Internet topology: connectivity of IP graphs

Andre Broido and kc claffy

Abstract—

In this paper we introduce a framework for analyzing local properties of Internet connectivity. We compare BGP and probed topology data, finding that currently probed topology data yields much denser coverage of AS-level connectivity. We describe data acquisition and construction of several IP-level graphs derived from a collection of 220M skitter traceroutes. We find that a graph consisting of IP nodes and links contains 90.5% of its 629K nodes in the *acyclic subgraph*. In particular, 55% of the IP nodes are in trees. Full bidirectional connectivity is observed for a *giant component* containing 8.3% of IP nodes.

We analyze the same structures (trees, acyclic part, core, giant component) for other combinatorial models of Internet (IP-level) topology, including arc graphs and placeholder graphs. We also show that Weibull distrbution $N\{X > x\} = a\exp(-(x/b)^c)$ approximates outdegree distribution with 10-15% relative accuracy in the region of generic object sizes, spanning two to three orders of magnitude up to the point where sizes become unique.

The extended version of this paper [BC01b] includes dynamic and functorial properties of Internet topology, including properties of and diffusion on aggregated graphs, invariance of a reachability function's shape regardless of node choice or aggregation level, analysis of topological resilience under wide range of scenarios. We also demonstrate that the Weibull distribution provides a good fit to a variety of local object sizes.

I. Introduction

As the Internet continues to grow, so does the diversity of connectivity among nodes. The number of different paths among a given set of nodes depends upon unknown but crucial interconnection points that are beyond control of individual users and end customers. We seek insight into measures of infrastructural redundancy and robustness through analysis of Internet *topology* at the IP address granularity.

In this study we examine a large (220 million) collection of experimental ICMP forward path (traceroute) probes, obtained over a month in late fall 2000. We explore Internet topology expressed as a directed graph of IP address nodes and observed forward links between them. Skitter [Skit98], our data collection tool, is run by CAIDA on more than 20 monitors around the globe, collecting forward path and round trip time to about 400,000 hosts, with two or more probes sent to each destination each day.

Mapping macroscopic Internet topology is a daunting task, and we recognize the presence of shortcomings in our data and analysis. However, to our knowledge this work represents the most complete and reliable account of global Internet topology available thus far.

The differences between our data sources and those previously analyzed are:

1. This forward IP path data was collected by CAIDA's skitter [Skit98], a lightweight ICMP traceroute [Jac89] tool explicitly designed and extensively tested to gather IP topology data. Other studies use mostly UDP traceroute, whose packets are more often filtered by firewalls.
2. The data set is several times larger than in any previously available study of Internet connectivity. It includes responses from 655K nodes. The number of traceroutes, 220M, is three orders of magnitude larger than in other published analyses [Paxson97,ZPS00,PG98,SSK98,SCHSA99,GT00].
3. The IP destination addresses probed are specifically selected to stratify the IPv4 global address space via a variety of methods described here and in [FCHM01].
4. Among globally routed network prefixes, over 50% contain IP addresses that replied to probes. Previous Internet mapping does not quantify global prefix coverage.
5. The data was collected over 28 days, providing more of a 'snapshot' than collections that use a longer time interval.

Our methods of data analysis differ from previous work in the following aspects:

1. We use directed graphs, which more accurately reflect observed Internet connectivity. Most previous studies analysed symmetric graphs.[1]

Authors are with CAIDA, San Diego Supercomputer Center, University of California, San Diego. {broido, kc}@caida.org.

Support for this work is provided by the Defense Advanced Research Project Agency (DARPA), through its Next Generation Internet program, by the National Science Foundation (NSF), and by CAIDA members.

[1] Since routing is based on policy, the reverse link, even when fea-

2. We reduce the graph to the set of all nodes reaching bidirectionally connected nodes (core) and restrict it to its largest strongly connected component. This subgraph is guaranteed to have minimally acceptable coverage.[2]
3. We use complementary cumulative distribution functions (ccdf's) rather than frequencies of object sizes, which are more relevant to operational questions such as probabilities of buffer overflow.
4. We compare goodness of fit between formulas and data using a relative accuracy metric, which applies to widely varying magnitudes of experimental values. Previous work used correlations or absolute error for cdf's (rather than ccdf's) approximation, or avoided this question entirely.

New concepts and results presented here include:

1. Selection of combinatorial models for Internet topology on IP, router, prefix and AS level, and a variety of IP-level graphs, including IP-only, arc and placeholder graphs.
2. An algorithm for extracting the bidirectionally connected part of the graph.
3. Structural analysis of observed IP graphs in acyclic (downstream) and strongly connected (backbone) portion.
4. Measures of node importance such as sizes of neighborhoods, cones, and stub trees rooted at a node.
5. Demonstration that Weibull distributions provide a good fit to a variety of local object sizes.

The algorithms and data collection techniques described here are an integral part of the processing used in CAIDA's AS core map [HBCFKLM00]. The extended version of this paper [BC01b] includes dynamic and functorial properties of Internet topology, including properties of and diffusion on aggregated graphs, invariance of reachability functions' shape regardless of node choice or aggregation level, analysis of topological resilience under wide range of scenarios. In that version we also give more details on the superior connectivity coverage given by CAIDA's available forward probed topology data [Skit98] over that of the best available BGP topology data [Meyer01]. Finally, we provide greater detail on Weibull fits to distributions of sizes of Internet topological objects: router interfaces, stub trees and cones, neighborhoods of radius 2, IP addresses within a prefix/AS, subprefix and sub-AS connected components.

sible as a physical connection, may not always carry response traffic.

[2]The Internet's transport protocol, TCP, requires bidirectional connectivity. Traceroute probes also measure bidirectional connectivity since they depend on reply packets. These packets, however, do not carry information on the return path other than a TTL value, from which one can make only limited inferences.

A. *Roadmap of the paper.*

Section II describes previous related work.

Section IV describes our methodology for extracting the core of an IP graph. Section V compares measures of structural richness for the center and periphery of several modifications of IP graphs. We find that despite presence of holes left by non-responses and bogus addresses, and possible discrepancy between responding, receiving and forwarding interfaces, IP-only graphs are conceptually simpler, allow more coherent and transparent results, less topology distortions and less computational overhead, than those which involve bypass arcs or added pseudo-nodes, or router interfaces identification. Fortunately, many results obtained for one type of graph hold qualitatively for other types. Section IX presents conclusions and describes our future plans for topology analysis.

II. Previous work.

Analysis of Internet connectivity was pioneered by Paxson in his PhD thesis [Paxson97] and follow-up study [ZPS00]. Paxson acquired data over several months via traceroutes among academic hosts. A smaller collection of data on Internet connectivity was gathered in 1995 by Pansiot and Grad [PG98].

Siamwalla et al. [SSK98] present heuristics found useful for discovery of Internet topology, including SNMP queries, DNS zone transfers and broadcast pings. They correctly concluded that topology obtained by traceroutes from one source may be too sparsely sampled to be legitimately representative and that many sources are necessary to observe cross-links. Savage et al. [SCHSA99] collected and analyzed data among dozens of traceroute servers in the Detour project. These two studies focused on analyzing the stability and optimality of paths. Each of these studies dealt with less than 290,000 traceroutes.

Bill Cheswick and Hal Burch began a large-scale Internet mapping project in 1997, and made available on their website data of traceroutes to about 100K selected destinations [CB00], including six best paths to each destination over approximately one year. Cheswick and Burch also developed a novel algorithm for IP address level graph layout [BCh99,PM01]. One limitation of this data set is its lack of coverage of a globally diverse set of networks. Their single source (Lucent) renders a connectivity coverage bias toward their transit provider. In the April 2000 data set, the number of destinations, 103K, and BGP prefixes (over 55% of 80K), were significant, but destinations that actually responded comprised 28% of recent probes (22% when probed from our network). The destinations were not chosen based on routability.

Govindan and Tangmunarunkit [GT00] developed Mercator, an Internet topology discovery tool to build a router-level Internet map by intelligent probing from a single workstation. One strength of its design is its few *a priori* assumptions about Internet topology. They offer several valuable caveats of Internet topology acquisition. However, Mercator is considerably slower at processing probes than skitter and uses source routing to discover cross-links not captured by standard traceroute. This practice tends to generate more user and ISP complaints[3] and is less practical for large-scale longitudinal studies.

Radoslavov e.a. [RTYGSE00] compares canonical graph models such as a grid or a tree, with the Mercator, AS and Mbone graphs and with topology generators. They focus on the impact of topological properties on the performance of various flavors of multicast protocols. They make many meaningful observations in spite of the relatively scarce data coverage.

Broder et al. [BKMRRSTW00] reported the most extensive experimental study of a large Internet-based graph (200 million nodes, 1.5 billion links) using web connectivity (URLs as nodes and html links between them). Some of their results are applicable to IP level connectivity. Indeed, several sets of connectivity data, with different node and link types, collection intervals, sizes and coverage have similar properties, although such similarity may be a consequence of their incompleteness.

When this work was in its final stage, we learned about several recent papers dealing with Internet structure and topology, [TGJSW01] [CJW01] [TGS01] [PSFFG01]. We hope to be able to review the work presented there in a future publication.

III. BGP TOPOLOGY DATA ANALYSIS

Several studies on Internet connectivity have used AS (autonomous system) data extracted from BGP routing tables [NLANR97,Meyer01,PCH01]. Compared to traceroute path data, BGP tables are easier to parse, process and comprehend. It is understandable that researchers who do not collect their own data try to study Internet topology using BGP AS connectivity.

[3] Internet providers often flag source routing as a security threat.

BGP data is useful for determining correspondence between IP addresses, prefixes and ASes [HBCFKLM00], and in analyzing different routing policies in the Internet [BC01a]. However, BGP connectivity does not qualify redundancy of different parts of the network. BGP tables only show the *selected* (best) routes, rather than all possible routes stored in the router. Nor does the BGP table show public and private exchange points within the infrastructure, or short-term AS path variation and AS load balancing. BGP data may also not be directly comparable to traversed path data due to the presence of *transit-only* ASes, i.e. ASes who do not announce global reachability of their networks but show up in forward AS paths. In addition to engineering factors, BGP behavior reflects contractual business relationships among Internet service providers, specifying which companies agree to exchange traffic. It does not guarantee that this traffic will actually traverse listed administrative domains.

As such, using BGP data to obtain a topology map incurs significant distortion of network connectivity. In building graphs of topology core, graphs obtained by parsing even many dozen backbone BGP tables are extremely sparse. They represent some downstream (backbone to customers) connectivity, but no lateral connectivity. For example, extracting the largest component of bidirectionally connected nodes from RouteViews data [Meyer01] yields less than 3% of all nodes, even when contributing routers number in dozens, carry full backbone tables, and are geographically and infrastructurally diverse. In contrast, for topology data gathered from active probing from many sites, the largest bidirectionally connected component comprises 8% of IP-level nodes and 35% of AS-level nodes. (See Section V.) BGP data thus represents a relatively meager projection of Internet connectivity. It is thus imprudent to infer Internet properties from BGP data alone. In particular, Internet vulnerability, e.g., resilience to attacks, cannot be reasonably inferred from BGP data.

IV. EXTRACTING THE INTERNET'S CORE

In [BC01c] we introduced background notions from graph theory that assist analysis of traceroute-based connectivity data. We will assume the reader is familiar with that description and use the terminology from that paper.

To extract the cyclic part from the IP graph, we define an iterative algorithm called *stripping*.

DEFINITION. A graph obtained by removing all nodes of outdegree 0 and *edges* of all terminal 2-loops is called the *transit* (level 1) subgraph of the original graph.

A *transit level n* subgraph is defined recursively as the transit subgraph of the level $n-1$ subgraph, i.e. level $n-1$ subgraph with all nodes of outdegree 0 and edges of all terminal 2-loops removed.

PROPOSITION. A node is in transit level n subgraph if there is a path of length n starting (outbound) at that node. Transit level of a node equals the maximum length of such a path.

Since a graph is finite, it has only finitely many transit levels. We call the intersection of all finite transit level subgraphs the *(combinatorial) core* of the graph. It is essentially the part of the graph containing all cycles and their interconnections, except some 2-loops.

In the example below, the node H has level 0. It will be stripped first. The node G and the edges of 2-loop C $\leftrightarrow$ D will be removed next, leaving D disconnected from the rest of the graph. At the next step, the node C will be stripped. The nodes A, B, E, F belong to the core. Any node that can reach them (not shown) also belongs to the core.

$$\begin{tikzcd}
A \arrow[r] & B \arrow[r] \arrow[d] & C \arrow[r, leftrightarrow] & D \arrow[d] \\
E \arrow[u] & F \arrow[l] \arrow[r] & G \arrow[r] & H
\end{tikzcd}$$

A node that belongs to a combinatorial core must have a minimum cycle of size 3. A node that is not in the core can reach only cycles of size 2. Node in the core are cyclic; nodes not in the core are acyclic. The example above has cyclic nodes A, E, B, F and acyclic nodes C, G, D, H.

The subgraph that contains all acyclic nodes, all edges inbound or outbound on these nodes and all nodes that belong to these edges, has no cycles other than 2-loops. This graph will be called the *acyclic subgraph* of the graph. Note that the core and the acyclic part intersect node-wise, but not edge-wise. In the diagram above, nodes B and F belong to the core and to the acyclic subgraph.

The core of the graph consists of connected components of various sizes, starting from 1. In our analyses of Internet cores (IP, prefix, AS graphs and their variations) one component, the *giant component*, is significantly (200 times) larger than all other components.

Stripping of trees and of chains from the graph, as a means of finding its core, was previously used for BGP AS graphs in [Fa99] and for Mercator data in [RTYGSE00].

A. Connected components

2-loops. Removing 2-loops together with the acyclic part allows for filtering of connectivity noise caused by *multipath* packet propagation when packets follow paths of various lengths between the source and responding node.

$$\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[r] & C \arrow[r] & D \arrow[d] \\
E \arrow[rrr] & & & F
\end{tikzcd}$$

In the example above, a traceroute from A returns D at hop 3 and F at hops 2 and 4, which makes D $\leftrightarrow$ F appear as a 2-loop in the graph, when the route rapidly alternates between two paths.

Multipath routing is easier to observe when the number of hops is large. Disposing of terminal 2-loops together with the acyclic part of the graph potentially can reduce connectivity inflation at the edge of the network which is due to multipath propagation in the backbone and/or at the edge. It causes only limited pruning in the backbone, where nodes are contained in longer cycles.[4]

Giant component. The Internet evolved by combining smaller networks using the IP protocol into one giant network. All graphs reflecting its structure possess a giant connected component [BKMRRSTW00]. However, this component may not cover the whole graph.

One experimental weakness in the data acquisition setup where a few monitors collect paths to many (four to five orders of magnitude more) destinations is that it captures mostly one-way connectivity, from sources to destinations, and from backbone to downstream customers. Lateral connectivity is much harder to observe. (Some [GT00] use source routing to attempt to capture lateral connections.) This constraint reduces the bidirectionally connected part of the graph to a smaller portion: 1/10 of all nodes for forward IP graph, and about 1/5 for the

[4]Strongly connected components can be obtained algorithmically by raising connectivity matrix of the graph to powers in Boolean arithmetic until saturation occurs. Our implementation finds connected components, shortest path distributions, reachability functions etc. for a graph of 60K nodes on a high end workstation using 1.05G of RAM in 32 minutes of wall clock time.

forward AS graph. For the BGP AS graph, the bidirectionally connected component is about 1/30 the size of the original graph.

Bidirectional connectivity is present even in single-monitor graphs, as a side effect of policy-conformant paths not being selected on the basis of minimum hop count. Using several monitors has the advantage of traversing the backbone in various directions, which increases the sampled bidirectional connectivity, and provides an arguably more legitimate representation of the Internet 'core'.

B. Cones and stub trees.

We will now introduce several measures that quantify node's importance in the overall connectivity of the graph.

Cones. Recall that the merged traceroute graph contains large portions where only downstream connectivity can be observed. These parts usually contain only 2-loops, some of which can represent TTL noise caused by multipath propagation (see above). Being nearly acyclic, the subgraph composed of these parts is close to a partial order.

DEFINITION. A *cone* of a node A consists of all nodes reachable from A via the acyclic subgraph.

The root A need not itself be acyclic. In fact, cones are most useful for comparison of giant component nodes.

A cone consists of those nodes that one or more traceroute paths find downstream from the cone's root node, and for which an upstream connection has not been observed. A root node is viewed as important when the cone is large, since many downstream nodes can be reached through it. Unlike the nodes in a tree, these downstream nodes can possibly be reached through several cones. Cone sizes can thus overestimate node's importance. To reduce this bias, we will study *stub trees* in conjunction with cones.

DEFINITION. We will call a subgraph a *stub* tree if it is connected to the rest of the graph only through its root.

Trees and cones are useful measures for representing the structure of the acyclic subgraph. This subgraph contains trees and denser part with some node indegrees greater than 1 and with some 2-loops. Below we provide statistics for the major variations of the Internet IP graphs.

C. Subprefix and sub-AS connected components.

For network engineering purposes and protocol design it is important to know the diameter of a network in IP hops. Traceroute data may not always provide a reasonable answer since the the longest path through the network may not be followed by any actual traceroute in available set.

It is possible to simply approximate a network diameter with the maximum of the lengths of shortest paths consisting of links observed in several traceroutes. However, this estimate may be inaccurate. The coverage of the network by traceroutes may be incomplete, which will result in longer paths and size overestimation. Conversely, shortest paths can also underestimate the lengths of the policy-conformant paths and IGP-based network diameter. Nonetheless, this metric is the closest approximation to diameter that we can reasonably make.

To capture a greater extent of network connectivity, we will define two nodes in a prefix as connected at distance 2 if they are both reachable in one hop from a responding and valid IP node outside the network (common entry point), or they reach an outside node in one hop (common exit point.) We will allow arcs (bypasses of non-responding and bogus IP nodes) together with direct IP links. For simplicity, each arc hop will be counted as one IP hop.

DEFINITION. A subset of nodes which belong to one network is a *subnetwork (subprefix or sub-AS)* connected component, if it is weakly connected (connected by links followed in either direction) by IP arcs within the network, and by common entry and exit points.[5]

Use of arcs, symmetric links, and common entry/exit points for subnetwork components avoids connectivity underestimation, e.g. a case when a stub network is served by a border router with an interface numbered out of a different address block. Otherwise such a network would appear as having no connections between its IP nodes. In the extended version [BC01c], we include details of our data sources, idiosyncracies of the resulting observations, and how the data is affected by accumulating measurements over time. We omit this discussion here in the interest of brevity.

V. Structure of IP graphs

Resolution of ambiguities. Close to one-third of

[5]Connected components based on intra-AS IP links were independently and implicitly used in [TGSE01].

probed paths contain non-responses. Other traces contain private or invalid addresses. The paths can either be treated as broken (discontinuous) at these nodes or we can try to recover connectivity information from them.

To preserve as much connectivity information as possible, one method is to add *arcs* which bypass hops where the reply is missing or unacceptable. Arcs connect valid replying IPs and skip undefined nodes. To preserve the metric, we need to account for arc lengths (which measure how many hops are covered by an arc). For graphs of IP-level size this is computationally expensive. Alternatively, we can introduce a placeholder for each non-unique node, as follows.

DEFINITION. Let A and B be two responding valid IP addresses in a traceroute path, and let entries in between, $b_1, \ldots b_n$ be non-responses or bogus addresses. Placeholder for b_k is given the name A-i-b_k-j-B, where i and j are integers with $i + j = n - 1$. Placeholder nodes are connected according to their position in the path.

Placeholder graphs preserve both connectivity and hop metric. However, they have a much larger number of nodes than IP (IP-only or arc) graphs. Another problem, which they share with arc graphs, is that they overestimate local connectivity (node degrees), since they partially implement transitive closure of IP path from source to destination.

To clean up those graphs, we can use only placeholder chains whose connectivity is not duplicated by shorter or equally long IP paths. We will call this variation the *shortcut* placeholder graph. We will provide a numeric comparison for all three types of graphs later in this section.[6]

Router graphs. If the goal of Internet topology discovery is to build a router level map [GT00], then interfaces need to be identified with routers. A published technique [PG98,GT00] is implemented in CAIDA's *iffinder* [Keys00], which sends UDP packet to unused port and registers the replying source address, which is one of IP addresses of the interface on which a packet is sent [RFC1812]. The relation between address pairs makes up an *IP alias* graph, whose weakly connected components (connected components of its symmetrization) are viewed as routers.

Figure 1 presents *iffinder* data of 16-17 Feb. 2001

[6]The graphs analyzed here contain no addresses in 0-2, 224-255 (multicast and reserved), 10 or 192.168 (private) range. They contain 1538 addresses in 172.16-31 (private) range, which are present in 9938 links. Out of these, 67 addresses are in the giant component.

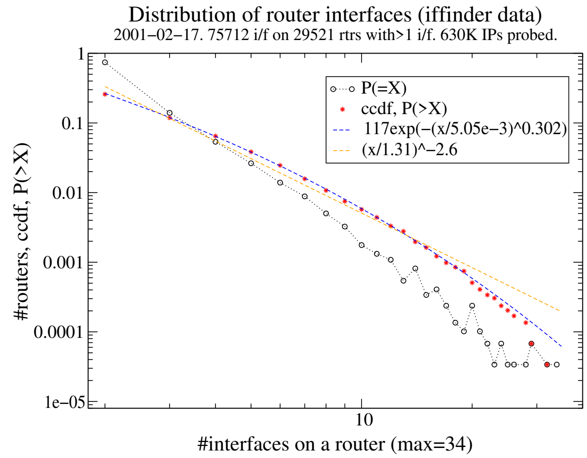

Fig. 1.

obtained by querying from CAIDA network 638K interfaces found in CAIDA's topology data of Nov-Dec. 2000. An interface was not queried after one alias was obtained. That is why the alias graph has only outdegrees 0 or 1. It therefore consists of inbound trees, which can be rooted on loops, and the loops have no edges pointing outside. 75712 IP nodes connected by 46716 alias links contain 29521 weakly connected components. Figure 1 shows the distribution of their size, i.e. the number of interfaces on a router. It is closely followed by Weibull distribution $117\exp(-(x/0.00505)^{0.302})$. We discuss Weibull approximations in Section VIII.

Among connected interface components, 484 have a 2-loop, 33 a 3-loop and 8 a 4-loop. There are 111 inbound trees rooted on 2-loops, and 17 trees rooted on 3- and 4-loops. Except for 525 loops and attached trees, 29K components are non-attached trees, with an overwhelming majority (22K) being just standalone pairs. All trees have height between 1 and 6 alias pairs. This characteristic is not what common sense reading of [RFC1812] suggests.

Cp.sz	2	3	4	5	6	7	8	9	10	11-34
#cpt.	21902	4122	1584	778	410	260	148	95	52	160

We checked one 4-loop by traceroute and found four IP addresses in Europe which indeed refer to each other in a round robin fashion. The IP addresses have 24 bits in common; they may address the same physical interface.

Alias resolution has two limitations. It may be impossible to find all identifications, and it is hard to quantify how many are found. We did not use *iffinder* results for merging traced path data,[7] since aggregation using in-

[7]This problem is discussed in [CJW01] which analyzes twice as many alias pairs as here.

complete equivalence introduces more ambiguities than it resolves. The identification is also separated by two months from the measurement interval, further compromising the integrity of the merged graph.[8] Note also that interfaces are individual devices, with their own individual processors, memory, buses, and failure modes. It is reasonable to view them as nodes with their own connections.

An interface address returned by traceroute with a TTL expired message may be different from the interface entered or exited by a packet on its forward path toward the destination. This difference occurs when routing is locally asymmetric on an IP level, i.e. the outbound interface address differs from the receiving interface address. Discrepancy in IP addresses can introduce ambiguity in traceroute data, especially if the returned address is in a CIDR block that does not belong to the operator of the router [CJW01]. We do not currently have reliable data quantifying the extent of this ambiguity. It is possible that it is of the same order of magnitude as that caused by the use of private and unrouted addresses in the backbone.

Stripping. We obtain the graphs by parsing traced paths. In that parsing, non-responses, bogus and private addresses are treated in accordance with the graph type. Unidirectional (downstream-only) connectivity represented by the acyclic subgraph is filtered out at the next stage by recursive removal of nodes and 2-loops with outdegree 0.

As we strip increasing transit levels from the graph, the number of nodes removed from the acyclic subgraph shrinks in a quasi-exponential way (like a regular fanout, e.g. a tree with k branches at every node). The corresponding average rate is a global characteristic associated with data set and graph type rather than of individual nodes or transit levels.

Two types of IP graphs have qualitatively different fanouts. Pure IP (IP-only), arc and router graphs have a fanout factor close to 2, like that of a complete binary tree. For the IP graph, the average decrease in the number of nodes in the graph between levels 0 and level 10 is 2.046; between 0 and 14, 1.975. Placeholder graphs have smaller average fanout (1.43, for nodes removed between level 1, and level 31), close to the square root of IP graph fanout. Their core height is also about twice that of IP

[8]Network prefixes currently change at a rate of 3-6% per month [BC01a]. Individual IP addresses become unreachable at a rate of 1-2% per month [Fomenkov00].

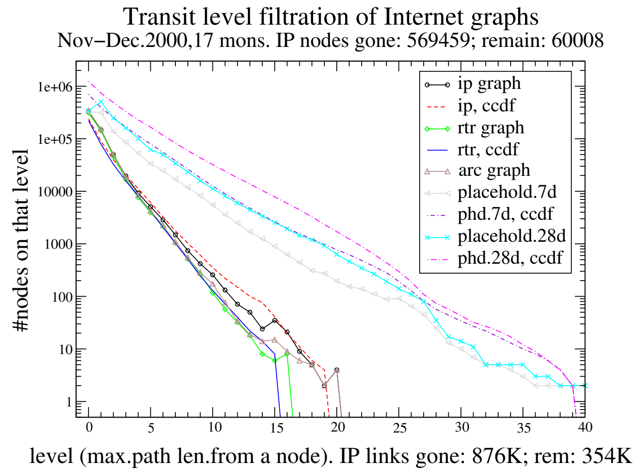

Fig. 2.

graphs. The validity of approximation by an exponential function is confirmed by the fact that ccdf's for removed nodes follow an exponential with about the same decay rate, in agreement with calculus, $\int_x^\infty e^{-at}dt = e^{-ax}/a$.

A. Structural statistics for IP graphs

Four candidate graphs for Internet representation on IP level are: IP-only (pure IP); IP arcs; placeholder; and shortcut placeholder graph. The tables below compare these graphs for the 28 days of data in Nov-Dec. 2000.

Graph type	IP-only	IP arcs	Placehd
nodes	629647	654945	2431590
links	1230572	1929445	4093701
links/nodes	1.95	2.95	1.68
outdeg.0 nodes	330752	341790	340259

max.tree height	9	9	21
tree nodes (no roots)	348354	252411	221866
perc. tree nodes	55.3%	38.5%	9.1%
non-tree nodes	281293	402534	2209724
non-tree links	886265	1677035	3871836
2-loops removed	2486	2226	2316
core height	21	21	≈ 40
core nodes	60008	73688	847346
Perc. core nodes	9.53%	11.25%	34.8%
core links	354250	710999	1629824
links/nodes	5.9	9.65	1.92
outdeg 1 core nodes	21931	24025	788351
outdeg$\geq$ 2 core nodes	38077	49663	58995
indeg 1 core nodes	10338	7848	772851
max in-core outdeg	569	767	850
outdeg.geo.mean	2.82	3.48	1.14
geo.mean, outdeg$\geq$ 2	5.12	6.35	6.23

Placeholder graphs have a large number of nodes, especially those with out- and indegree 1, caused by their construction which avoids accounting for arc lengths. The number of branching (outdeg$\geq$ 2) nodes in their core is comparable to that of IP and arc graphs (59K). Algorithms that account for variable arc lengths could probably perform well on these graphs.

We have also collected a 7-day placeholder graph for a week before the software upgrade. This graph has 1336707 nodes, 2115274 links, tree height 19 and core height 36.

The shortcut placeholder graph for 28 days of data contains 1.78M nodes and 3.07M links. Its level 17 transit subgraph has 615K nodes and 1.26M links, of which 563K belong to outdegree 0 nodes. The number of nodes with outdegrees 2 or more in this subgraph is 51869, and the largest outdegree is 772.

Core metrics. IP-only graph vs. IP arcs graph.

	IP-only	IP arcs
core diameter	32	31
g.c. nodes	52505	67939
g.c. IP links	324933	682235
links/nodes	6.19	10.04
g.c. placehd.nodes	613783	704088

The number of outdegree 1 nodes in the IP-only giant component is 18456; of indegree 1, 10293. The maximum outdegree is 563; indegree, 690. The large number of nodes of indegree and outdegree 1 in the core suggests that it contains many subgraphs that are inbound (fan-in) and outbound (fanout) trees. These trees inside the giant component may represent traffic aggregators (concentrators) and deaggregators in the backbone. This question needs further analysis. Removing edges with outdegree 1 reduces the giant component to 29890 nodes (57%); thus, 4159 nodes (8%) in original g.c. connect back to it only through the nodes of outdeg.1. The reachability of g.c. drops to to 48564 nodes, i.e. by 7.5%.

Routed-only IP graph. If we remove all unrouted (non-advertised in BGP) nodes, the IP-only graph reduces to a routed-only IP graph with 620184 nodes and 1203385 links. Its core has height 21 transit levels and contains 57998 nodes and 341290 links, with 21334 nodes of outdegree 0 and geometric mean of outdegrees 2.82 (for outdegrees of 2 or more, 5.12.) The diameter of the core is 32; the average shortest path length 6.74 and the giant component has 50175 nodes. Qualitatively it is identical to the graph with unrouted IP nodes; quantitative differences are under 5%.

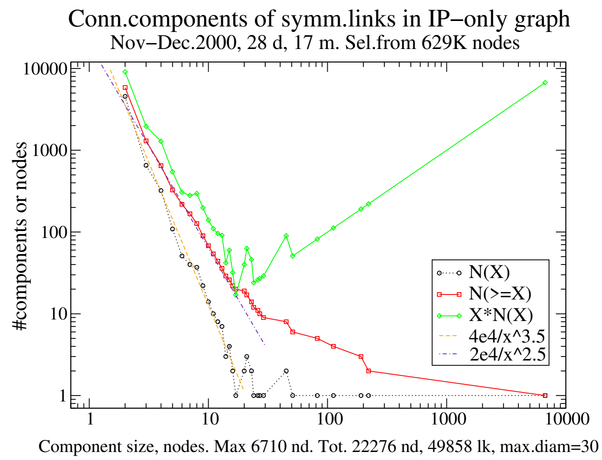

Fig. 3.

Symmetry of links. Most graph theory is developed for symmetric graphs. As already discussed, one cannot assume symmetry in traceroute data. The router must send ICMP messages (e.g., 'TTL expired', in traceroute) with the "IP source address...one of the addresses associated with the physical interface over which an ICMP message is transmitted [or] the router-id used instead." [RFC1812, 4.3.2.4].[9] Getting different interface IP addresses from the same router is possible, when return routes differ for different monitors. In particular, traceroutes that pass two adjacent routers in different order are likely to encounter not two pairs of interfaces matching each other in reverse order, but four interfaces with four different IP addresses.

In the IP-only graph with 629K nodes, 22276 nodes have at least one symmetric link. The number of links for which the reverse link is also in the graph is 49858 (8%). The maximum size of the subgraph connected by all these links is 6710; the number of connected components is 5855. The concentration of symmetric links is slightly higher in the giant component, in which 14482 nodes have symmetric links and 37902, or 11.7% of links (18951 pairs of links) are symmetric. Symmetric connectivity is an order of magnitude less than directional connectivity of the IP graph.

Figure 3 shows the distribution of these symmetrically connected components, whose frequencies can be approximated by power functions $4e4/x^{3.5}$ and ccdf by $2e4/x^{2.5}$. The matching powers (3.5 vs. 2.5) is in agreement with calculus, $\int_x^\infty t^{-c-1}dt = x^{-c}/c$ which sug-

[9] [GT00] say that they checked and confirmed this property for equipment from two major router vendors.

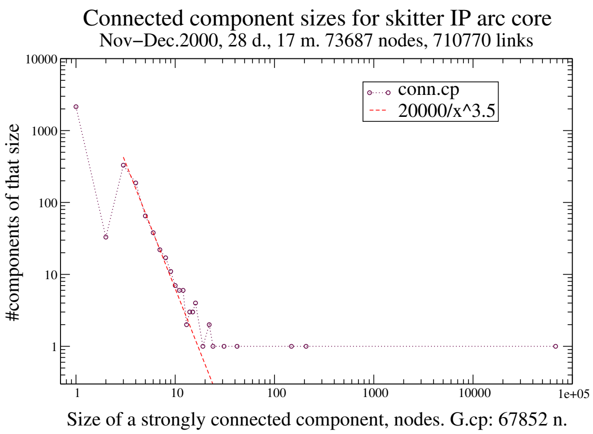

Fig. 4.

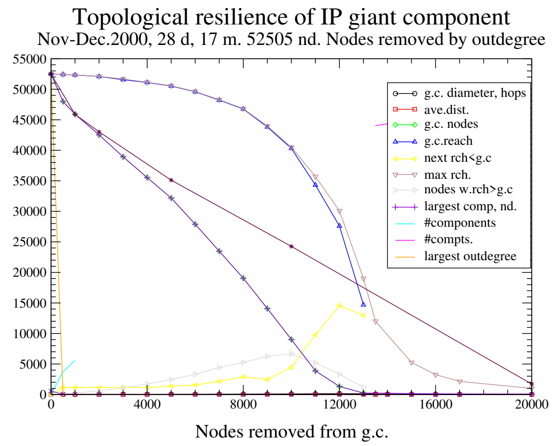

Fig. 5.

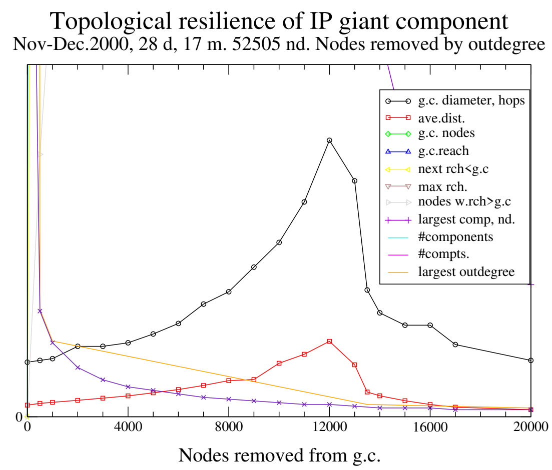

Fig. 6.

gests that the fit is not an illusion. In the followi sections we will show that many *local* object sizes well with the Weibull distributions. The symmetric cc nected (sub)components reflect global connectivity of graph, i.e. they are nonlocal objects, so we do not e pect Weibull to apply. Note also that the approximati for the arc core's connected components sizes (discuss next) has about the same power.

Core component sizes. The size distribution for cc nected components of the arc core (Figure 4) is clc to the power function $2e4/x^{3.5}$ for component sizes b tween 3 and 10 nodes.

We observe that count of 2-loops stripped from t core (and missing almost completely from the distrib tion) comes close to the count predicted by this formula. It differs from one reported in [BKMRRSTW00] for the Web's weakly and strongly connected components, in that it falls off more rapidly (has larger exponent). CAIDA's monitors mutually probe one another, which makes the IP graph as a whole weakly connected. Two mid-size strongly connected components (208 and 148 nodes) come from the networks of a Japanese electronics firm (two /24s) and a New Hampshire ISP (one /24) in which every host can forward packets to at least one other host.

VI. Topological resilience.

Resilience of the graphs to removal of nodes has been the subject of a number of recent studies [AJB00], [CEAH00] [CNSW00] [PSFFG01]. We tested properties of the giant component (combinatorial backbone) of the IP-only graph with respect to removal of nodes with largest outdegrees, or those with smallest average distance to the rest of the graph.[10]

In our experiment, outbound edges of nodes are deactivated in order of decreasing outdegree. Forwarding via these nodes becomes impossible, although they may themselves still remain reachable.

It turns out (Figure 5, bottom curve) that the IP giant component size decays smoothly, almost linearly relative to the number of deactivated nodes. It takes a lot of node removal to destroy it completely. For instance, when the top 10000 out of 52505 (about 19%) nodes in the network have forwarding disabled, we are able to reach 40310 (77%) nodes and still have a giant component of size 9020 (17.2% of total) nodes. To our knowledge, this property of the IP topology graph does not match any theoretical analysis. It completely disappears only when

[10] The results presented here are preliminary. We will integrate more complete analysis before the workshop.

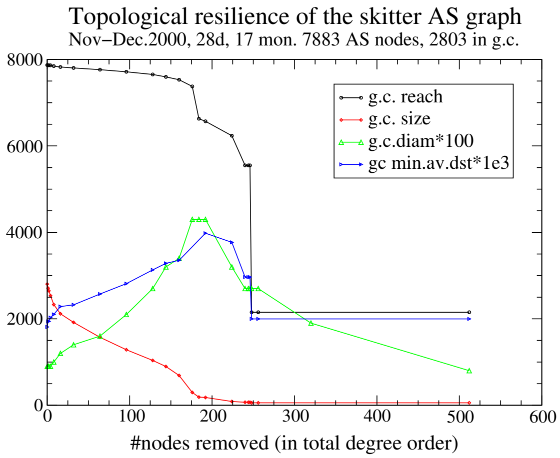

Fig. 7.

25% of nodes have their forwarding deactivated. On the other hand, reachability of nodes by the giant component has a concave slope with accelerating decrease of connectivity.

Width measures of the giant component, such as diameter and average distance (Fig. 6) increase as nodes are deactivated, and have a characteristic jump where the giant component finally breaks down. This behavior is qualitatively similar to that described in [AJB00] for models of scale-free networks. However, unlike [AJB00], the curves appear to be highly asymmetric around a critical point.

Removal of nodes in the order of average distance to the rest of the giant component (middle curve) has significantly smaller impact than removal by outdegree.

Decay of the forward AS graph when nodes are removed in order of total degree (indegree+outdegree) (Figure 7), on the other hand, is qualitatively similar to predictions of percolation theory [CNSM00]; the authors of [CNSM00] assume a power function as the outdegree distribution. The powers that they use for simulation and numeric evaluation of formulas, are close to those presented in studies of outdegree of web URL graphs [BKMRRSTW00] and BGP AS graphs [Fa99]. It starts steep (with large derivative and concavity) and then changes to linear decrease when the most well-connectes nodes are removed. The end of the decay has a drop on the forward AS curve that does not match [CNSM00] predictions.

VII. Distributions and approximations

Networking infrastructure operates under constant resource pressure, in particular with respect to bandwidth and computational resources in routing and switching equipment. For example, continued prefix table growth [Hou01] [BC01a], threatens to upset a delicate balance between equipment investment and operating margins, both in terms of router memory and the computational power required for timely route selection and table maintenance.

Limited buffer size and other resource constraints are a commonplace in networking and computing. We can quantify the cost of resource optimization by the number of objects lost (e.g. denied service) due to size cutoff at x, $N\{X > x\}$, where X is object size. The fraction of objects lost is expressed by the complementary cumulative distribution function (ccdf), $Q(x) = P\{X > x\}$. [11]

A potential problem with ccdf is that the values of experimental ccdf have different statistical significance (since they are sums of varying number of data samples), and they are not independent. For ccdf studied here the latter is not much of a concern: the tails are small enough so that they do not change ccdf's order of magnitude for x at the lower end of the scale.

Standard sample estimators (median, mean, mode, variance) lose their meaning for skew distributions associated with most Internet object sizes. Their goal is to compare data to a delta function, centered at certain representative point and spread around this point in a limited and relatively symmetric way. Internet data, however, often has most frequent values at the lower end of size spectrum, and its spread up from that size does not have any intuitive meaning either.

It may be better to approximate data with formulae that can accurately estimate probabilities of objects with widely varying sizes, especially large sizes, as these cause buffer overflows, server meltdowns and other undesirable phenomena. These probabilities can be quite small.[12]

Approximating the tail of the ccdf in uniform (Chebyshev) metric, e.g., as implied by the Kolmogorov-Smirnov test, can be misleading, since any fixed accu-

[11] We will often use $N\{X \geq x\}$ in place of the ccdf since it shows the data range at both ends of the plot: the total number of objects and largest object size.

[12] One exception: IP packets of 1500 bytes (largest size) comprise about 21% of observed packets [BCN00] [CAIDA01].

racy expressed as an absolute error will eventually become overwhelmingly coarse compared to the probability of tail events. We avoid this difficulty by using the *relative error*,

$$r = \max(y/a, a/y) - 1;$$

where y is function value and a its approximation.

Small values of relative error across sufficiently long ranges guarantee good approximation for both frequent and rare events. We use the relative error when comparing approximations for object size distributions. Minimizing relative error is equivalent to approximating the logarithm of the distribution in the uniform (absolute accuracy) metric.[13] When the relative error is small (10% or less), these quantities are also numerically close, since in that case $r \approx |\log(y/a)| = |\log(y) - \log(a)|$.

To specify a meaningful threshold for relative error, recall that memory upgrades are usually done in increments of 50% or more (e.g. from 256M to 384M.) A rule of thumb for the approximation error is then 20%, since 1.2/0.8 = 1.5, so upper and lower ends of the interval differ by not more than 50%. We will view approximation as good if its relative accuracy is under 20% in the interval of arguments where approximation applies.

Similarity of the data to distributions is often shown using log-log plots. For the region of large arguments (the right side of graph), the proposed approximations may not apply. This region can escape attention when viewed on log plots, even when it covers up to 70% of the argument's range. Hence, we need to specify not only the relative error, but also the region in which approximation is acceptable.

VIII. Weibull approximations

In [BC01c] we showed that many different measures of Internet object's inherent strength or complexity can be approximated by a *Weibull distribution* [Ext00] [Gr92]

$$N\{X > x\} = a \exp(-(x/b)^c)$$

The Weibull distribution is well known in reliability theory and other applied sciences. [FS97] [LS98] [Ext00] [Gr92] [JK75] [BKN00]. However it has not been used to approximate Internet object size distributions, except for traffic analysis [Nor95] [BH00].

[13]The relative error metric (both as a ratio and as a logarithm) was introduced by Chebyshev [Che1889].

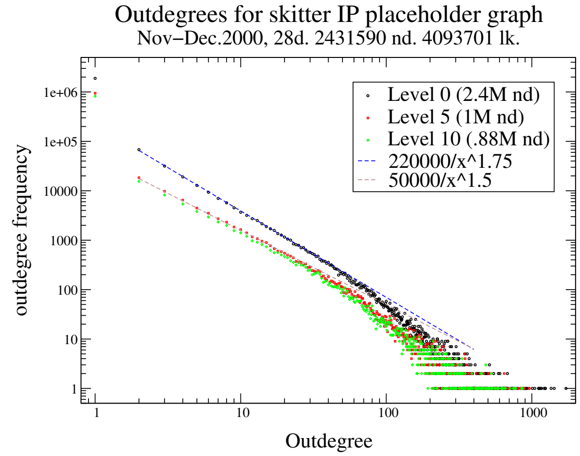

Fig. 8.

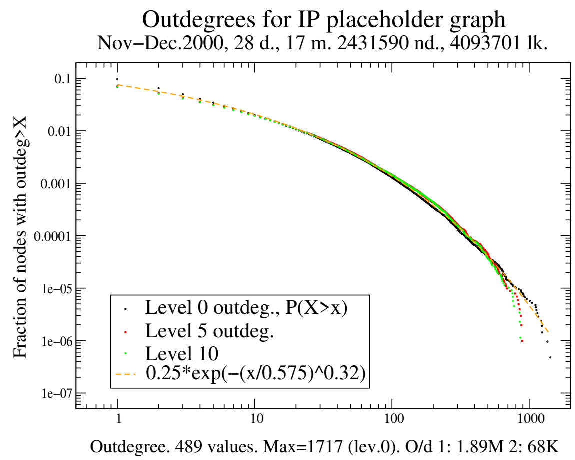

Fig. 9.

Figures 8 and 9 show statistics of outdegrees (outgoing edge counts). Non-responding intermediate nodes are given temporary unique names, so the number of nodes and links is larger than actually observed. The approximation we present is of medium quality (14% on half of the range.) It holds with different parameters for all other types of IP-derived graphs.

We start by observing that $P\{\text{outdegree} > 1\} \approx 9.73\%$. Over 90% of the vertices have outdegree 0 (these make up 14%) or 1. The initial (left-hand side) portion of the outdegree data in Figure 8, except for prominent outdegrees 1 and 0 (artifacts of the placeholder graph construction, though degree 0 is not plotted on the log axes), is visually similar to a power function. The distribution for the whole (level 0) graph looks close to $Cx^{-1.75}$. The distribution for the transit level 5 subgraph (nodes with outbound paths of 5 or more hops) appears to be close to $Cx^{-1.5}$.

Unlike the frequency plot, the ccdf for the whole graph

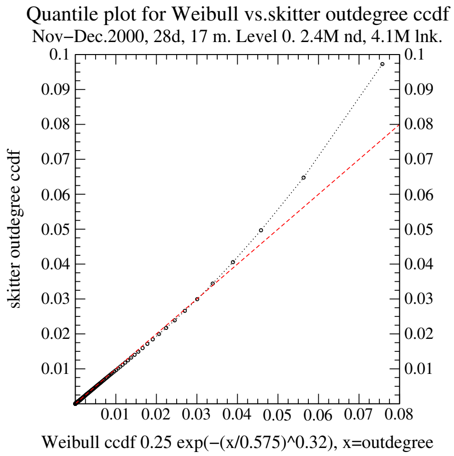

Fig. 10. (linear scale)

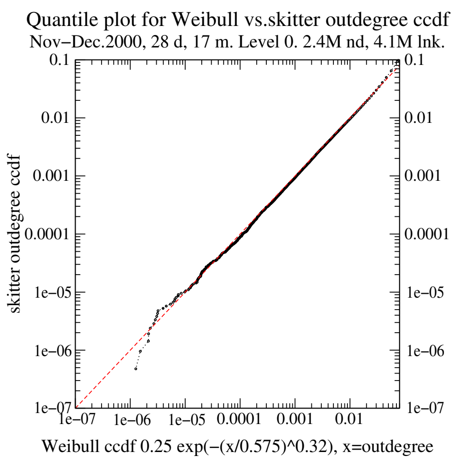

Fig. 11. (log scale)

appear to be close to $a\exp(-(x/b)^{0.32})$, with $a = 0.24$ and $b = 0.575$, [14] rather than a power function. Knowing this, it is easy to observe that the outdegree frequency plot starts to bend down around $x = 50$. This is one of the reasons why the ccdf has a shorter tail than that of a power function.

Figures 10 and 11 show quantile plots (linear and log scales, respectively) of the Weibull approximation to data against quantiles of the data's ccdf. Every data point (outdegree value) is assigned two coordinates: x, the Weibull distribution value at that argument; and y, the value of the ccdf of outdegree frequencies in the graph.

The quantile plot of Figure 10 shows that the outdegree distribution deviates from Weibull for the first few points, after which they become visually indistinguishable. The log plot emphasizes the tail of the distribution, where the data oscillates around Weibull approximation, in accordance with Chebyshev's theory.

To determine adequacy of fit, we searched over a uniform grid with $21 \times 21 \times 21$ (9261) values of Weibull parameters. The best fit occurs at values where factor $a = 0.233$, scale $b = 0.55$, and shape parameter $c = 0.317$, i.e. $0.233\exp(-(d/0.55)^{0.317})$. This fit minimizes logarithmic error, i.e. the logarithm of the ratio between the approximation and the data. Approximating data "by hand" was therefore reasonably close to the optimum fit.

Comparing the empirical data with the approximation shows that Weibull's relative error in $[4, 839]$ is under 14%. This interval contains 466 outdegree values out of possible 489. Weibull does not approximate, however, the last 20 values which belong to $[870, 1717]$. (The 20% cutoff is exceeded at 1009.) It covers only half of the total data range on linear scale. On the other hand, for data between 7 and 300 (half the range on log scale) with 293 outdegree values, relative error is under 7.5%. Compared to data, Weibull starts at 4 with underestimation, changes sign at 113, then again at 534, after which oscillations become more frequent. The largest deviations from data are observed at 1 (40%) and second from the end value 1427 (176%). [15]

Note that each object size greater than 690 (of which there are 31) occurs only once in the distribution. In this and many other examples. Weibull applies for generic sizes, and loses accuracy for the sizes with unique object counts.

Finding a power function as close to the data as Weibull is impossible for the following reasons. If three arguments $x_1 < x_2 < x_3$ have function values $y_1 > y_2 > y_3$, then the minimum of approximation error by a linear function

$$\min_{a,b} \max_{k=1,2,3} |y_k - ax_k - b|$$

over all a, b equals $\frac{1}{2}|y_2 - (y_1 + \frac{y_3-y_1}{x_3-x_1}(x_2-x_1))|$ i.e. half the vertical distance between y_2 and the line connecting (x_1, y_1) with (x_3, y_3). Taking ccdf values at 4, 100 and 839 (ends and mid-point of 14% accuracy interval) we find that the best approximation of these three points by a straight line in log coordinates cannot have relative error less than 120%, about 9 times more than Weibull.

The best approximation we found is $4.7(x/0.31)^{-1.55}$. Note that -1.55 is closer to the power with which individual frequencies are decreasing, -1.75, than to the

[14] Weibull applies only to the degrees over 1, thus $a < 1$, whereas $a = 1$ in textbook formula [Ext00].

[15] For the last value, ccdf=0 and the relative error is undefined.

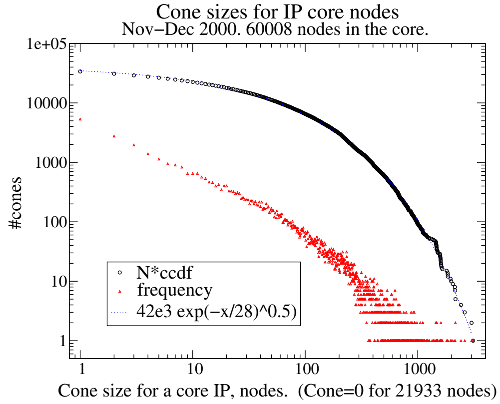

Fig. 12.

power -0.75 for the integral of $x^{-1.75}$. This mismatch is fairly common for distributions generally believed to be close to power functions.

Relative error of 120% means that the power function can be up to 2.2 times larger or smaller than the experimental ccdf for outdegrees. The magnitude of difference depends, of course, upon data type. For some kinds of Internet data, e.g. prefix counts in policy atoms [BC01a], the two functions can be much closer.

Figure 12 shows IP-only graph's core nodes' cone size distribution, i.e. number of nodes in the cone excluding root. There are 853 sizes, of which 46 are larger than 1271. The largest cone size is 3125 nodes. Our hand-picked Weibull approximation for cone sizes between 1 and 1271 nodes has relative error 15%. It is visually almost undistinguishable from the data in that range, Using computer search, we found that that in [1,1271] formula $41 \cdot 10^3 \exp(-(x/27.8)^{0.5})$ has relative error 11%.

The extended version of this report provides details on how we found Weibull to be a good fit for stub tree sizes, sub-prefix and sub-AS connected component sizes and number of components in prefix or AS [BC01b].

IX. Conclusions

We presented a structural description of Internet topology as represented by IP-level graphs obtained from 220M paths traced toward destinations covering over 50% of globally routable BGP prefixes [Meyer01]. Most IP nodes are found in downstream (backbone to end user) portion of the paths, which results in most of the graph (90%) being in an *acyclic subgraph*, and 55% of all nodes belonging to stub trees. The part of the graph with full bidirectional connectivity (giant strongly connected component), which includes the global 'IP core backbone', contains 8% of nodes.

We estimated how much connectivity information is lost due to non-responses and bogus addresses, and found that skipping gaps can add up to 30% nodes to the giant component. To correctly analyze the metric structure of these *arc* graphs it is necessary to account for links spanning more than one hop (splitting them results in explosion of nodes observed in *placeholder* graphs.) We did not pursue this approach since it overestimates the local connectivity of many nodes in the graph, and yet the change in the giant component size is small compared to the increase in the complexity of algorithms and computational overhead.

We presented a number of examples that confirmed that Weibull approximation fits several different Internet topology object size distribution, in particular those that are non-unique (assumed by more than one object). In the range of applicability, the relative error can be impressively small. Approaching ccdf with relative errors of 10% and even 7% in the central range of the distribution is common. For a plot spanning three to five decades on both axes this level of error produces a curve that is visually indistinguishable from the empirical data in the range where approximation applies. The Weibull approximation sometimes breaks down at the lower end of the size spectrum, often due to the exceedingly large number of small objects, typical for current state of the industry and network engineering design. The extent of this breakdown is different for different types of data. It is less pronounced for ccdf of the form $P\{X > x\}$.

The Weibull approximation generally appears to apply to local size measures (e.g., immediately adjacent connectivity), for parameters intrinsically controlled by an object and not dependent upon the global environment. Several open research questions remain as to whether there is a general cause or many unrelated reasons for the Weibull approximation to hold, and whether good approximation is an exclusive property of this formula. In future work we will analyze other ways of approximating Internet data, including three- and four-parametric families of functions which generalize power functions.

X. Acknowledgments

This paper contains a summarized version of the results that we obtained in 1999-2001, and presented at ISMA workshop in December 2000 [BC00a][BC00b], NRDM workshop in May 2001 [BC01a] and many other occasions. We want to thank all partici-

pants of these meetings for their feedback and appreciation. Also many thanks to CAIDA folks Evi Nemeth, Dan Anderson, Dan Plummer Marina Fomenkov, Brad Huffaker, Ken Keys and David Moore of CAIDA for their help in measurement and analysis tool development and helpful feedback.

References

[AJB00] R.Albert, J.Jeong, A.-L.Barabasi. Error and attack tolerance of complex networks. Nature, v.405, 27 July 2000, 378-381

[BC00a] A.Broido, kc claffy. Graphs That Make the Net Work. ISMA Winter 2000 Workshop, San Diego, Dec.2000 http://www.caida.org/outreach/isma/0012/agenda.xml.

[BC00b] A.Broido, kc claffy. The Internet's Core: Top IPs, Prefixes, and ASes. ISMA Winter 2000 Workshop, San Diego, December 2000.
http://www.caida.org/outreach/isma/0012/agenda.xml.

[BC01a] A.Broido, kc claffy. Complexity of global routing policies. Proceedings of Network-Related Data Management workshop, Santa Barbara, May 25, 2001, 18 p.

[BC01b] A.Broido, kc claffy. Internet topology, 30 pp, in preparation, 2001,
http://www.caida.org/outreach/papers/topologylocal/.

[BC01c] A.Broido, kc claffy. Internet topology: properties of IP graphs IEEE SPIE conference, Denver, Colorado, August 2001.

[BCh99] H.Burch, B.Cheswick. Mapping the Internet. IEEE Computer, 32(4), April 1999.

[BCN00] A.Broido, kc claffy, E.Nemeth. Packet arrivals on rate-limited links. CAIDA, 2000.
http://www.caida.org/~broido/coral/packarr.html

[BH00] S.Bodamer, J.Charzinsky. Evaluation of effective bandwidth schemes for self-similar traffic. ITC Specialist Seminar on IP Traffic Measurement, Modeling, and Management. Monterey, California, September 14, 2000

[BKMRRSTW00] A.Broder, R.Kumar, F.Maghoul, P.Raghavan, S.Rajagopalan, R.Stata, A.Tomkins, J.Wiener. Graph structure in the web. Comput.Netw,, 33, 2000. In Proceedings of The Ninth International World Wide Web Conference, Amsterdam, The Netherlands, May 2000. Available from http://www9.org/.

[CAIDA01] CAIDA's 'Packet Sizes and Sequencing' site,
http://www.caida.org/outreach/resources/learn/packetsizes

[CB00] B.Cheswick, H.Burch. Internet Mapping Project.
http://cm.bell-labs.com/who/ches/map

[CEAH00] R.Cohen, K.Erez, D.ben-Avraham, S.Havlin. Resilience of the Internet to random breakdowns. Physical Review Letters, 85 (21), Nov.2000.

[CJW01] H.Chang, S.Jamin, W.Willinger. Inferring AS-level topology from router-level path traces. Proceedings of SPIE workshop on Scalability and Traffic Control in IP Networks Denver, Aug.2001.

[CNSW00] D.S.Callaway, M.E.J.Newman, S.H.Strogatz, D.J.Watts. Network Robustness and Fragility: Percolation on Random Graphs. Phys.Rev.Lett.85 (25), Dec.2000

[Che1889] P.L.Chebyshev. Approximate expression for the square root of a variable through simple fractions. Zapiski Imp.Akad.Nauk, vol.LXI, 1, 1889. (Acta Mathematica, XVIII, 1894, p.113-132.) See also: Oeuvres, publ. de A.Markoff et N.Sonin. Chelsea Pub.Co., NY, 1961, vol.II.

[Downey01] A.B.Downey. The structural cause of file size distributions, to appear.

[Ext00] Extreme value distributions. In: Engineering statistics handbook, Ch.8. National Institute of Standards, 2000. http://www.itl.nist.gov/div898/handbook/apr/section1/apr163.htm

[Fa99] M.Faloutsos, P.Faloutsos, and C.Faloutsos. On power-law relationships of the Internet topology. ACM SIGCOMM, Cambridge, MA, Sept.1999

[FCHM01] M.Fomenkov, kc claffy, B.Huffaker and D.Moore, "Macroscopic Internet topology and performance measurements from the DNS root name servers", submitted. http://www.caida.org/outreach/papers/rssac2001a/

[Fomenkov00] M.Fomenkov. Internet's 'death rate'. CAIDA internal presentation, 2000.

[FS97] U.Frisch, D.Sornette. Extreme deviation and applications. J.Phys. I France, 7, 1155-1171, 1997.

[Gao00] On Inferring Automonous System Relationships in the Internet. IEEE Global Internet, Nov 2000. http://www-unix.ecs.umass.edu/ lgao/ globalinternet.ps

[Gr92] S.Gran. A course in Ocean Engineering. Developments in Marine Technology, Vol. 8. Elsevier, 1992. See also: A course in Ocean Engineering. Det Norske Veritas, 2001. http://www.dnv.com/ocean/course.htm

[GT00] R.Govindan, H.Tangmunarunkit. Heuristics for Internet map discovery. In Proceedings of IEEE Infocom, Tel Aviv, Israel, March 2000.

[Ha75] F.Harary. Graph Theory. Addison Wesley, 1975.

[HBCFKLM00] B.Huffaker, A.Broido, kc claffy, M.Fomenkov, K.Keys, E.Lagache, D.Moore, Skitter AS Internet Graph. Published by CAIDA. 1st ed: Apr.2000. 2nd ed: Oct.2000.

[HFMC01] B.Huffaker, M.Fomenkov, D.Moore, kc claffy. Macroscopic analyses of the infrastructure: measurement and visualization of Internet connectivity and performance. PAM 2001: A workshop on passive and active measurements on the Internet. Amsterdam, 23-24 Apr.2001.

[Hou01] G.Houston, 'Analyzing the Internet's BGP Routing Table', *The Internet Protocol Journal,* Mar.2001, 4(1) http://www.telstra.net/gih/papers/ipj/4-1-bgp.pdf

[Huffaker99] B.Huffaker, 'skdump' (a utility for reading skitter files.) Developed at CAIDA, 1999.

[Jac88] V.Jacobson, 'traceroute'. ftp://ftp.ee.lbl.gov/traceroute.tar.Z

[JK75] N.Johnson, S.Kotz. Distributions in statistics: Continuous univariate distributions. Wiley, 1975.

[Keys00] K.Keys, 'iffinder' (a tool for mapping interfaces to routers.) Developed at CAIDA, Sept.2000.

[Lang92] Serge Lang, Algebra, 3rd edition, Addison-Wesley, 1992.

[LS98] J.Laherrere, D Sornette (1998), "Stretched exponential distributions in Nature and Economy: 'Fat tails' with characteristic scales", European Phys.Journal, B2:525-539. http://xxx.lanl.gov/abs/cond-mat/9801293

[Merc00] Mercator, http://www.isi.edu/∼govindan/.

[Meyer01] Meyer, D., University of Oregon Route-Views Project, 2001. http://www.antc.uoregon.edu/route-views/

[Moore99] D.Moore, 'netgeo', IP geography server, developed at CAIDA, 1999.

[NLANR97] NLANR routing tables. http://moat.nlanr.net/Routing/rawdata

[Nor95] I.Norros. On the use of fractional Brownian motion in the theory of connectionless networks. IEEE Journal on Selected Areas in Communication, 13, No.6, Aug.1995, 953-962.

[NSSW00] C.Nuzman, I.Saniee, W.Sweldens, A.Weiss. A compound model for TCP connection arrivals. ITC Specialist Seminar on IP Traffic Measurement, Modeling, and Management. Monterey, California, Sept.14, 2000.

[PCH 2001] Sean McCreary, Bill Woodcock. PCH RouteViews archive. http://www.pch.net/documents/data/routing-tables

[PSFFG01] C.Palmer, G.Siganos, M.Faloutsos, C.Faloutsos, P.Gibbons. The connectivity and fault-tolerance of the Internet topology. Proceedings of the Workshop on Network-Related Data Management. Santa Barbara, May 2001.

[PG98]. J.-J.Pansiot, D.Grad. On routes and multicast trees in the Internet. ACM SIGCOMM Computer Communication Review, 28(1), Jan.1998

[PM01] Peacock maps. http://www.peacockmaps.com/index.html.

[PST99] G.Phillips, S.Shenker, H.Tangmunarunkit. Scaling of multicast trees: Comments on Chuang-Sirbu scaling law. Proc.of the ACM SIGCOMM, Sept.1999.

[RTYGSE00] P.Radolavov, H.Tangmunarunkit, H.Yu, R.Govidan, S.Shenker, D.Estrin. On characterizing network topologies and analyzing their impact on protocol design. Tech.Report 00-731, USC Computer Science Dept, 2000. http://www.isi.edu/ hongsuda/publication/

[RFC1771] Y.Rekhter, T.Li. A Border Gateway Protocol 4 (BGP-4) RFC 1771, March 1995. ftp://ftp.isi.edu/in-notes/rfc1771.txt

[RFC1812] F.Baker, ed. Requirements for IP version 4 routers. RFC 1812, 1995. http://www.faqs.org/rfcs/rfc1812.html.

[Skit98] Daniel McRobb, kc claffy, Skitter. CAIDA, 1998. http://www.caida.org/tools/measurement/skitter/.

[Skit01] CAIDA skitter monitor locations. http://www.caida.org/tools/measurement/skitter/monitors.xml.

[TGSE01] H.Tangmunarankit, R.Govindan, S.Shenker, D.Estrin. The impact of Routing Policy on Internet Paths. Proceedings of INFOCOM, Anchorage, AK, April, 2001.

[TGJSW01] H.Tangmunarunkit, R.Govindan, S.Jamin, S.Shenker, W.Willinger. Network topologies, power laws, and hierarchy. Tech.Report 01-746, Comp.Sci.Dept,USC, submitted for publication.

[TGS01] H.Tangmunarunkit, R.Govindan, S.Shenker. Internet path inflation due to policy routing. Proceedings of SPIE conference, Denver, CO, Aug.2001.

[Wessels00] D.Wessels. Squid chache logs. NLANR, 2000. ftp://ircache.nlanr.net/Traces/

[Willinger01] W.Willinger, private communication, March 2001.

[Woodcock01] Bill Woddcock, private communication, May 2001.