical

Category: (# of logical links)	Sub-Category	Description	Empirical Evidence	Analysis
0	Partial peering teardown	A few but not all of the physical links between two ASes fail	eBGP session resets	
	AS partition	Internal failure breaks an AS into a few isolated parts	Problem in Sprint backbone	Section 4.6
1	Depeering	Discontinuation of a peer-to-peer relationship	Cogent and Level3 depeering	Section 4.2
	Teardown of access links	Failure disconnects the customer from its provider	NANOG reports	Section 4.3
> 1	AS failure	An AS disrupts connection with all of its neighboring ASes	UUNet backbone problem	
	Regional failure	Failure causes reachability problem for many ASes in a region	Taiwan earthquake, etc	Section 4.5

Table 5: Failure model capturing different types of logical link failures.

link or AS failures in the affected region. In addition to local networks in the region, other parts of the Internet whose traffic traverses the region are also impacted. Well-known examples include 911 attack [1], Hurrican Katrina [26], as well as the recent Taiwan earthquake [2].

As evidenced by various real events, the Internet is susceptible to certain types of failures, especially when critical nodes (UUNet problem) or links (Cogent and Level3 depeering) are involved. In Section 4, we use our simulation tool to conduct a more systematic evaluation of the impact of different types of failure on the Internet.

3.1 Case study: Taiwan earthquake

Given the known disruption to the Internet due to the recent Taiwan earthquake [2], we perform a more detailed study of its impacts in the region on the third day after the earthquake happened. The earthquake occurred in December 2006 near Taiwan, damaging several undersea cable systems in Asia. Many networks in Asia were affected, causing degraded performance, and network connectivity problems in Asia were globally felt for weeks.

We first collected BGP data for that period of time from RouteViews and RIPE which captures the earthquake effects based on the number of ASes or prefixes that experience path changes (or even complete withdrawals). In addition, given that the effect of the earthquake was relatively long-lasting due to the long repair time, we augment our analysis with traceroute probes. In particular, we probe from PlanetLab hosts [27] located in several Asian countries and other areas of interest: China, Singapore, Taiwan, Japan, South Korea, US, and Australia. The goal is to understand possibly abnormal paths with long delays and to locate the bottleneck causing the slowdown.

We summarize our findings. Most affected prefixes belong to networks in Asian countries around the earthquake region. For example, 78-83% of the 232 prefixes announced from a large China backbone network were affected across 35 vantage points. Most of the withdrawn prefixes were reannounced about 2 to 3 hours later. We found that many affected networks announced their prefixes through their backup providers. For example, before the event all the vantage points went through AS1239 to reach China 169 backbone (AS4837). After the earthquake, backup paths through networks such as AS3320, AS7018, and AS1239 are used.

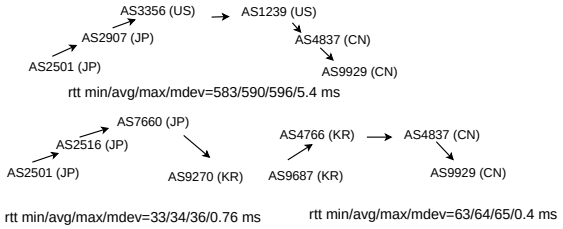

Figure 3: Path from Japan to China

	AU2	CN2	HK2	JP2	KR2	SG2	TW2	US2
AU	11	657	433	271	335	392	304	229
CN	570	150	41	446	318	83	286	475
HK	288	219	2	127	137	40	446	337
JP	152	450	117	21	44	94	137	169
KR	287	203	655	40	5	468	378	172
SG	391	412	37	208	355	90	360	267
TW	270	559	456	32	280	471	1	182
US	242	205	251	190	194	296	188	8

Table 6: Latency matrix among Asian countries in msec (from educational to commercial networks)

We identified several AS-level links experiencing problems. For example, before the event, all the vantage points traversed AS1239 to reach a Singapore network, AS4657. After the earthquake, they instead choose other ASes such as AS209 and AS2914.

By actively performing traceroute probing from 8 PlanetLab nodes in 8 distinct Asian countries, we found that interestingly, traffic between some network prefixes in Asia are routed via other remote continents during the period after the earthquake. For example, The Taiwan Academic Network to China Netcom were routed from Taiwan to NYC before reaching China Netcom. The roundtrip delays can exceed 550ms due to the long distance and congestion. During normal period though the AS level path is the same, packets are routed within the east pacific area. As shown in Figure 3, we found that from the PlanetLab node in Japan to a China commercial network, the path goes through the US, taking a long time to travel over excessive physical distances. However, two networks in South Korea have direct connections to both Japan and China networks. Hence, if the networks in Korea can provide temporary transit services for both China and Japan, we obtain an overlay path through Korea with a much shorter physical distance.

To generalize our analysis, we obtained a latency matrix among Asian countries and the US from educational to com-

mercial networks shown in Tables 6. Based on this, we identify that at least 40% of paths with long delays can be significantly improved by traversing a third network. The best improvement reduces latencies from 655ms to only around 157ms (from KR to HK2 when asking JP to provide the transit service). More details of the study are presented in [19].

4. IMPACT ANALYSIS OF FAILURES

We now analyze each failure type to understand the impact at the Internet scale. Note that we focus on *logical* link failures only, which corresponds to failures of one or more physical links. Such failures are not unlikely as evidenced in the past. In what follows, unless otherwise specified, a link implicitly means a logical link, and a node refers to an AS.

4.1 Evaluation metrics

A failure disrupts the traffic that traverses the failed network component and the traffic has to be rerouted via a different path to reach its destination. To quantify failure impact, we define the following two metrics:

- *Reachability impact*: In the worst-case failure scenario, no alternative path can be located between the source and the destination. We define two types of reachability impact: the *absolute* reachability impact R^{abs} and the *relative* reachability impact R^{rlt}. R^{abs} is the number of AS pairs that lose reachability to each other during the failure. In addition, We define R^{rlt} as the percentage of disconnected AS pairs over the maximum number of AS pairs that could possibly lose reachability.

- *Traffic impact*: After the failure, the traffic that used to traverse the old failed link is shifted onto the new paths. The shifted traffic could lead to serious network congestion. Due to the lack of accurate information on actual traffic distribution among ASes, we instead estimate the amount of traffic over a certain link as the number of the shortest policy-compliant paths that traverse the link, denoted as *link degree* D. We compute the link degree D of all links before and after the failure, and estimate the effects of traffic shift by calculating these 3 metrics: (1) the *maximum* increase of D among all links T^{abs}, (2) the *relative* increase of D of this link T^{rlt}, and (3) the maximum relative increase in D of the failed link T^{pct}. Suppose the link A is failed, and most of its traffic is shifted to link B. The three metrics are computed as follows.

$$T^{abs} = D_B^{new} - D_B^{old}, T^{rlt} = \frac{T^{abs}}{D_B^{old}}, T^{pct} = \frac{T^{abs}}{D_A^{old}} \quad (1)$$

The first two quantify the impact of traffic shift on individual links while T^{pct} captures the evenness of re-distributed traffic for the failed link. Although the link degree cannot exactly quantify the traffic impact in each failure because of the uneven traffic distribution in the Internet, it, which computes the increased number of AS paths that traverse each link, provides a good estimate on the amount of shifted traffic.

AS	174	209	701	1239	2914	3356	3549	3561
174	/	/	/	/	/	/	/	/
209	100	/	/	/	/	/	/	/
701	87	91	/	/	/	/	/	/
1239	79	91	85	/	/	/	/	/
2914	100	93	100	85	/	/	/	/
3356	100	95	100	85	100	/	/	/
3549	82	99	82	85	100	87	/	/
3561	87	92	100	89	100	100	100	/
7018	92	100	92	100	92	92	92	100

Table 8: R^{rlt} (%) for each Tier-1 depeering

4.2 Depeering

Today's Internet core consists of a group of large ISPs known as Tier-1 ASes which are the top service providers. Their customers can reach each other via the peer-peer links among the Tier-1 ASes, so these peering links are critical to maintaining the Internet connectivity. In this section, we analyze the effects of peering (particularly the Tier-1 peering) link failures on network reachability and traffic shift.

Table 7 presents the number of single-homed customers with and without the stub ASes for each Tier-1 AS, where *single-homed* refers to customers that can only reach only one Tier-1 AS through uphill paths. If all the physical peering links between two Tier-1 ASes stop working, *i.e.,* a logical link failure, their respective single-homed customers can only reach each other using the lower-tier peering links.

We first analyze how each Tier-1 depeering affects loss of network reachability due to unreachable AS pairs of single-homed ASes of the Tier-1 ASes involved. Because of the rich connectivitity in the Internet, some pairs of the single-homed ASes of the depeered Tier-1 can still reach each other via low-tier peering links. We use the relative reachability impact $R_{i,j}^{rlt}$ to quantify the impact,

$$R_{i,j}^{rlt} = \frac{\#\ of\ disconnected\ pairs}{1/2 \times S_i \times S_j}, \quad (2)$$

where S_i and S_j indicate the number of single-homed ASes for the two depeered Tier-1 ASes i and j. Table 8 presents the results for our graph without stub ASes. Tier-1 depeering disrupts connections among most single-homed customers. Overall, 89.2% of pairs of Tier-1 ISP's single-homed customers suffer from reachability loss, while the remaining pairs manage to detour using lower-tier peers or siblings. If we consider the stub ASes, 298493 (93.7%) out of 318562 single-homed AS pairs lose reachability.

We examine pairs of single-homed customers that remain connected after depeering. Among all 744 connected pairs, 86% of them traverse peer-peer links, and the remaining 14% have common low-tier providers.

Second, we investigate the effects of Tier-1 depeering on traffic shift. We observed, on average, the maximum traffic increase of a link, *i.e.,* T^{abs} is 3040 (with maximum of 11454), which corresponds to 22% (with maximum of 62%)

Tier-1 AS	174	209	701	1239	2914	3356	3549	3561	7018
# of single-homed customers without stubs	16	13	9	13	11	30	15	10	9
# of single-homed customers with stubs	193	229	45	47	43	162	53	55	49

Table 7: Number of single-homed customers for Tier-1 ASes

# of perturbed links	0	2k	4k	6k	8k
% of disconnected ASes	89.2	88.6	87.9	87.2	86.3

Table 9: Effects of perturbing relationship.

of the traffic of the depeered link (*i.e.*, T^{pct}) being shifted. Our results also show the relative traffic increase T^{rlt} could reach up to 237% with an average increase of 61%, indicating that the traffic shift might impose a serious burden on certain links.

We also analyze depeering of lower-tier peering links. Even though they do not impact network reachability due to the ability to use Tier-1s to reach each other, we examine the traffic impact. We pick 20 most utilized non-Tier-1 peer-to-peer links, and simulate the path changes after the failure of each link. Our results show that the average maximum traffic increase T^{abs} is 14810, and the corresponding T^{pct} and T^{rlt} are 35% and 379%, respectively, indicating that lower-tier peering links can also introduce significant traffic disruption.

4.2.1 Effects of missing links

As we discussed in Section 2.2, our topology graph, constructed solely from BGP measurement data, cannot capture all the links in the Internet. We add the newly-discovered links in graph UCR to examine how it affects the simulation results.

A total of 10847 links are added, containing 8059 (74.3%) peer-peer links, 2753 (25.4%) customer-provider links, and 35 (0.3%) sibling links. For comparison purposes, we use the same set of single-homed ASes in our analysis. 5892 (85.5%) pairs of ASes experiencing loss of reachability in the new graph, compared to 6143 (89.2%) pairs of ASes in the old graph. As expected, adding new links slightly improves the resilience under Tier-1 depeering as the new links can be used to locate alternative paths.

4.2.2 Effects of relationship perturbation

Next, we evaluate how perturbing the relationship described in Section 2.4 affects the analysis results. We have a candidate set of 8589 peer-peer links which can be changed to customer-provider links. In our evaluation, we test 4 different scenarios in which 2000, 4000, 6000, and 8000 peer-peer links in the candidate set are randomly selected and changed to customer-provider or provider-customer links. For each test scenario, we randomly generate 5 different graphs.

For comparison purposes, we consider the same set of single-homed ASes and evaluate how the perturbation affects the connectivity between any pair of these ASes. Table 9 presents the percentage of single-homed AS pairs that lose reachability under different scenarios. As shown in the table, perturbing the relationship slightly improves the resilience of the network as the perturbed provider-customer links either make single-homed ASes become multi-homed or provide better lower-tier connectivity. *The quite limited improvement also indicate that these single-homed customers have very limited access links to reach Tier-1 ASes and uninformed, random relationship perturbation does not improve their routing resilience much.*

To summarize, Tier-1 depeering disrupts the reachability of only a small number of ASes that are single-homed to the affected Tier-1 ASes, nevertheless, these affected ASes experience severe damage as they can no longer reach 89% of the rest of the ASes.

4.3 Teardown of access links

After the analysis of failures of peer-peer links, we now study how the failure of customer-provider links (also known as *access link*s), which counts for 77% of all AS links in the Internet, affects the network reachability. The robustness of connectivity of an AS can be captured by the similarity of its paths reaching the *Tier-1* ASes, given that Tier-1 ISPs are so richly connected; thus, reaching them is very important. For example, in the Tier-1 depeering analysis, ASes with uphill paths to multiple Tier-1 ASes can survive the depeering disruption without losing reachability to other ASes.

Path similarity can be defined as the number of commonly-shared links among all the paths under consideration. In particular, nonzero path similarity means that failing *a single link* can disrupt reachability. For instance, similarity of 2 implies that there exists two commonly shared links among all possible paths; therefore breaking any of the two links will create disruption.

We now describe how to calculate the path similarity of each AS to the set of all Tier-1 ASes to evaluate the robustness of the connectivity and to identify critical links. We first transform this problem into a max-flow-min-cut problem [28]. We solve the minimum-cut problem by using an approach based on the "push-relabel" method [28] and then present our analysis for scenarios with the BGP policy imposed and also those without policy restrictions. Moreover, we study the impact of failures of commonly-shared links which tend to be critical for the network.

Since our focus is on finding cases of nonzero path similarity, we transform the problem into a max-flow-min-cut problem by assigning a capacity of 1 for every link in the graph. The solution identifies the maximum flow that can be transferred between a source s and a sink t. Because each link has a capacity of 1, once we have a solution with a maximum flow value of 1, there has to be at least one link shared by all paths between s and t.

In our analysis, we have one source and multiple sinks.

```
function find_path(src, dst, last, link_set)
# if returns TRUE, paths exist between src and dst;
# link_set is the set of links shared by these paths
  if (src = dst)
    ret = TRUE; link_set = {(last, dst)}
  else
    S = {all links}; ret = FALSE; # initialize S and ret
    foreach x ∈ {src's providers or siblings}
      if (find_path(x, dst, src, S_x) = TRUE)
        S = S ∩ S_x; ret = TRUE;
    link_set = S ∪ {(last, src)};
  return ret;
```

Figure 4: Algorithm to locate shared links among all paths from src to dst.

The source can be any non-Tier-1 AS while the multiple sinks are the Tier-1 ASes. We create a supersink t and add a directed link from each Tier-1 AS to t with a capacity value of ∞. We perform the analysis for both conditions of BGP policy constrained path selection and no policy restrictions. For the latter, we transform our topology into an undirected graph. For the former, since we consider the uphill paths of each non-Tier-1 AS to Tier-1 ASes, which do not contain any peer-peer links, we remove all peer-to-peer links from the topology, while keeping each customer-to-provider link as a directed link pointing from the customer to the provider, and making each sibling link undirected. All links in the converted graph have capacity value of 1 except for the links to the supersink.

Under no policy restrictions, 703 (15.9%) out of 4418 non-Tier-1 ASes have a min-cut value of one and can thus be disconnected from the network by removing only one of the commonly-shared links. This implies that *despite apparent physical redundancy, a fairly large number of networks on the Internet are vulnerable to significant reachability disruption caused by a single access link failure even without policy restrictions.*

Under BGP policy restrictions, 958 (21.7%) of 4418 ASes have a min-cut value of 1, and about 255 (6%) of the ASes are susceptible to single link failures even though they have physical connectivity. This indicates *BGP policies severely limit network reachability under failures, and relaxing policies can help alleviate the failure impact.*

Recall that stub ASes excluded from our topology graph tend to have even more limited connectivity due to being single-homed. In our graph, we exclude 21226 stub ASes, 7363 (34.7%) of which have only one provider and are thus subject to a single access link failure. Considering the stub ASes, at least 8321 (32.4%) of the ASes are vulnerable to single access link failure.

The default s-t max-flow-min-cut solution only generates one possible cut. We develop a recursive algorithm for finding the set of all commonly-shared links among all possible paths between a given non-Tier-1 AS and the set of Tier-1 ASes, shown in Figure 4. By remembering partial results, the running time complexity of this algorithm is $O(|V| + |E|)$.

# of shared links	0	1	2	3	4
percentage	78.3	18.3	3.1	0.3	0.02

Table 10: Number of commonly-shared links.

# of nodes	1	2	3	4	5	>5
percentage	92.7	4.5	1.6	0.1	0.3	0.7

Table 11: Number of ASes sharing the same critical link.

Table 10 shows the percentage of the number of shared links from any non-Tier-1 AS to all Tier-1 ASes. Most of the ASes that share link(s) have only 1 common link while few nodes share as many as 4 links to reach Tier-1 ASes. This implies that *the attack of a randomly selected link is unlikely to significantly disable the targeted AS's connectivity from other networks.* We also collect the statistics on the links that are commonly shared by any of these ASes.

Table 11 presents statistics on the number of AS nodes that share the same critical link. Removing each of these links disrupts the connectivity of all of the ASes that share the link. More than 90% of the links are shared by only one AS while few links are shared by more than 10 ASes to reach the set of Tier-1 ASes. This indicates *a single logical failure has a limited scale of impact as ASes rarely share a common critical access link.*

To capture the impact of removing shared links, we study failure scenarios in which any of the 20 most shared links is disabled. We estimate the impact by using previously defined metrics. Upon failure, the affected AS(es) can no longer reach the Tier-1 ASes and their reachability to other networks solely relies on their alternate lower-tier connectivities. We use the relative reachability impact R_l^{rlt} for failed link l as our metric,

$$R_l^{rlt} = \frac{\#\ of\ disconnected\ pairs}{1/2 \times S_l \times (S - S_l)}, \quad (3)$$

in which S_l and S indicate the number of ASes that share the failed link l and the total number of ASes in the graph, respectively. For the 20 scenarios analyzed, our results show that the average value of R^{rlt} is 73.0% with standard deviation of 17.1%. *Failures of shared access links disrupt most of the reachability for ASes that share the removed links.* In the few cases when reachability is not impacted, the corresponding pairs of ASes use low tier links similar to depeering to route around the failed link.

For the traffic impact, the maximum increase T^{abs} among the 20 failures is 53179, accounting for 50.3% of the total traffic shift, *i.e.*, T^{pct}.

4.3.1 Effects of missing links

Similar to the depeering analysis, we evaluate how the addition of new links learned from graph UCR affects our conclusions. With added links, our results show that 678 (15.3%) of the ASes have min-cut value of 1 under no policy restriction showing an increase of 25 (0.6%) ASes no longer sharing common links. Under policy restrictions, however, 956 (21.6%) of the ASes have min-cut value of 1, *i.e.*,, only 2 (0.05%) additional ASes becomes insusceptible to single link failures with additional links. For the failures of the

# of perturbed links	0	2k	4k	6k	8k
# of ASes with min-cut 1	958	928.6	901.3	873.5	848.9

Table 12: Perturbing relationships: improved resilience.

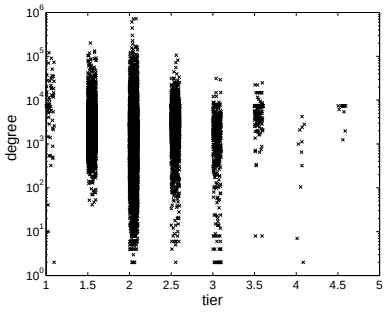

Figure 5: Link degree vs. link tier.

same 20 shared links, the average of R^{rlt} is 68.7% with standard deviation of 14.3%.

We can conclude that *although the addition of new links increases the physical connectivity of networks, it only slightly improves the resilience for access link failures as the added links, most of which are peer-peer links, have a limited access to reach the affected ASes.*

4.3.2 Effects of relationship perturbation

We next discuss how relationship perturbation affects the min-cut analysis results. Similarly, we simulate failures on 4 different graphs in which we change 2000, 4000, 6000, and 8000 peer-peer links in the candidate set to customer-provider/provider-customer links. We randomly generate 5 tests for each scenario. We focus on min-cut analysis under BGP policy restrictions.

Table 12 presents the min-cut analysis results for 4 relationship perturbation scenarios. *Changing peer-peer links to customer-provider/provider-customer links improves the overall network resilience as the perturbed links provides ASes extra flexibility in choosing paths to other networks.*

To summarize, despite the apparent physical redundancy, a surprisingly large number of ASes are vulnerable to a single access link failure, which we believe is the most common failure in today's Internet. Even worse, BGP policies severely further limit the network resilience under failure: about 35% of the ASes can be disconnected from most of the rest of the network by a single link failure.

4.4 Failure of heavily-used links

Shared links to reach Tier-1 ASes can be considered as one type of *critical* links. We also analyze the impact of failures of another type – links used by many networks or heavily-utilized links based on their topological location.

Figure 5 is a scatter plot of the link degree vs. link tier. *Link tier* is calculated as the average of tier values of the two ASes of the link. For example, if the link is between a Tier-1 AS and a Tier-2 AS, the link tier is 1.5. *Link degree* D, as defined in Section 4.1, is the number of AS pairs traversing the link. As shown in the figure, the most heavily-used links are within Tier-2. This is expected as core links carrying significant amount of Internet traffic have high link degrees.

In our simulation, we select 20 most heavily utilized links as failure targets, excluding Tier-1 peer-to-peer links which have been studied in Section 4.2. These 20 links either reside in Tier 2 or connect between Tier-1 and Tier-2 ASes and are traversed by 0.9% up to 5.2% of paths between all AS pairs. In each simulation run, we remove one of these 20 links and estimate the failure impact. In particular, we examine how those AS pairs that used to traverse the broken link fail over to new paths. Our analysis shows that 18 out of 20 failures do not disrupt reachability between any AS pairs. In fact, the two cases that impact reachability involve two shared links as evaluated in Section 4.3.

For the 20 failures studied, the maximum T^{abs} is 113,277 with an the average of T^{abs} 64,234 while the maximum T^{pct} is 77.3% with the average of T^{pct} 38.0%. These values indicate significant, uneven traffic re-distribution that may require traffic engineering to reduce potential congestion.

4.5 Regional failures

We now present simulation-based analysis of a particular regional failure scenario. We first describe the method to determine the set of affected ASes and links before presenting the analysis on the failure impact.

Motivated by several real incidents such as the 9/11 attack and the 2003 Northeast blackout, our regional failure simulates the scenario when all ASes and links traversing New York City (NYC) are broken. Unlike the previous scenarios that focus on single link failures, regional failures usually affect multiple links and tend to have larger impact.

We first use NetGeo [29] to approximately identify the set of ASes and links that can be affected by events in NYC. NetGeo provides a set of geographic locations for each AS. Because our analysis is based on the AS-level granularity, we select ASes located in NYC only and thus ignore partial AS failure for simplicity. To identify relevant links, we first choose links whose both end points share a single common location in NYC. In addition, NYC might also be critical to links with a single end point in NYC. For example, we observe that South African ISPs connect to New York as their main exchange point to the rest of the Internet even though NetGeo indicates they only reside in South Africa.

To capture such long-haul links connecting NYC to a remote region, we perform traceroute from PlanetLab hosts located different foreign countries to 35 PlanetLab ASes located near NYC. If traceroute results exhibit any stops in NYC, we include the corresponding AS links. Due to limited probing, our analysis may miss some links impacted by the failure. A total of 268 ASes and 106 links (56 of them are customer-to-provider links; the remaining are peer-to-peer links) are selected to fail concurrently in our simulation.

Our simulation shows that this example regional failure disrupts the reachability between 38,103 AS pairs, which mainly involve only 12 ASes, which we separate into 2 sets according to their failure patterns.

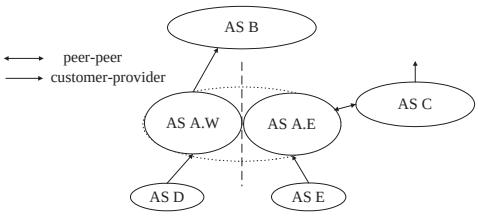

Figure 6: An example of AS partition

Case 1: One AS (located in South Africa) used to have 2 providers and 2 peers. The failure disabled its links to both of its providers, leaving it with only 2 peers to connect to the rest of the Internet.

Case 2: This set includes 11 ASes located in one of the European countries. Similar to the previous case, the failure caused breakage of their provider link(s). However, these ASes do not have peers, leaving them isolated from the rest of the Internet due to the failure.

In both cases, the affected ASes experience the failure of its shared access link(s) as discussed in Section 4.3 as their paths to Tier-1 ASes are disrupted. Regional failures cannot cause Tier-1 depeering due to their rich geographic diverse peering. *Most damage caused by the regional failures is due to the failure of critical access links.*

We also evaluate the potential impact of the failure on traffic caused by traffic shift from paths that used to traverse the affected region. This imposes extra traffic load on links in other regions. we fonnd T^{abs} to be as high as 31,781.

4.6 AS partitions

In this section, we examine scenarios when failures break an AS into two or more isolated parts and disrupt connectivity among these AS partitions. We first describe our analysis method before presenting the results.

First, we use an example in Figure 6 to illustrate how an AS partition disrupts reachability. AS A is partitioned into two parts, $A.E$ and $A.W$. A direct effect is that the communication between its separate parts is disrupted as $A.E$ and $A.W$ cannot reach each other unless their neighbors can provide extra connectivity to bypass the failure. (Special configuration, *e.g.,* tunneling, needs to be set as the neighbors cannot use the AS number to distinguish the partitions.) As described previously, the reachability resilience of an AS is indicated by the diversity of its uphill paths to the Tier-1 ASes. No reachability will be disrupted unless one of its partitions, AS $A.E$ as well as its single-homed customer E, loses connection to its only provider AS B. As such, *the AS partition becomes equivalent to the failure of an access link as discussed in Section 4.3.* Note that even though AS C in the example can no longer reach $A.W$, it can still reach $A.W$ through its provider(s).

In our analysis, we simulate a special case of AS partition in which a Tier-1 AS is separated into two parts. Due to the lack of detailed AS specific geographical information such as peering location, it is very challenging to model a network partition accurately. Since a Tier-1 AS spans over most of the country, we simulate the partition by breaking the AS into 2 parts: east region and west region. Based on its geographical presence from NetGeo data, we classify each neighboring AS of the target Tier-1 AS into 3 types: "east neighbor", "west neighbor" and "other neighbor" which resides in both regions. The failure only affects east or west neighbors. The Tier-1 AS in our simulation contains 617 AS neighbors, 62 of which in the east and 234 in the west.

In the simulation, we transform the old Tier-1 AS into two pseudo ASes. The east/west neighbors connects to only one of these new ASes while the rest of the neighbors have links to both ASes. Because Tier-1 ASes peer at many locations, the partition does not break any of the peering links. Failure only affects the communication between the single-homed ASes in the east and those in the west. To estimate the reachability impact, we choose R^{rlt} as the metric and S_i and S_j are the number of single-homed customers in east and west, respectively. Our results show that the partition disrupts 118 pairs of ASes with R^{rlt} 87.4%.

5. RELATED WORK

Several previous work [30, 31] on understanding the resilience of the Internet to faults are based on a simplified topology graph without policy restrictions and thus may draw incomplete conclusions. They also do not provide suggestions on improving failure resilience. We build on previous work [32] on analyzing how location of link failures affect the Internet and extend it to realistic topologies with routing policies as well as more general failure models. Our work also makes contribution in developing more accurate Internet routing models by focusing on the structure of the network. We take a different approach from recent work [33] by modeling routing decisions based on policies while accommodating multiple paths chosen by a single AS. Unlike previous studies focusing on obtaining complete AS topologies [9, 8], our focus is understanding how the topological structural properties affect routing resilience to failures.

In the area of understanding network resilience, a common method for analyzing network resilience is to compute the number of node or link disjoint paths between any pair of ASes, *i.e.,* path diversity of the Internet. Teixeira *et al.* [34] studied the path diversity problem both inside an AS (Sprint network) and across multiple ASes based on the CAIDA topology. In comparison, we present a more systematic evaluation of the resilience problem based on more complete and accurate topology data. Previous study by Erlebach *et al.* [35] also proposed using the min-cut analysis to compute the maximum disjoint paths between a pair of ASes, which is shown to be NP-hard. Instead of developing approximation algorithm, our paper simplifies the path diversity problem by precisely locating critical links between an AS and the set of Tier-1 ASes. Our technique is shown to be efficient and capable of identifying weakness in the Internet.

6. CONCLUSIONS

We have presented a comprehensive framework to analyze the resilience of Internet routing to common types of failures captured by our failure model which is developed based on empirical analysis. Our efficient simulation tool enables us to study how network topologies and routing policies influence network failure resilience measured using basic metrics of network reachability and traffic impact. Our results reveal that today's Internet might not be as resilient as we thought to be. 32% of the ASes are vulnerable to a single AS link failure. We demonstrate how restrictions imposed by routing policies can prevent network reachability under various failures, thus disallowing routing to fully take advantage of the underlying network physical redundancy. 255 (6%) non-stub ASes can no long reach other ASes during an AS link failure even though the physical connectivity might be available to bypass the failure.

In our future work, we plan to develop techniques which can be used to improve the network resilience. For example, we have learned that BGP policies restricts the paths each network takes to reach other networks, therefore, relaxing these policy restrictions could benefit certain ASes, especially under extreme conditions, such as failures. How and when we relax BGP policy is an interesting problem to pursue. In addition, we will explore the possibility of incorporating the traffic distribution matrix into our analysis to make a better estimate of the traffic impact caused by failures.

7. REFERENCES

[1] "Internet Routing Behavior on 9/11." `http://www.renesys.com`.

[2] "Asia scrambles to restore communications after quake." `http://www.iht.com/articles/2006/12/28`.

[3] H. Wang, Y. R. Yang, P. H. Liu, J. Wang, A. Gerber, and A. Greenberg, "Reliability as an interdomain service," in *Proc. ACM SIGCOMM*, August 2007.

[4] "University of Oregon Route Views Archive Project." `http://www.routeview.org`.

[5] "RIS Raw Data." `http://www.ripe.net/projects/ris/rawdata.html`.

[6] "Public route servers." `http://www.bgp4.net`.

[7] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz, "Characterizing the Internet hierarchy from multiple vantage points," in *Proc. IEEE INFOCOM*, 2002.

[8] R. Cohen and D. Raz, "The Internet dark matter - on the missing links in the as connectivity map," in *Proc. IEEE INFOCOM*, April 2006.

[9] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy, "A systematic framework for unearthing the missing links: Measurements and impact," in *Proc. NSDI*, April 2007.

[10] "CAIDA AS Relationships." `http://as-rank.caida.org/data/`.

[11] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, kc claffy, and G. Riley, "AS Relationships: Inference and Validation," *ACM Computer Communication Review*, vol. 37, no. 1, 2007.

[12] J. Xia and L. Gao, "On the Evaluation of AS Relationship Inferences," in *Proc. IEEE Global Internet Symposium*, 2000.

[13] L. Gao, "On Inferring Autonomous System Relationships in the Internet," in *Proc. IEEE Global Internet Symposium*, 2000.

[14] Z. M. Mao, L. Qiu, J. Wang, and Y. Zhang, "On AS-Level Path Inference," in *Proc. ACM SIGMETRICS*, 2005.

[15] G. Battista, M. Patrignani, and M. Pizzonia, "Computing the Types of the Relationships Between Autonomous Systems," in *Proc. IEEE INFOCOM*, March 2003.

[16] W. Muehlbauer, S. Uhlig, B. Fu, M. Meulle, and O. Maennel, "In search for an appropriate granularity to model routing policy," in *Proc. ACM SIGCOMM*, August 2007.

[17] X. Dimitropoulos and G. Riley, "Modeling Autonomous System Relationships," in *Proceeding of 20th Principles of Advanced and Distributed Simulation (PADS)*, 2006.

[18] L. Gao, T. G. Griffin, and J. Rexford, "Inherently safe backup routing with BGP," in *Proc. IEEE INFOCOM*, 2001.

[19] "Internet Routing Resilience Project Page." `http://www.eecs.umich.edu/~wujz/irrf/`.

[20] S. M. Bellovin and E. R. Gansner, "Using Link Cuts to Attack Internet Routing." May 2003.

[21] J. Wu, Z. M. Mao, J. Rexford, and J. Wang, "Finding a needle in a haystack: Pinpointing significant bgp routing changes in an ip network," in *Proc. NSDI*, May 2005.

[22] "The backhoe: A real cyberthreat." `http://www.wired.com/science/discoveries/news/2006/01/70040`.

[23] "ISP spat blacks out Net connections." `http://www.networkworld.com`.

[24] "NANOG mailing list." `http://www.merit.edu/mail.archives/nanog/`.

[25] "UUnet backbone problems slow down the Net." `http://www.itworld.com`.

[26] "Impact of Hurricane Katrina on Internet infrastructure." `http://www.renesys.com`.

[27] "PlanetLab." `http://www.planet-lab.org`.

[28] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. Cambridge, MA: The MIT Press, 2001.

[29] "NetGeo - The Internet Geographic Database." `http://www.caida.org/tools/utilities/netgeo/index.xml`.

[30] R. Albert, H. Jeong, and A.-L. Barabasi, "Error and attack tolerance of complex networks," *Nature*, 2000.

[31] R. Cohen, K. Erez1, D. ben Avraham, and S. Havlin, "Resilience of the Internet to Random Breakdowns," *Phys. Rev. Lett.*, 2000.

[32] X. Zhao, B. Zhang, A. Terzis, D. Massey, and L. Zhang, "The Impact of Link Failure Location on Routing Dynamics: A Formal Analysis,," in *Proceedings of ACM SIGCOMM Asia Workshop*, 2005.

[33] W. Muhlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig, "Building an AS-Topology Model," in *Proc. of ACM SIGCOMM*, 2006.

[34] R. Teixeira, K. Marzullo, S. Savage, and G. M. Voelker, "Characterizing and measuring path diversity of internet topologies," in *Proc. ACM SIGMETRICS*, September 2003.

[35] T. Erlebach, A. Hall, L. Moonen, A. Panconesi, F. Spieksma, and D. Vukadinovic, "Robustness of the Internet at the Topology and Routing Level," *Lecture Notes in Computer Science*, vol. 4028, 2006.