The Origin of Power Laws in Internet Topologies Revisited

Qian Chen Hyunseok Chang Ramesh Govindan Sugih Jamin Scott J. Shenker Walter Willinger

Abstract— **In a recent paper, Faloutsos et al. [1] found that the inter Autonomous System (AS) topology exhibits a power-law vertex degree distribution. This result was quite unexpected in the networking community and stirred significant interest in exploring the possible causes of this phenomenon. The work of Barabasi and Albert [2] and its application to network topology generation in the work of Medina et al. [3] have explored a promising class of models that yield strict power-law vertex degree distributions. In this paper, we re-examine the BGP measurements that form the basis for the results reported in [1]. We find that by their very nature (i.e., being strictly BGP-based), the data provides a very incomplete picture of Internet connectivity at the AS level. The AS connectivity maps constructed from this data (the *original* maps) typically miss 20–50% or even more of the physical links in AS maps constructed using additional sources (the *extended* maps). Subsequently, we find that while the vertex degree distributions resulting from the extended maps are heavy-tailed, they deviate significantly from a strict power law. Finally, we show that available historical data does not support the connectivity-based dynamics assumed in [2]. Together, our results suggest that the Internet topology at the AS level may well have developed over time following a very different set of growth processes than those proposed in [2].**

I. Introduction

Recent studies concerning Internet connectivity at the level of *Autonomous Systems* (ASs) have attracted considerable attention. For example, the empirically derived power-law relationships in the Internet's AS topology, originally due to Faloutsos et al. [1], suggest that the random and hierarchical graph models that have until recently been used to generate Internet-like topologies may not capture critical features or relevant structures inherent in actual networks. Barabasi and Albert in [2] show that such power-law graphs can arise from a simple dynamic model that combine incremental growth with a preference for new nodes to connect to existing ones that are already well connected (subsequently referred to as *the BA model*). Citing [1], Barabasi and Albert also state that this intuitively appealing growth model applies to the Internet's AS graph and therefore explain why AS graphs exhibit power-law vertex degree distributions. In this paper, we revisit the measurements used in the original discovery of power laws in AS topology [1]. In re-examining this data, we ask the following two questions: (1) Are the measurements used in [1] sufficiently complete to establish a strict power law for AS vertex degree distributions? and (2) How can the available measurements be used to establish the validity of Internet topology models at the AS level such as those proposed by Barabasi and Albert [2]?

The measurements that form the basis for the original power-law study [1] consist of BGP routing tables collected by the route server `route-views.oregon-ix.net` (henceforth, *the Oregon route server*) [4]. The use of this data for the purpose of studying the Internet's AS connectivity structure raises the following important issue. As a protocol, BGP is *not* about physical connectivity at the AS level; instead, BGP is *all* about AS peering relationships. Consequently, the AS connectivity structure gleaned from the Oregon route server data may provide a very incomplete picture of the physical connectivity that exist in the actual Internet. A more complete picture of Internet connectivity may potentially question or even invalidate some of the earlier findings. In Section II, we provide qualitative and quantitative insight into the degree of incompleteness of AS connectivity graphs constructed from BGP routing tables obtained solely from the Oregon route server. In particular, we observe that while the measured AS vertex degrees are clearly *highly variable* in the sense that they typically vary over 3–4 orders of magnitude, the vertex degree distributions resulting from more complete AS graphs do not conform to strict power-law distributions but are instead consistent with distributions where only the tails decay like a power.

In Section III we outline a general framework for validating the explanatory aspect of any proposed AS-level connectivity models. The discovery of an empirical phenomenon, such as the one reported in [1], is often followed by proposed explanations. Such proposed explanations may put forth a dynamical model that identifies a set of more elementary mechanisms as the main cause of the said phenomenon. A critical feature of the validation framework outlined in Section III is that it "closes the loop" between the discovery process and the proposed explanatory model. This "closing of the loop" is achieved by requiring that the proposed model also conforms to measured data at the level where the more elementary mechanisms are verifiable. In this particular case, we want to verify the network growth model put forth in [2] to explain the discovery of power law in AS vertex degree distribution reported in [1].[1] When conformance is verified, the network model in question can be expected to provide considerable new insights into the way networks organize themselves. The model can then be exploited for various engineering purposes. On the other hand, if the verification fails, the proposed model can not be the explanation of the observed phenomenon. Even if the proposed model were able to re-produce said phenomenon, it would have failed the verification if the associated elementary mechanisms turn out to be inconsistent with the measured data.

By performing a careful analysis of a collection of historical AS maps, we show in Section III that measured data are largely *not* consistent with the elementary mechanisms proposed in the BA model (described in Section III-A). As such, the BA model does not provide a valid explanation for the empirically derived power law for the AS vertex degrees. Recent generalizations of the BA modeling approach allow for vertex degree distributions that are more flexible than strict power-law distributions [5]. Despite this recent generalizations, and in light of our findings regarding AS vertex degree distributions that are not strictly power-law, we argue in Section IV that: (1) our general validation framework remains applicable, and (2) the connectivity evolution rules underlying the generalized BA modeling approach still do not conform to the data.

The BA model attempts to explain the highly variable vertex degree distribution of the AS topology through the detailed dynamics of how connections between ASs are established. Our results and observations suggest that this approach may be doomed. It is more likely that the Internet's AS topology has evolved over time according to a very different set of mechanisms or growth processes. We discuss the implications of our findings, as well as possible alternative approaches for explaining the high variability phenomena associated with the Internet's AS topology in Section IV.

II. On the Completeness of BGP-derived AS Maps

A number of recent studies characterize AS-level topology of the Internet by exploiting connectivity information contained in BGP routing tables. These studies, including [1], [2], mainly obtain their data from the Oregon route server. The Oregon route server connects to several operational routers within the Internet solely for the purpose of collecting BGP routing tables. From Nov. 1997 to Mar. 2001, these routing tables have been archived on a daily basis by the National Laboratory for Applied Network Research (NLANR) [6]. Presently, archives of the Oregon routing tables are available from sites such as the Packet Clearing House (PCH) [7] (starting from Feb. 2001) and `routeviews.org` [8] (starting from Apr. 2001). By making these data sets available to the public, both the Oregon route server and the archival sites are providing invaluable service to the research community. The question we explore in this section is, "How complete is the AS-level topology captured by the Oregon route server?"

If the actual Internet AS-level topology were known, the completeness of a topology captured by the Oregon route server could be checked by comparing it with the actual topology. Since the actual topology is not known, how do we check the completeness of the topology inferred from the Oregon route server data? Our initial study investigating this question is reported in [9]. In this section, we first summarize the methodology and main results of [9], then we present additional results not reported in [9] that further support our conclusion that AS maps constructed solely from Oregon route server data contain only a portion of AS connectivities on the Internet.

[1] Note that in this original context, the available data sets consist only of measurements collected by the Oregon route server. Further, the validity of the strict power-law relationship for the AS vertex degrees is taken for granted.

TABLE I
41 BGP Views

AS Name (AS#)	Description	Router Location
GTE (1)	Backbone	MA, US
STARTAP (10764)	Research network	IL, US
Telstra (1221)	ISP	Australia
VTX (12350)	ISP	Switzerland
SprintLink (1239)	Backbone	VA, US
C&W-Europe (12541)	Backbone	VA, US
INIT7 (13030)	ISP	Switzerland
LAN (15600)	ISP	Switzerland
Telstra USA (16779)	ISP	CA, US
CERFnet (1740)	Backbone	CA, US
Ebone (1755)	Backbone	Sweden
AT&T-GNS (2685)	Backbone	NY, US
XO Comm. (2828)	ISP	CA, US
KPNQwest (286)	Backbone	Europe
ESnet (293)	Research network	CA, US
Tiscali (3257)	ISP	Germany
RIPE NCC (3333)	Internet Registry	Netherlands
C&W (3561)	Backbone	VA, US
Exodus (3967)	ISP	CA, US
Global Online (4197)	ISP	Japan
Globix Corp. (4513)	ISP	OH, US
NETINS Inc. (5056)	ISP	IA, US
JIPPII (5409)	ISP	Germany
Broadwing (6395)	ISP	TX, US
Teleglobe Canada Inc. (6453)	ISP	QC, CA
Abovenet (6461)	ISP	CA, US
SCIFI (6667)	Backbone	Finland
C&W-Switzerland (6893)	Backbone	Switzerland
UUNET (701)	Backbone	NJ, US
AT&T WorldNet (7018)	Backbone	NJ, US
Zocalo (715)	ISP	CA, US
Colt (8220)	Backbone	UK
IBS (8271)	ISP	Switzerland
Teleglobe Europe (8297)	ISP	QC, CA
IXPRIME (8327)	ISP	Switzerland
Dolphins (8758)	ISP	Switzerland
AGRI (8843)	ISP	Switzerland
Carrier1 (8918)	ISP	Europe
ECS-IP (8938)	Backbone	Switzerland
SolNet (9044)	ISP	Switzerland
NEXTRANET (9177)	ISP	Switzerland

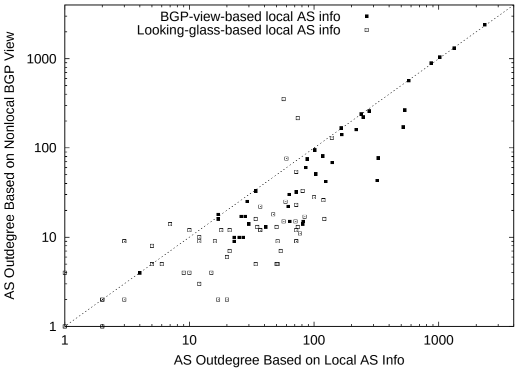


Fig. 1. The completeness of non-local BGP views.

While we do not have a complete map of the Internet at the AS level, we can re-construct the neighbor(s) each AS is connected to (i.e., its connectivity map) if we have access to its BGP routing table. We call the connectivity map of an ASi so constructed the *local* view of ASi. Table I lists the 41 ASs for which we have access to their BGP routing tables. A BGP routing table contains the list of all destination address ranges accessible to an AS. For each destination address range, the path vector (called *AS_PATH*) from the AS to the destination is also enumerated. Hence ASi can infer some portion of ASj's connectivities by observing its *AS_PATH*s that traverses ASj. We call the connectivities of ASj so inferred by ASi the *non-local* view of ASj. Next we ask, if we try to infer the connectivities of one of the 41 ASs (say, ASj) from the BGP routing tables of the other 40 ASs, i.e., by merging the 40 non-local views of ASj, how complete would the inferred connectivities of ASj be compared to its own local view?[2]

In addition to the above 41 ASs from which we are able to obtain their full BGP routing tables, we are able to collect summary BGP peering relationship from 70 different ASs that maintain *Looking Glass* sites. *Looking Glass* sites are maintained by individual ISPs to help troubleshoot Internet-wide routing problems. Since the BGP information obtained from *Looking Glass* sites do not include *AS_PATH* information, we can only use it to construct local AS views, which we then compared against non-local views constructed from the 41 ASs for which we have full BGP tables.

Figure 1 plots the number of AS connectivities as seen from the AS's local view (on the x-axis) against that seen from the aggregated non-local views (y-axis). In the case of the black dots, the local views are obtained from the local AS's full BGP routing table. In the case of the white dots, the local views are obtained from the local AS's *Looking Glass* information.[3] From this data, we conclude in [9] that the Internet maintains a much richer connectivity than can be observed by aggregating a handful of BGP routing tables.

A. Perusing the Internet Routing Registry

Our results in the previous section suggest that to construct a complete AS map, we must obtain full BGP routing tables from all existing ASs. Unfortunately, for security and commercial reasons, not all ASs are willing to reveal their full BGP routing tables. We next try to construct AS connectivity maps by perusing the data available in the Internet Routing Registry (IRR) [10]. The IRR maintains individual ISP's (Internet Service Provider) routing information in several public repositories to coordinate global routing policy. We study two of the biggest IRR databases in this paper, the ones maintained by the Routing Arbiter Project (RADB) and by Réseaux IP Européens (RIPE) (see [11]). IRR's routing policy database expresses routing information at various levels (e.g., individual address prefixes or ASs, etc.). We illustrate how the routing information of the IRR is expressed in Routing Policy Specification Language (RPSL) with the following two hypothetical database records.

```
route:    1.2.3.0/24
desc:     Foo.com
origin:   AS1
changed:  admin@foo.com 20010313
source:   RADB
```

```
aut-num: AS1
as-name: FOO-ASN
desc:    Foo Primary AS
import:  from AS2 action pref=100;
         from AS3 action pref=200;
         accept AS4
export:  to AS2 announce AS4
         to AS3 announce ANY
changed: admin@foo.com 20010313
source:  RADB
```

The left record states that on Mar. 13th, 2001, address range "1.2.3.0/24" belongs to AS1. The right record, which expresses AS1's routing policies, indicates that AS1 has two peering neighbors AS2 and AS3 with which it exchanges route reachability information of AS4. From this policy specification, we can infer that AS1 has AS2 and AS3 as neighbors.

Traditionally, publication of an ISP's routing policies in the IRR has been voluntary. However, many European exchange points [12], [13], [14], [15] now require members to register their routes and peering policies in the RIPE database. Other non-European ISPs have also relied more on the IRR to filter route announcement at border routers [16]. Comparing the RADB and RIPE databases, we found that out of the 2,673 ASs registered with RADB as of May 25th, 2001, 2,039 (76.3%) publish their routing policy;

[2]We do not claim that the local views themselves are complete. We simply want to know how complete are the non-local views relative to the local views.

[3]Individual *Looking Glass* sites may not list all of an AS's existing neighbors, hence the number of connectivities inferred represent only the *minimum* number of actual neighbors. This explains why we can see white dots lying above the diagonal line in Figure 1.

TABLE II
VALIDITY ANALYSIS OF RIPE DATABASE

	# of records (ASs)
All records	4,203
− void	3,917
− void − obsolete	1,582
− void − obsolete − incomplete	1,026

whereas in the RIPE database, 4,203 ASs (93.6%) out of 4,492 registered ASs publish their routing policy.

The IRR records are manually entered and updated; we present some data in [9] indicating that RIPE records are kept up-to-date more frequently than RADB records. Based on the relative completeness and freshness of the RIPE database, we decide to use it in our study. In order to avoid including any incorrect or outdated information from the RIPE database, we apply the following checks to individual database records[4]:

Void records. We consider an IRR record *void* if the AS described by the record was once present in the Oregon BGP tables (dating from Nov. 1997) but has disappeared from the tables at a certain point in time.[5]

Obsolete records. To find *obsolete* records, we first construct an *AS reference graph* from the registry records. The AS reference graph is a directed graph where each node corresponds to a registered AS and an edge corresponds to published peering relationship between an AS and one neighbor. An edge is directed from node A to node B if the registry record of AS A specifies AS B as a peering neighbor. If all the records were up-to-date, all edges on the graph must be bidirectional. A unidirectional edge indicates at least one of the two incident ASs has outdated information in database. If the update time of the two incident ASs of a unidirectional edge is more than 1 month apart, we consider the record updated earlier *obsolete*.

Incomplete records. We consider a record *incomplete* if the AS described has a neighbor in the Oregon BGP routing tables that was not registered with IRR. To detect unregistered neighbors, we scan the daily BGP routing tables starting from the date of the record's last update to the end of our data set (May 25, 2001). If an unregistered neighbor of an AS is found, the AS's record in the IRR is considered incomplete.

Only records that are not void, not obsolete, and not incomplete are considered *valid* records. Table II shows that only 1,026 ASs (about 9% of all known ASs) have *valid* records in the RIPE database. Nevertheless, as Table III shows, perusing the IRR database allows us to identify an extra 4,860 edges (or about 17.42%) over the most complete AS map constructed from all the BGP information we can obtain (compare the last and second-to-last rows of the table). The table shows the number of nodes (ASs) and edges contained in the AS map constructed from the various sources, cumulatively. The first row, labeled "Oregon BGP dump" lists the number of nodes and edges found in the AS map constructed *solely* from Oregon BGP routing tables. The second row ("+ RSs") lists the number of nodes and edges found in *all* publicly available full BGP routing tables (listed in [17]). In essence, this row represents the most complete AS map one can construct *solely* from publicly available BGP routing tables. The AS map reported in the third row was constructed from the AS map in the second row plus the *Looking Glass* (LG) data. Finally, the AS map reported in last row includes the data from the RIPE database. The "%inc" numbers in parentheses denote the percentage of increase in number of nodes and edges with respect to the Oregon-based AS map of the first row.

[4] These checks are more stringent than the ones we used in [9].

[5] To prevent artifacts caused by the finite time frame of our data, an AS is considered void only if its disappearance dated more than three months from the date of the last entry of our data set (May 25, 2001).

TABLE III
AS GRAPH STATISTICS

Source	# of nodes (%inc)	# of edges (%inc)
Oregon BGP dump	11,174	23,409
+ RSs	11,268 (0.84%)	26,324 (12.5%)
+ RSs + LG	11,320 (1.3%)	27,899 (19.2%)
+ RSs + LG + RIPE	11,456 (2.5%)	32,759 (40.0%)

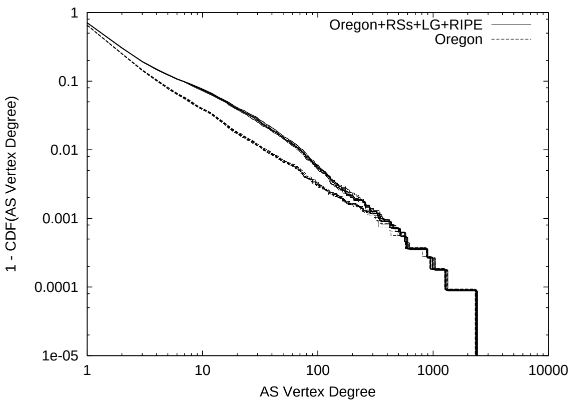


Fig. 2. Frequency distribution of AS vertex degree over time.

B. AS Vertex Degree Distribution Revisited

We currently have 9 instances of the *Looking Glass* and RIPE data sets. These were collected once a week, on the same day of the week, for 9 consecutive weeks starting Mar. 2001. In Figure 2 we plot the complementary distribution functions $F^c(x) = 1 - F(x)$, where $F(x)$ is the cumulative distribution function of the AS vertex degree corresponding to one of these 9 data sets. All 9 graphs lie very close to each other and form the upper, curved line in the figure (labeled once as "Oregon+RSs+LG+RIPE"). For comparison, we also plot the $F^c(x)$'s corresponding to the 9 AS maps constructed *solely* from Oregon BGP routing tables. The BGP routing tables were obtained at times corresponding to the time we collected the LG and RIPE data sets. These 9 graphs also lie very close to each other and form the lower, straight line of the figure (labeled once as "Oregon"). As is clear from the figure, the AS vertex degree distribution of the "Oregon" data sets agrees with the strict power-law curve reported in [1]. However, the more complete, though not necessarily complete, AS maps constructed from sources beyond the "Oregon" data set show more ASs with vertex degrees ranging from 4 to 300, resulting in a curved line in the distribution. Nevertheless, the distribution is certainly *heavy-tailed* or *highly-variable* in the sense that the observed vertex degrees typically range over three or four orders of magnitude; furthermore, the *tail* of the distribution may fit a power law.

III. BA-like Models and AS Maps

Several recent papers in the physics and biology literature have attempted to uncover the mechanisms that cause massive graphs such as Web linkage [18], telephone call [19], or bibliographic citation [20] graphs to exhibit phenomena similar to the power-law vertex degree distribution observed in the AS map constructed from the Oregon data set. Such works include the papers [21], [2], [22], [23], [24], [20]. Of these works, [2], [23] have attracted the most attention in the networking community as their authors propose a very appealing construction of network topologies. This construction was later used to form the basis for the claimed error intolerance and attack vulnerability of the Internet [23]. It also underlies the the network topology generator described in [3].

In this section, we explore the question, "Do BA-like models explain why AS maps have highly variable degree distributions?" After a short description of the BA model, we first revisit the context in which it was proposed (as far as the applicability to the Internet is concerned), namely for AS maps that exhibit strict power laws in their vertex degree distributions and have been re-constructed using only the Oregon data sets. By applying a generally applicable *validation framework* for Internet models, we demonstrate that while the BA model is able to produce power-law degree distributions, it does not explain why this type of distribution arises in the AS context. We argue later in Section IV that the same conclusion still holds when the strict power-law assumption on the degree distribution is replaced by the high variability observed on more complete AS maps.

A. The Barabasi-Albert (BA) Model

The Barabasi-Albert (BA) model [2], [22] consists of three generic mechanisms that drive the evolution of graph structures over time to produce graphs with power-law vertex degree:

1. **Incremental growth.** Incremental growth follows from the observation that most networks develop over time by adding new nodes and new links to the existing graph structure.
2. **Preferential connectivity.** Preferential connectivity expresses the frequently encountered phenomenon that there is higher probability for a new or existing node to connect or reconnect to a node that already has a large number of links (i.e., high vertex degree) than there is to (re)connect to a low-degree vertex.
3. **Re-wiring.** Re-wiring allows for some additional flexibility in the formation of networks by removing links connected to certain nodes and replacing them by new links in a way that effectively amounts to a local type of reshuffling connections.

Thus, starting with some initial graph structure, at every step during the evolution of the proposed BA model, each of following local events has some probability of taking place. The first event consists of adding a single new vertex, together with m new edges that connect the node to the existing graph, in agreement with the preferential connectivity assumption. The second event consists of adding m new links, independent of the new node addition above, by randomly selecting m preexisting vertices with uniform probability as origin nodes and connecting to m preexisting destination nodes following the preferential connectivity rule. Finally, the third event consists of re-wiring m existing links by random uniform selection of m vertices, and for each of them, removing a given link and reconnecting to a different node in agreement with the preferential connectivity property. Evolving according to this algorithm, the authors of [2], [22] showed that the resulting graph attains a steady state, where, for example, the distribution of the node degree remains unchanged over time and follows a power law with an exponent that is a function of the input parameters.

B. A Framework for Validating the BA Model

To apply the BA model to the AS map, we first map the elements and processes of the model to elements and processes observed on the AS map. The BA model constructs a graph structure that grows over time. Given an initial topology of the graph at time t_0, its topology at some later time t_i is determined by four primitive events: node birth, node death, link birth, and link death. Each node is born with some number of links; over time, a node gains some links and loses some links, and finally some nodes leave the graph. In the AS map case, a "node" is an "AS" and a "link" is a direct peering relationship between two ASs.

An AS is "born" when a new Internet Service Provider (ISP) or a large institution with multiple stub networks joins the Internet. When an AS is born, a number is assigned to the AS, and BGP routers begin to see the new AS number in their routing tables. A link is "born" when an existing AS decides to increase (or change) its connectivity by peering with another AS. When a link is born, a new path connecting the ASs incident to the new link appears in the routing tables of BGP routers.

By "validating the BA model," we mean to determine empirically whether the characteristics of birth/death processes in the AS map agrees with the elementary mechanisms underlying the BA model (i.e., incremental growth, preferential connectivity, and re-wiring). To that end, we first need to describe the data set we use in our empirical study.

C. Identifying the Births and Deaths of ASs and Links

In most studies involving the Oregon data sets, analyses are done on snapshots of the AS map taken at various points in time. To validate the BA model, however, requires the study of the AS map *as it evolves through time.* Since the Oregon data set is the only archive of daily snapshots of the AS map dating back to Nov. 1997, this is the only data set available to us for this study. To validate the BA model requires the identification of births and deaths of ASs and links. One immediate concern we face in doing this identification is to differentiate actual birth/death events from artifacts of the data collection process. As previously described, the Oregon route server connects to several operational routers within the Internet to collect their routing tables. Section II shows that the connectivities of an AS are only partially captured by the BGP routing tables of other ASs. We found that the set of operational routers from which the Oregon route server collects routing tables changes over time, both in membership and in number. When a link between two ASs observed on an earlier snapshot of the AS map disappears from later snapshots, it could be caused either by actual termination of the peering relationship between the two ASs or simply by a change in the Oregon route server's set of peer routers. Our first task is thus to determine the effect such ambiguities may have on our study.

Between Nov. 1997 and Nov. 2000, the Oregon route servers have peered with a total of 51 operational routers, with a maximum of 27 and minimum of 11 at any one time. Of these, only 8 have maintained steady relationships between Nov. 1998 to Nov. 2000 (Table IV). Table V lists the coverage of the AS map constructed from the routing tables of these 8 steady peers (limited AS map) against that constructed from the routing tables of all peers (full AS map). All the analyses reported in this paper have been done on both the limited and full AS maps. We did not find qualitatively significant differences in the results. This ensures us that the ambiguity in identifying actual births/deaths caused by artifacts of the data collection process does not have a detrimental effect on the results of our analysis. For the rest of the paper, we present only data obtained from the limited AS maps dating from Nov. 1998 to Nov. 2000. We will refer to this set of limited AS maps as "our data set."

TABLE IV
EIGHT PEERS OF THE OREGON ROUTE SERVER

IP Address	AS#	Description	Geographic Location
4.0.0.2	1	BBNPlanet	Palo Alto, CA
12.127.0.249	7018	ATT	Chicago, IL
192.121.154.25	1755	EBONE	Europe
192.41.177.192	2548	DIGEX	MAE-EAST
193.0.0.56	3333	RIPE NCC	Amsterdam, NL
195.66.225.254	5459	LINX	London, UK
202.232.1.8	2497	IIJ	Japan
204.70.4.89	3561	C&W USA	San Francisco, CA

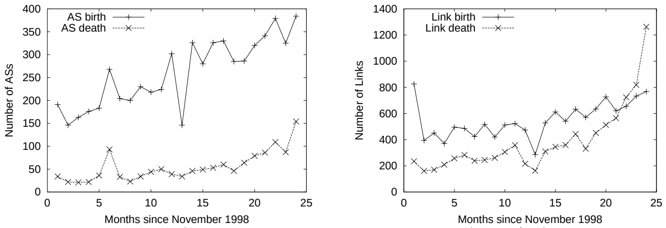
TABLE V
COVERAGE OF AS MAP FROM THE 8 STEADY PEERS

Peer	November 1998		November 2000	
	#ASs (%)	#Links (%)	#ASs (%)	#Links (%)
All	4,320 (100)	7,877 (100)	9,592 (100)	19,035 (100)
8 Combined	4,292 (99.4)	7,298 (92.7)	9,536 (99.4)	17,291 (90.8)

We mention two other caveats regarding our data set. First, we found that some of the BGP routing tables collected by NLANR were truncated prematurely, for example, the routing tables from Dec. 1999 are mostly less than half the expected size. We exclude all such truncated tables from our study. Second, we found instances in our data set where an AS (or link) disappears for a period of time. The authors of [25], [26] reported frequent losses of connectivity/reachability on the Internet due partly to routing instabilities caused by BGP implementation bugs, hardware glitches, and human errors. These *outages* can last anywhere from several minutes, to several hours, to longer than a week [27]. Outages are not modeled as contributing the AS map evolution and growth in the BA model. Hence we discount outages from our study of the birth/death processes, i.e., we assume that ASs/links are present even during their outages. In our data set, about one third of ASs and two thirds of links experienced at least one outage.

D. On the Incremental Growth Assumption

More formally, we define a birth to be an *actual* birth if we have not seen the born AS or link in our data set prior to the birth event. Similarly, a death is an *actual* death if the dead AS/link never again appears in our data set subsequent to the death event. To account for potential outages that started prior to the start of our data set or extend beyond the end of our data set, we discount all birth events



(a) ASs (b) Links

Fig. 3. Number of monthly births and deaths.

TABLE VI

VERTEX DEGREE DISTRIBUTION OF NEW OR DEAD ASs.

New	Degree	1	2	3	4	5	6	7	10	11	14
ASs	Frequency	5591	816	23	4	1	1	1	1	1	1
Dead	Degree	1	2	3	4	5	6	7	9	12	48
ASs	Frequency	1184	204	22	6	4	1	1	1	1	1

in the first C days and all death events in the last C days of our data set. We experimented with C values of 10 and 30 and did not see any significant differences in the results reported below. We decided to use $C = 10$.

Fig. 3 shows the number of monthly AS and link births and deaths in our data set.[6] From this figure we note that even after accounting for dead ASs and links, which should be included in an Internet AS topology model, the Internet AS graph does indeed grow incrementally by the addition of new nodes and links (though there were more link deaths than births in the last few months of our data set). Tables VI lists the vertex degree distribution of the new and dead ASs encountered in our data set. We note that a vast majority of new ASs are born with vertex degree 1 or 2. The same is true for dead ASs. Hence even though the data concerning the evolution of the AS map conforms to the incremental growth assumption, the choice of vertex degrees of the new ASs follows a distribution that heavily favors degrees 1 and 2, instead of the fixed number m assumed in the BA model [2]. Since the BA model does not include death events, and due to space constraints, we will present only birth events data in the remainder of the paper.

E. On the Preferential Connectivity Assumption

According to the preferential connectivity mechanism underlying the BA model, when a new node joins the network, the probability of its connecting to an existing node (henceforth, "target node" or "peer") follows a *linear preferential model*, $k_i / \sum k_j$, where k_i is the vertex degree of the target node and $\sum k_j$ is the total vertex degrees of all nodes in the graph before the addition of the new node. To check the practical relevance of this mechanism, we start with a simple but illustrative experiment. This experiment allows for a qualitative comparison between the actual connectivity preferences seen in the Internet and those assumed by the BA model.

AS Birth. Starting with the AS map of Nov. 1998, consider the next AS (node u) that joins the network. Node u joins the network with initial vertex degree m_u. Before we actually let node u join the network, we simulate the addition of node u with target AS(s) selected using the linear preferential model. We record the vertex degrees of the m_u target ASs so chosen, and label them $\widehat{k_i^u}, 1 \leq i \leq m_u$. Next we actually add node u to the network, record the vertex degrees of its actual target ASs, and label them $k_i^u, 1 \leq i \leq m_u$. We repeat the above process for 1,000 nodes added to the Internet between Nov. 1998 and May 1999. Fig. 4 plots the k_i^u's of the 1,000 nodes, and Fig. 5 plots the corresponding $\widehat{k_i^u}$'s. Note that both figures use log-scale on the y-axis.

From these figures, it is clear that in the real Internet, new ASs have a much stronger preference to connect to high vertex degree ASs than predicted by the linear preferential model (compare the density of points at the extreme ranges on the y-axis of the two figures). It may seem counter intuitive that the linear preferential model should prefer nodes with small vertex degrees. After all, by the linear preferential model, the probability of connecting to nodes with large vertex degrees is proportionally higher than the probability of connecting to nodes with small vertex degrees. There are, however, a much larger number of nodes with vertex degrees 1 and 2 than there are nodes with large vertex degrees (see Fig. 2). This explains why, despite the higher probability of connecting to nodes with large vertex degrees, we see a predominance of target nodes with small vertex degrees under the linear preferential model.

To turn the qualitative insight gained from Figs. 4 and 5 into a quantitative statement, we need a metric for analyzing the evolution of connectivity graphs. Here we encounter a problem: the transition probabilities that describe how the graph evolves from one state to another are themselves dependent on the graph's current state [19].[7] To illustrate, when a node joins the network, what is the probability that it will connect to a node with vertex degree K? The answer depends on how "big" K is at the time the new node joins relative to the other nodes. Also, once a new node connects to an existing node of vertex degree K, it increases the vertex degree of the target node, further complicating the analysis. Such changing AS vertex degree conditions are reflected in Figs. 4 and 5 in terms

[6]The link birth and death numbers do not include links births/deaths associated with AS births/deaths.

[7]To use Markov-model to model AS graph evolution would require a multi-stage Markov chain for each of the $>$ 11,000 ASs on the Internet.

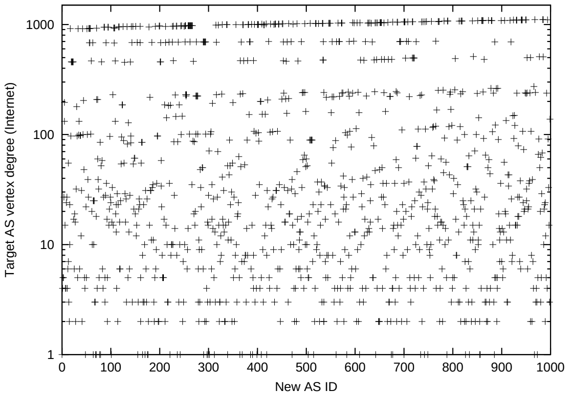


Fig. 4. Actual target node's vertex degree(s).

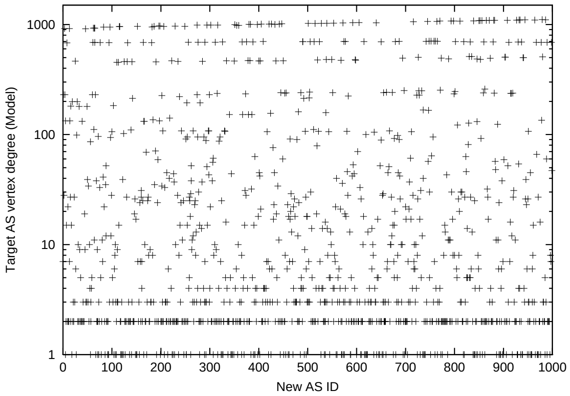


Fig. 5. Target node's vertex degree(s) by the linear preferential model.

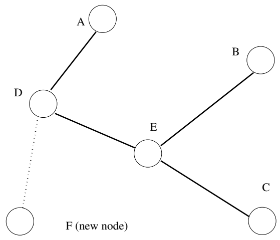


Fig. 6. A sample AS map.

of the slightly upward slanted "lines" that are more apparent for higher degrees, i.e., as the network gets bigger, the vertex degrees of the high rank nodes increases over time.

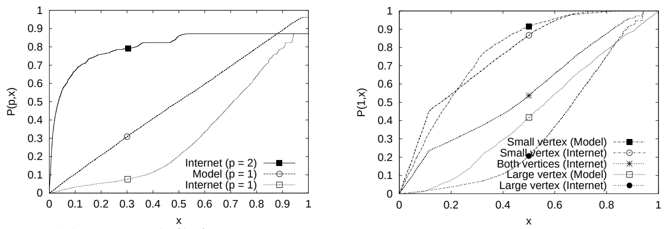
To deal with this problem more formally, we adopt in this paper the following metric for studying and analyzing connectivity preferences over time. For each node i, we calculate its *degree ratio* $C_i = k_i/\sum_j k_j$. Next, we sort all existing nodes of a given graph in a monotonically increasing order by their vertex degrees (hence the corresponding C_i). When a new node u joins a graph, for each target node v that node u connects to, we record target node v's *cumulative degree ratio* $X_v = \sum_{i:rank(i)\leq rank(v)} C_i$ (where $rank(i)$ is the rank of node i in the sorted array). We use the following example to illustrate the calculation of the cumulative degree ratio.

Assume there is a network consisting of 5 nodes: A, B, C, D and E and their connectivity graph is as shown in the Fig. 6. A new node F joins the network and connects to node D. The cumulative degree ratio of node D X_D can be calculated as follows: First, for each existing node i, calculate C_i. In this case, $C_A = \frac{1}{8}, C_B = \frac{1}{8}, C_C = \frac{1}{8}, C_D = \frac{2}{8}, andC_E = \frac{3}{8}$. Next, we sort all nodes according to their degrees in increasing order. For nodes with the same degree, we break ties arbitrarily (e.g., alpha numerically). The sorted result is A, B, C, D, E. Finally, we calculate target node D's cumulative degree ratio, $X_D = \sum_{i:rank(i)\leq rank(D)} C_i = C_A + C_B + C_C + C_D = \frac{5}{8}$.

We calculated the X_i of all 6,440 AS births in our data set. Next, we calculate the cumulative distribution of X_i. In other words, we ask, when a node u connects to its target node i, what is the value $P(x) = Prob[X_i < x]$, $0 \leq x \leq 1$? If a new node chooses its target node v based on linear preference, then $(X_v, P(X_v))$ must lie along the straight line $y = x$, for the following reasons. Recall the definition of $X_v = \sum_{i:rank(i)\leq rank(v)} \frac{k_i}{\sum_j k_j}$. By this definition, for node i, if $X_i \leq X_v$, node i must appear before node v, or is node v, in the sorted array. Thus, $P(X_v)$ is the sum of the probabilities that any node i such that $rank(i) \leq rank(v)$ is chosen as the target node. By the linear preferential model, each node i has the probability $k_i/\sum_j k_j$ to be chosen. Thus $P(X_v) = \sum_{i:rank(i)\leq rank(v)} \frac{k_i}{\sum_j k_j}$ and $X_v = P(X_v)$. We illustrate this by going back to the example involving Fig. 6. As previously computed, $X_D = \frac{5}{8}$. Then $P(X_D) = \sum_{i:rank(i)\leq rank(D)} \frac{k_i}{\sum_j k_j} = C_A + C_B + C_C + C_D = \frac{5}{8}$.

In view of Figs. 4 and 5, which show that new nodes are less likely to connect to low-degree target nodes than is predicted by the linear preferential model, we expect to see the empirical $P(X_v)$ vs. X_v plot to stay well below the diagonal. We can generalize the metrics above to also check for non-linear preference model, as follows: $C_i(p) = \frac{{k_i}^p}{\sum_j {k_j}^p}$, $X_v(p) = \sum_{i:rank(i)\leq rank(v)} C_i(p)$ and $P(p, x) = Prob[X_v(p) \leq x]$. (Hence $C_i(1) = C_i$, $X_v(1) = X_v$ and $P(1, x) = P(x)$.)

Fig. 7 (a) shows the metric $P(p, x)$ of the target nodes of the 6,440 ASs born in our data set, for p=1 and 2. The figure also includes a plot labeled "Model (p = 1)" of $P(1, x)$ of target nodes of a network that was grown purely by incremental growth and linear preferential connectivity to a size comparable to that of the Internet of Nov. 2000. (Note that the marks on the curves in the figure serve *only to*



(a) New ASs' peers (b) New links' vertices

Fig. 7. $P(p, x)$ distribution (p = 1 in (b)). There are 6,440 AS births in (a) and 13,558 link births in (b). The markers are only for visual identification of the curves, it does *not* mean that we have only a single data point for each curve.

help visual identification of the curves. We have 6,440 data points for each curve, not a single data point.) In agreement with our earlier results reported in Figs. 4 and 5, we see that the target nodes' $P(1, x)$ distribution of the Internet is more skewed towards large vertex ASs than predicted by the BA model. The $P(2, x)$ curve illustrates that simply changing the value of p does not make the data fit the BA model.

From the data presented so far, we conclude that while incremental growth by node addition does play a role in the evolution of the AS map, the vertex degree distribution and peering characteristics of new ASs are more complex than the simple models used in the BA construction.

Link Birth. In addition to connectivities (links) formed between a new AS and its target AS(s), new links can also be formed between existing ASs. We now characterize link births that are *not* associated with AS births. There are 13,558 such link births in our data set. In the BA model, a link birth is modeled by a two-step process: first a node is chosen with probability $1/N$, where N is the number of nodes in the graph, then a new link is added to connect this node and a peer i chosen according to the linear preferential model $C(1, i)$.

Following [28], we classify ASs into three categories: (1) provider ASs provide Internet access to other ASs, (2) customer ASs purchase Internet access from provider ASs, and (3) peer ASs agree to exchange traffic between themselves. Customer ASs tend to be smaller than provider ASs (in terms of vertex degree, number of routers, geographic area coverage, traffic carried, etc.) and peers tend to be of comparable sizes. An instance where we may see a new connection between existing ASs is when a customer purchase an extra access service, for fault tolerance or capacity reasons. Another instance could be that an AS grows in importance and becomes a peer to a clique of other ASs. In such cases, it is usually the AS with the smaller vertex degree that initiates a new connection to another AS of a larger vertex degree. Subsequently, in our characterization of link births, we group the two vertices associated with a newly born link into the "Small vertex" and the "Large vertex" groups.

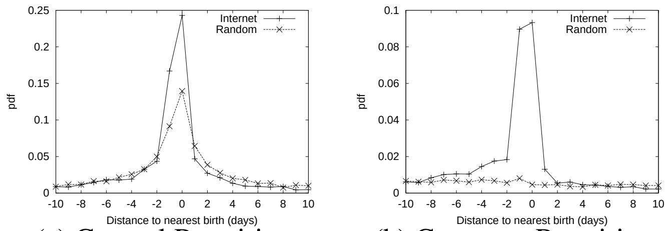
Fig. 7 (b) plots the $P(1, x)$ of the vertices of the 13,558 new links in our data set. The "Both vertices (Internet)" curve plots the $P(1, x)$ of both vertices as a single distribution. The "Small vertex (Internet)" curve plots the distribution of the smaller of the two vertices of the new links, and the "Large vertex (Internet)" curve plots the distribution of the larger of the two vertices. For comparison with the BA model, we conduct the following experiment: we add a single link to the Nov. 1998 AS map according to the BA model and record the vertex degrees of the smaller and larger of the two vertices the link connects. We repeat the experiment 1,000 times and plot the resulting two distributions in the same figure (the curves labeled "Small vertex (Model)" and "Large vertex (Model)"). We note that while the "Small vertex" distribution under the BA model may be considered similar to that of the Internet, the "Large vertex" distribution of the BA model is very different from that of the Internet.

F. *On the Relevance of the Re-wiring Mechanism*

In addition to incremental growth and preferential connectivity, the evolution of the graph structure in the BA model is driven by a third mechanism, namely *re-wiring*. Re-wiring under the BA model consists of uniformly selecting m nodes, removing one of their links, and connecting the node to another existing node according to the linear preferential connectivity model. In the Internet, a re-wiring could be approximated by a customer changing its choice of access provider.

The data on customers changing their ISPs is generally not available. Hence it is not practical to validate the BA re-wiring mechanism directly against empirical data. Instead, we approximate link re-wiring by the following heuristic: if a link birth happens within 10 days of a link death (before or after), and one of the two vertices of the dead link is a vertex of the new link, we consider the link rewired. We call this the *General Re-wiring* model. (Recall that by our definition, a link birth/death precludes link addition/removal associated with node birth/death.) If we restrict the definition of re-wiring to a customer AS changing provider, and assume that customer ASs usually have smaller vertex degree than provider ASs, we can have a narrower definition of re-wiring: given the two vertices of a link, a link is considered rewired only if the AS with the smaller vertex stays the same in both the link death and link birth events. We call this the *Customer Re-wiring* model.

Fig. 8 (a) shows the probability that among all the 10,260 link death events in our data set, one can find a link



(a) General Re-wiring (b) Customer Re-wiring

Fig. 8. Link Death/Birth Correlation in Two Re-wiring Models

birth correlated with the link death 10 days before (-10 on the x-axis) to 10 days after the death of a link when the *General Re-wiring* model is assumed. A correlated link birth occurring before a link death could be interpreted as a customer not relinquishing its current connection to the Internet until a new one is set up and stable. Next, in order to investigate the degree of link death/birth correlation under our Re-wiring models, we perform the following experiment. For every link death, we identify the *anchoring* AS, i.e., the AS incident to both the dead and new links. Independent of this link, the AS has a lifetime of its own. We re-position the death event uniformly along the lifetime of the AS (or the whole trace if the AS's lifetime is longer than the trace length); then we ask, how far away is the birth event with this AS as the anchor that is closest to the re-positioned death event? There is always at least one such birth event, namely the birth event originally associated with the death event before the re-positioning. For example, if we observe that the link between AS1 and AS2 dies on day 0 and a new link between AS1 and AS3 is born on day 10, we keep the birth on day 10 and randomly place the death event within AS1's lifetime. Once all repositions are done in such a way, we re-generate the death/birth correlation probability using the *modified* dataset (line labeled "Random" in Fig. 8).

We observe that in both figures, there is a noticeable correlation beyond random coincidences of the birth and death events only for births on the day or the day before the deaths. In both cases, not more than 20% of all link deaths can be meaningfully correlated with a nearby link birth, which strongly suggests that re-wiring may not be a significant factor in the evolution of the Internet AS topology.

IV. Conclusion: On the Need for Alternative Models

Recall that the original intent of the BA model (when applied to the Internet's AS graph) was to explain the empirically derived power law degree distribution of the AS topology. This explanation is in terms of the detailed yet simple dynamics of how connections between AS are established (i.e., incremental growth, preferential connectivity, and re wiring). However, we have demonstrated in Section II that power law degree distributions, while consistent with AS maps that rely solely on the Oregon data sets (the *original* maps, are not consistent with the *extended* maps that offer a more complete picture of the actual AS connections that exist in today's Internet. While the degree distributions of these extended AS maps are certainly *heavy-tailed* or *highly-variable* in the sense that the measured vertex degrees typically range over three to four orders of magnitude, only their tails can in general be expected to conform to a power law. In addition, we concluded in the previous section that as far as the Oregon-based AS maps are concerned, the detailed dynamics underlying the BA modeling approach does not explain the structure of the vertex degree distributions of the resulting AS maps.

We are fully aware of the possibility that the original BA models can be modified in a number of ways (see for example [5]). These modifications may not only produce AS maps with highly variable degree distributions (so as to model more realistic and complete AS connectivity maps), but also accommodate different types of preferential attachment rules (so as to provide a good fit with historical AS data), etc. However, any such resulting model would still seek to explain the highly variable degree distributions in terms of the detailed dynamics of network growth—just as the original BA model. Moreover, pursuing such modifications typically means sacrificing simplicity and parsimony for flexibility. Sacrificing parsimony in modeling is essentially self-defeating when aiming for truly explanatory topology models. A parsimonious model that can be simply understood has the highest potential of providing novel insights into how AS connectivity evolves over time and what basic mechanisms are responsible for shaping the structure of future topology. While highly parameterized BA-type models can be expected to fit a particular AS data set very well, it will be, in general, hopeless to give any physical meaning to all the parameters that come with such models.

Motivated by the observed discrepancy between actual AS growth (of the original maps) and the growth dynamics following the BA-model, and by the strong correlation between measured vertex degree and vertex "size" that suggests a possible economic-based connection [29], we are led to explore alternative modeling approaches that promise to retain the simplicity and parsimony of the original BA model but allow for greater flexibility otherwise. In particular, we look for ways to explain the observed highly variable vertex degree distributions of actual AS maps in terms of an alternative set of basic mechanisms for growing AS-type graph structures. To this end, if the results presented in Section III can teach us a lesson, these alternative mechanisms should ideally be flexible enough to account for non-generic and context-specific design details.

One such alternative approach that moves beyond the "design-free" BA framework concerns a recently proposed construction due to Carlson and Doyle [30], [31], called HOT (for *highly optimized tolerance*). HOT is a mechanism for creating power laws that focuses on systems which are optimized by engineering design, for example, and at the same time provide robust performance despite uncertain environments. This approach suggests that power laws in such systems are due to tradeoffs between yield, cost of resources, and tolerance to risk. It also suggests that these tradeoffs lead to highly optimized designs that perforce allow for a wide range of event sizes, in particular for occasional extreme sizes. More importantly, Carlson and Doyle show that characteristic features of HOT systems include (a) high efficiency, performance, and robustness to designed-for uncertainties; (b) hyper-sensitivity to design flaws and unanticipated perturbations; (c) non-generic, specialized, structured configurations; and (d) power laws. This newly proposed framework is in sharp contrast to SOC or *self-organized criticality*, which can only claim the last one of these four properties as its hallmark, but has nevertheless been widely viewed in the past as the origin for power laws in complex systems.[8] The Carlson-Doyle model demonstrates that by simply adding an element of "design" to SOC, the characteristics of the underlying systems completely change, but the power-law relationships are typically maintained. These properties make the Carlson-Doyle model very appealing from a networking perspective (for a first networking-related application of HOT related to Web layout design, see [33]), because it seems to overcome, at least in theory, some of the apparent shortcomings of the BA approach as reported in this paper. At the same time, while we have demonstrated here how to validate the BA framework against measured AS data, verifying the causes underlying the HOT mechanism against our data appears challenging, to say the least.

REFERENCES

[1] C. Faloutsos, P. Faloutsos, and M. Faloutsos, "On Power-Law Relationships of the Internet Topology," in *Proceedings of the ACM SIGCOMM*, Sept. 1999.

[2] Albert-László Barabási and Réka Albert, "Emergence of scaling in random networks," *Science*, pp. 509–512, Oct. 1999.

[3] Alberto Medina and Ibrahim Matta, "Brite: A flexible generator of internet topologies," Tech. Rep. BU-CS-TR-2000-005, Boston University, Boston, MA, 2000.

[4] Route Views, "University of Oregon Route Views Project," http://www.antc.uoregon.edu/route-views/.

[5] Réka Albert and Albert-László Barabási, "Statistical Mechanics of Complex Networks," Submitted for publication, 2001.

[6] National Laboratory for Applied Network Research (NLANR), "Route Views archive," http://moat.nlanr.net/Routing/rawdata.

[7] Packet Clearing House (PCH), "Route Views archive," http://www.pch.net/documents/data/routing-tables.

[8] Routeviews.org, "Route Views archive," http://archive.routeviews.org.

[9] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "On Inferring AS-Level Connectivity from BGP Routing Tables," Submitted for publication, June 2001.

[10] Merit Network, "Internet Routing Registry," http://www.irr.net.

[11] IRR, "List of Routing Registries," http://www.irr.net/docs/list.html.

[12] The London Internet Exchange (LINX), "Memorandum of understanding," http://www.linx.net/joining/mou.thtml.

[13] The CERN Internet eXchange Point (CIXP), "Technical operational environment and principles," http://wwwcs.cern.ch/public/services/cixp/cernixp.technical.html.

[14] The Internet Neutral Exchange (INEX), "Memorandum of understanding," http://www.inex.ie/inexmou.html.

[15] INX-Hamburg, "Peering policy," http://www.inx-hh.de/html/i_peering.html.

[16] K. Oberman and J. Haas, "Route registry: who uses them?," Posting to the NANOG mailing list, October 25th, 2000.

[17] Traceroute.org, "Public route server list," http://www.traceroute.org/.

[18] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tompkins, and J. Wiener, "The Bow-Tie Web," in *Proceedings of the 9th International World Wide Web Conference*, 2000.

[19] William Aiello, Fan Chung, and Linyuan Lu, "A random graph model for massive graphs," in *Proc. of the 32rd Annual ACM Symposium on Theory of Computing*, 2000, pp. 171–180.

[20] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, "Classes of Small-World Networks," *Proceedings of the National Academy of Sciences*, vol. 97, no. 21, October 2000.

[21] D. J. Watts and S. H. Strogatz, "Collective Dynamics of Small-World Networks," *Nature*, vol. 363, pp. 202–204, 1998.

[22] Réka Albert and Albert-László Barabási, "Topology of Evolving Networks: Local Events and Universality ," *Physical Review Letters*, vol. 85, pp. 5234–5237, 2000.

[23] Réka Albert, Hawoong Jeong, and Albert-László Barabási, "Error and attack tolerance of complex networks," *Science*, pp. 378–381, July 2000.

[24] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts, "Network robustness and fragility: Percolation on random graphs," *Physical Review Letters*, vol. 85, no. 25, pp. 5468–5471, December 2000.

[25] Vern Paxson, "End-to-end routing behavior in the internet," in *Proceedings of ACM SIGCOMM '96*, Aug. 1996, pp. 25–38.

[26] C. Labovitz, R. Malan, and F. Jahanian, "Internet routing instability," *Proceedings of ACM SIGCOMM*, September 1997.

[27] S. Donelan, "History: Lengthy outages," Posting to the NANOG mailing list, January 24th, 2001.

[28] L. Gao, "On Inferring Autonomous System Relationships in the Internet," in *Proc. IEEE Globecom*, San Francisco, CA, 2000.

[29] H. Tangmunarunkit et al., "Does size determine degree in as topology?," Submitted for publication, July 2001.

[8] In fact, the BA approach considered in this paper can be viewed as an incarnation of the general SOC framework [32].

[30] J. Doyle and J. M. Carlson, "Highly Optimized Tolerance: A Mechanism for Power-Laws in Designed Systems," *Physical Review Letters*, 1999.

[31] J. Doyle and J. M. Carlson, "Power Laws, Highly Optimized Tolerance and Generalized Source Coding," *Physical Review Letters*, 2000.

[32] P. Bak, *How Nature Works: The Science of Self-Organized Criticality*, Copernicus, June 1996.

[33] Xiaoyun Zhu, Jie Yu, and John Doyle, "Heavy tails, generalized coding, and optimal web layout," *IEEE infocom*, April 2001.