On Inferring AS-Level Connectivity from BGP Routing Tables

Hyunseok Chang
Department of EECS
University of Michigan
Ann Arbor, MI 48109-2122
hschang@eecs.umich.edu

Ramesh Govindan
Information Sciences Institute
University of Southern California
Los Angeles, CA 90292-6695
govindan@isi.edu

Sugih Jamin
Department of EECS
University of Michigan
Ann Arbor, MI 48109-2122
jamin@eecs.umich.edu

Scott J. Shenker
ACIRI/ICSI
1947 Center St., Suite 600
Berkeley, CA 94704-1198
shenker@aciri.org

Walter Willinger
AT&T Labs-Research
180 Park Ave.
Florham Park, NJ 07932-0971
walter@research.att.com

Abstract—**Recent studies concerning the Internet connectivity at the AS level have attracted considerable attention. These studies have exclusively relied on BGP measurements collected by the National Laboratory for Applied Network Research (NLANR) in deriving some of the unexpected and intriguing results. The NLANR data sets reflect AS peering relationships, as reported by BGP, seen from a handful of vantage points in the global Internet. The possibility that these data sets may provide only a very sketchy picture of the complete inter-AS connections that exist in the actual Internet has received surprisingly little scrutiny. By augmenting the NLANR data sets with BGP summary information from a large number of Internet *Looking Glass* sites and with routing policy information from Internet Routing Registry (IRR) databases, we find that (i) a significant number of existing AS connections remain hidden from most BGP routing tables, (ii) the AS connections to Tier-1 ASs are in general more easily observed than those to non-Tier-1 ASs, and (iii) there are at least about 25–50% more AS connections in the Internet than what recently considered BGP-derived AS maps contain (but only about 2% more ASs). These findings point out the need for an increased awareness of and a more critical attitude toward the applicability and completeness of given data sets at hand when establishing the generality of any particular observations about the Internet.**

Keywords—**Internet topology, BGP routing tables**

This project is funded in part by NSF grant number ANI-0082287 and by ONR grant number N000140110617. Sugih Jamin is further supported by the NSF CAREER Award ANI-9734145, the Presidential Early Career Award for Scientists and Engineers (PECASE) 1999, and the Alfred P. Sloan Foundation Research Fellowship. Additional funding is provided by AT&T Research, and by equipment grants from Sun Microsystems Inc. and Compaq Corp.

Correspondence to hschang@eecs.umich.edu.

I. INTRODUCTION

For the past two years, there has been a significant increase in research activities related to studying and modeling the Internet's topology, especially at the level of *autonomous systems* (ASs). To date, these activities consisted mainly of analyzing and modeling measurements to infer the Internet's AS connectivity graph and describe its properties [1], explaining the origins and causes of some of the observed surprising features [2], [3], and building generators that produce graph structures that match those of the measured AS connectivity graphs well (see for example [4], [5], [6]). A closer look at the measurements that form the basis for all these studies reveals that the data sets used consist of BGP routing tables collected by the route server `route-views.oregon-ix.net` [7]. This Oregon route server connects to several operational routers within the Internet to collect their routing tables. Starting in Nov. 1997, these routing tables have been archived on a daily basis by the National Laboratory for Applied Network Research (NLANR) [8]. Both the Oregon route server and NLANR have provided invaluable service to the research community in making these data sets available to the public. However, the use by researchers of these data sets for the purposes 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, BGP-derived AS connectivity may yield a very incomplete picture of the physical connectivity that exists in the actual Internet. We are not aware of any study that has addressed this completeness of BGP-derived AS maps. In particular, none of the above mentioned papers indicates awareness of this problem and of the potentially damaging effect it could have on some of the reported findings about the reported Internet's AS graph.

Our main objective in this paper is to shed light on the completeness of Internet AS maps reconstructed from BGP route advertisements. To this end, we augment the NLANR data with (i) full BGP table dumps from a dozen additional public route servers, (ii) a selection of Internet *Looking Glass* sites that provide BGP summary information, and (iii) the Internet Routing Registry (IRR). By processing the available BGP dumps, we end up with about 40 BGP viewers (see Section II for a definition of a BGP viewer and a description of the Looking Glass sites), all residing in different ASs, that allow us to explore the question of how well the peering relationships maintained by a certain AS ("local view") are observed by other ASs ("non-local view"). We observe that a significant number of existing AS connections are commonly hidden from most BGP routers. Considering that the ability to infer the existence of an actual AS connection from routing tables depends largely on existing AS relationships, we found that the AS connections to Tier-1 ASs are in general more easily observed than those to non-Tier-1 ASs (Section II-C). In short, the findings reaffirm our earlier comment that BGP is all about AS peering relationships and suggests that the actual Internet maintains a much richer connectivity at the AS level than has been previously reported.

To quantify to some degree the difference between the BGP-derived AS connectivities and the actual inter-AS connections in place in today's Internet, we also consult IRR databases that maintain individual ISP's routing policy information in several public repositories. The IRR's goal in maintaining these databases is to coordinate and facilitate the setting of global routing policy. The IRR repository at RIPE contains reasonably up-to-date entries. We find that AS graphs reconstructed from NLANR data, the Looking Glass sites, as well as RIPE information have typically about 25%–50% more edges (and about 2% more nodes) than their counterparts that rely solely on the NLANR data.

The implications of our findings are twofold. First, they clearly demonstrate the need to be more aware of and critical towards relying on any single data repository. Even when the data is by itself of the highest overall quality, its applicability and sufficiency must be evaluated in terms of the particular needs of any given study. Second, as far as published AS connectivity studies are concerned, our findings have practical as well as theoretical implications. For example, the finding reported in [1] that says measured AS graphs exhibit power law vertex degree distributions can be interpreted qualitatively to mean simply that these vertex degrees are highly variable, i.e., they typically vary by over three or so orders of magnitude. This qualitative interpretation is *not* disputed by our findings. However, our findings *are* in conflict with the power law claims in [1] when vertex degree distributions are expressed quantitatively, such as in $f_d \propto d^{-\alpha}(d = 1, 2, \ldots; \alpha > 0)$, relating the vertex degree d to its frequency. In fact, our findings state that the vertex degree distributions of more complete snapshots of the AS graph do not conform to power law distributions, but are instead consistent with the more flexible class of heavy-tailed distributions, where only the tails are assumed to decay as a power law. Clearly, this latter distinction has direct implications for the generation of AS-like graphs or for the more challenging question of explaining the origins and causes of heavy-tailed vertex degree distributions in the Internet context. To illustrate, the work by Barabasi and Albert [2], [3] takes the quantitative power law observations at face value and provides a suite of results, including constructions that attempt to explain the causes that lead to power law vertex degree distributions. The applicability of these results and constructions to the Internet has been claimed in [3] based on the power laws reported in [1] (see, however, [9]). Given our finding of the absence of a strict power law for the vertex degree distributions, such a theory may say little about actual Internet topology at the AS level.

The rest of the paper is structured as follows. In Section II, we introduce the notion of a representative BGP viewer and explore in detail how well peering relationships maintained by an individual AS are being observed by other ASs. To quantify the degree of incompleteness of BGP-derived AS maps, we include in Section III information from the IRR and use its RIPE database to obtain a more complete picture of the physical connections that exist between the different ASs (and of which BGP only sees a certain fraction). We conclude in Section IV by commenting on some of the lessons learned from pursuing this study of Internet connectivity at the AS level.

II. The Completeness of BGP-derived AS-level Topology

If the actual Internet AS-level topology were known, the completeness of a topology constructed from the NLANR data sets could be checked by comparing it with the actual topology. Since the actual Internet AS-level topology is not known, how do we go about checking the completeness of the topology inferred from NLANR data? The approach we have adopted in this paper is as follows. The BGP routing table obtained from an AS contains information about that AS's connectivities to other ASs. It also contains information on the connectivities between other

TABLE I
BGP DUMP FROM PUBLIC ROUTE SERVERS

Name	AS#	# distinct next hops	# neighbor ASs	Data size (MB)
NC1	-	43	35	291.5
NC2	-	42	16	131.3
C1	7018	24	1	145.8
C2	3967	199	279	127.0
C3	6539	7	1	57.1
C4	3549	3,089	447	49.5
C5	8709	19	187	47.4
C6	1740	3	1	26.0
C7	8220	42	331	20.3
C8	4197	12	82	18.3
C9	3257	1	1	8.0
C10	1	1,175	495	7.8

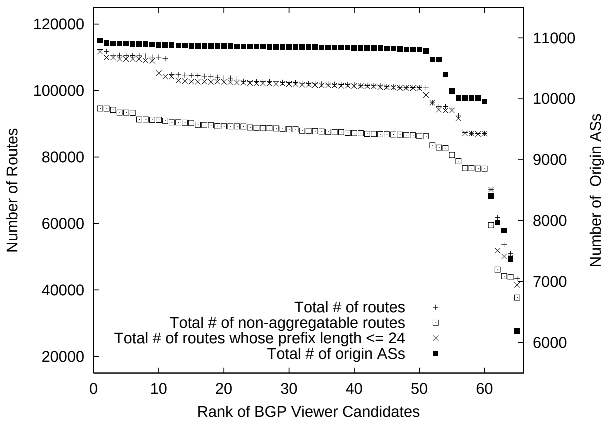


Fig. 1. The Number of Routes (Origin ASs) vs. Rank

ASs. Assume that the BGP routing table collected at an AS X contains the most complete vertex degree information obtainable of AS X.[1] The BGP routing table obtained from AS Y will have (see) some, but most likely not all, of the connectivities between AS X and other ASs. Similarly, the BGP routing table obtained from AS Z will see some but not all connectivities between AS X and other ASs. Adding together the observations from ASs Y and Z, we will likely get a more complete count of AS X's vertex degree, though probably by no means the complete count. Considering that the NLANR data set is the repository of the BGP routing tables obtained from several ASs, the question we ask in this section is, "How many BGP routing tables from different distinct ASs do we have to aggregate before we see the same vertex degree of AS X as reported by AS X's BGP routing table?" To answer that question, we must first collect BGP routing tables (*BGP views* henceforth) from several distinct ASs.

A. Available BGP Routing Tables

In addition to the Oregon route server, the Swiss Network Operators Group (SwiNOG) also provides access to a non-commercial route server that collects and makes publicly available BGP route table dumps [10]. We have also obtained address prefixes and AS path information from UUNET which forms one of our data set. As of April, 2001, ten commercial ISPs (Internet Service Providers), residing in different ASs, also allow public access to their route servers providing full BGP table dump. As a very first step of our study, we collected BGP routing tables from all these route servers. Due to the different nature of the BGP information available at Oregon and SwiNOG from that available at the commercial route servers, we will denote the Oregon and SwiNOG route servers the "collector" route servers, and call the others "operational" route servers. Table I lists the characteristics of the router servers. The column "# distinct next hops" lists the number of distinct next hop routers found in each BGP routing table, the column "# neighbor ASs" lists the number of distinct ASs that those routers reside in. A router interface that connects to a network access point (NAP) sees all other routers connected to the same NAP, hence in some cases we see a larger number of neighbor ASs than the number of distinct next hops in the table. Several entries in the table show routers seeing only a single neighbor AS. These are commercial routers that connect only to other, topologically distributed, routers residing in the same AS (belonging to the same commercial entity). In the table we also see a couple of routers with thousands of distinct next hops. Router server C10 turns out to be a "route-reflector" that receives BGP announcements from many internal BGP (iBGP) routers, each of which is connected to several external BGP (eBGP) routers, and the number of "distinct next hops" reported are the aggregate number of eBGP routers "connected" to it through iBGP routers. We suspect the same to be true for C4, however, we cannot confirm this and do not use C4 in our analysis. Finally, observe the variability in BGP route table sizes ("Data size"). In particular, note the lack of proportionality between BGP routing table sizes and the number of distinct next hop routers. These observations led us to do a more careful analysis of the data available from each source and to use in our analysis only sources from *BGP viewer*s satisfying the definition below.

B. BGP Viewer Selection

We define the union of all the address space reachable in all the BGP routing tables we have as the *known ad-*

[1] Due to policy routing, the BGP routing table exported by an AS may not have its complete vertex degree, e.g., private peering agreements may not be advertised. The assumption is not that we have the complete vertex degree but that we have the most complete vertex degree obtainable.

TABLE II
41 BGP VIEWERS

AS Name (AS#)	Description	Viewer 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

dress space. Next, we define a *BGP viewer* to be either an "operational" route server or a *peer* of a "collector" route server; the BGP routing table of a *BGP viewer* must cover the whole known address space. When a BGP router does not have a path to a given address space, it can sometimes forward all packets addressed to that address space to a default router. A BGP routing table is said to be *default-free* if it does not employ default routing. By requiring our *BGP viewer*s to cover the whole known address space, we stipulate that only routers with default-free routing tables can be used as our data source.

While the routing table of a peer of a "collector" route server contains the whole address space reachable through that peer, reflecting the address space reachable from the peer's AS, the address space reachable through an "operational" route server, and therefore the address space reachable from the AS in which the "operational" route server resides, must be constructed from the routing tables of all its peers. An "operational" route server may see advertisements for a given address space from several of its peers, for instance:

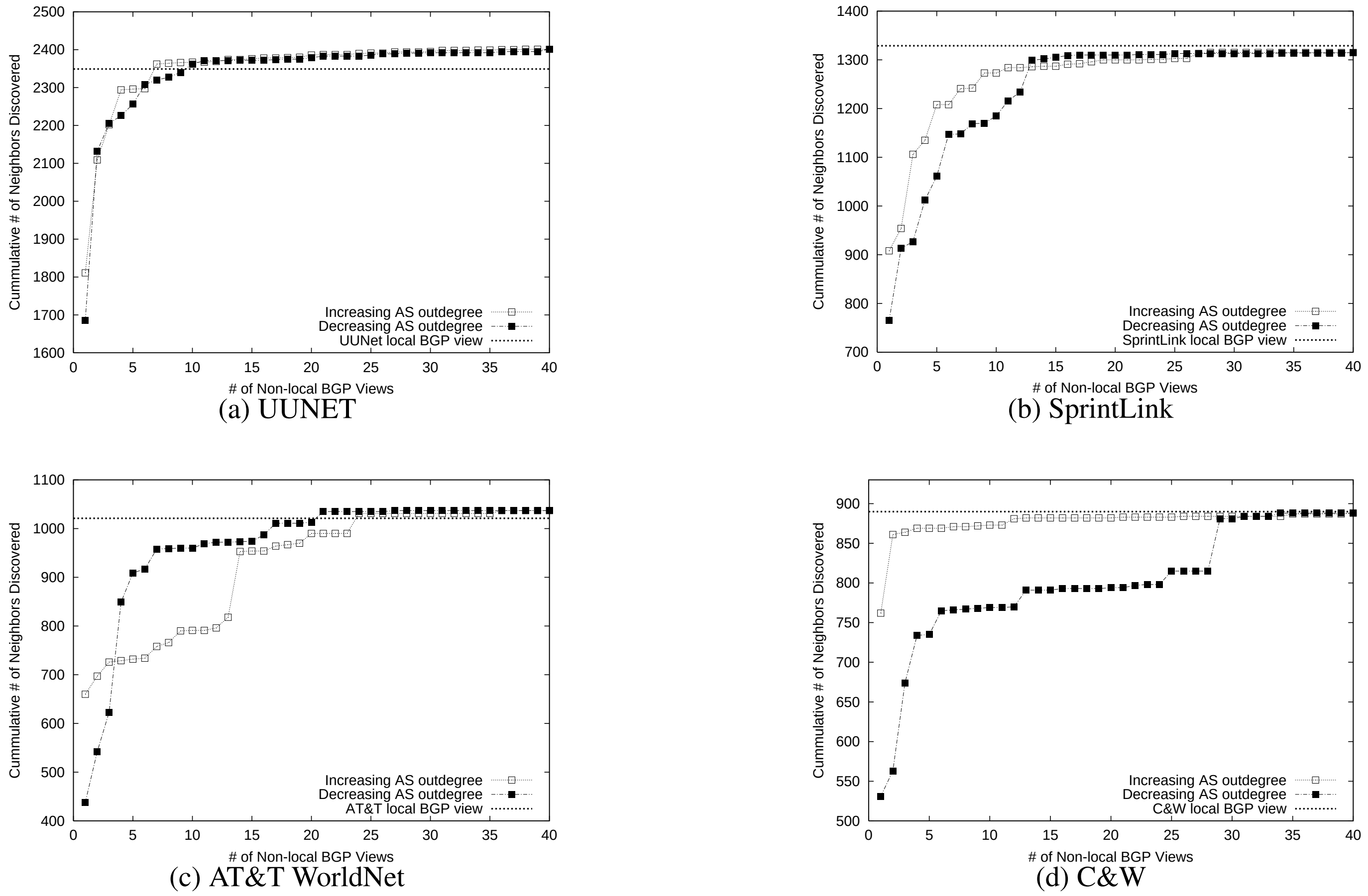
```
* i205.145.52.0  204.255.168.133  701 16758
* i              208.48.18.10     701 16758
* i              204.255.168.129  701 16758
* i              208.48.18.10     701 3744 16758
*>i              137.39.5.157     701 16758
* i              204.255.168.137  701 16758
```

In this example, the "operational" route server can reach the address space 205.145.52.0 through six of its peer routers, whose addresses are listed in the second column. The remaining columns list the ASs (the AS path) a packet destined for that address space must travel through for each of the alternatives. The best AS path for each address prefix, according to the local routing policy set by the administrator of the AS, is marked with '>' in conventional BGP routing tables. When a BGP router re-advertises a particular route, it will advertise only the best path, after prepending its own AS to the AS path. Therefore, to construct a BGP routing table of an "operational" route server, we use only the best entry for each individual address prefix. Recall that our goal in this section is to answer the question, "How many BGP routing tables from distinct ASs must we aggregate to capture all the vertex degree reported by the BGP routing table of a given AS?" Our construction of "operational" route servers' BGP routing table reflects our intention to construct the BGP routing table of an individual AS that can then be used to answer the question we pose. That is, our goal here is not to infer the AS-level topology from the individual "operational" route servers. If our goal had been to infer the AS-level topology from the individual "operational" route servers, we would have constructed a BGP routing table consisting of *all* AS paths from all peers of the route servers instead of just the best path. Doing so, however, will only bring us back to our original question of how complete such an AS-map would be.

To summarize, for "collector" route servers, each of their peers is a potential *BGP viewer*; whereas for "operational" route server, we have only a single potential *BGP viewer* whose routing table must be constructed from the routing tables of all its peers. Thus in this study we have 10 potential *BGP viewer*s from the 10 "operational" route servers and 85 candidates from the two non-commercial "collector" route servers. In addition, we have verified that the UUNET routing information we obtained was from a default-free BGP router, which should be considered an "operational" route server.

To qualify as a BGP viewer, a candidate's routing table must cover all of the known address space. To compare the coverage of the address space among our BGP viewer candidates, we look at four different measures in each of the candidates' routing tables: (1) number of routes, (2) number of non-aggregatable routes, (3) number of routes whose prefix length are less than or equal to 24,[2] and (4)

[2]In [11], the authors approximate the set of globally visible prefixes



(a) UUNET (b) SprintLink

(c) AT&T WorldNet (d) C&W

Fig. 2. Marginal Utility of Non-local BGP Views: Tier-1 ASs

number of origin ASs. An aggregatable route is a redundant route which could be removed from BGP tables by route aggregation; e.g., when a BGP table sees two routes, 12.0.0.0/8 (3786-1-7018) and 12.1.140.0/24 (3786-1-7018), the latter is a redundant route. The origin AS is the AS where the address prefix resides, in this example it is AS7018.

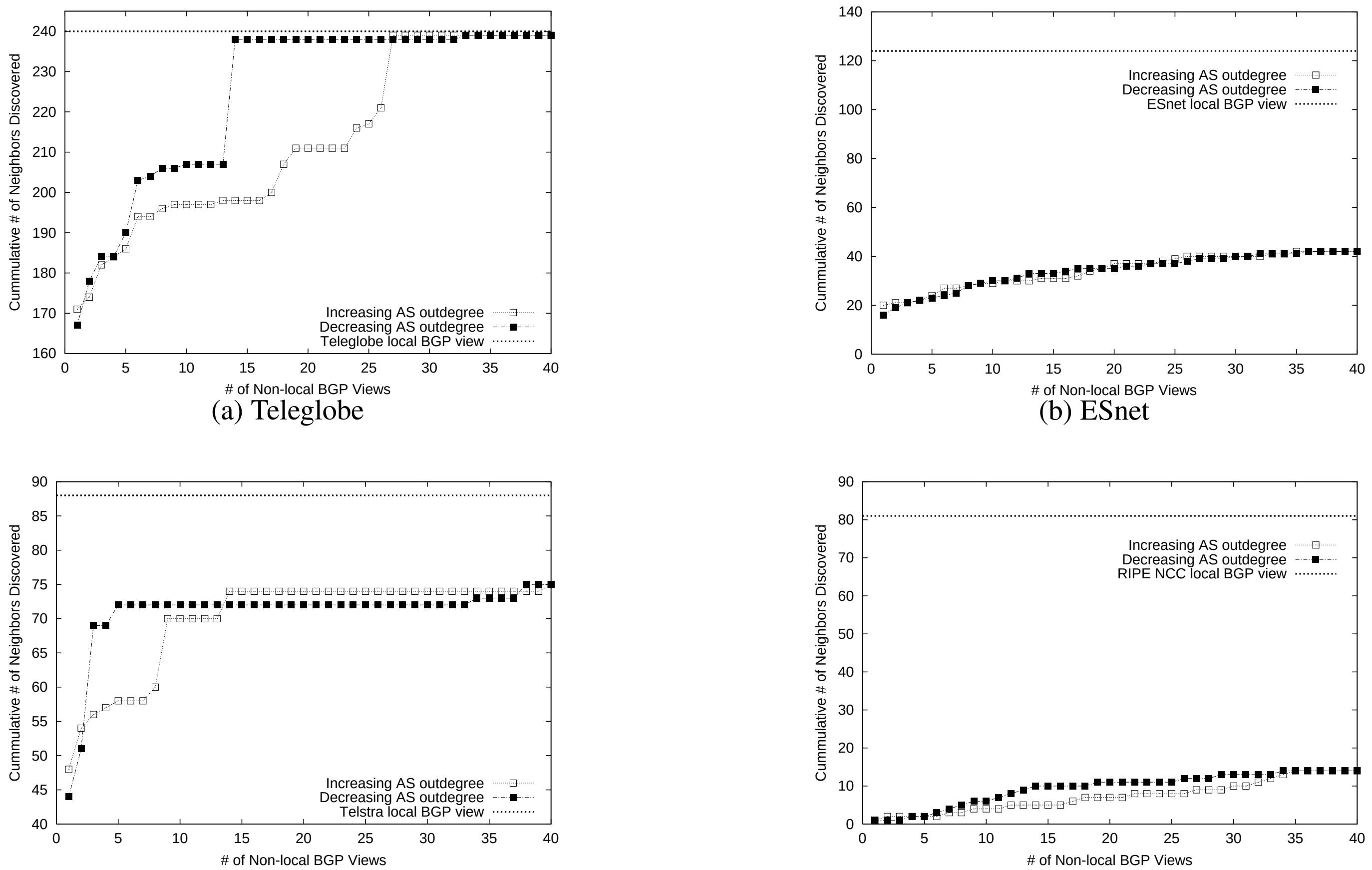
Next we sort these four measures for the 96 candidates in decreasing order and plot the top 65 in Figure 1. From the figure, it can be seen that all four measures visibly decrease after the 51st rank or so. It turns out that the four different measures pick out the same set of 51 BGP tables. Thus of the 96 candidate BGP viewers, only 51 of them satisfy our definition. These 51 BGP viewers reside in 41 distinct ASs. For our study, we use only 41 of these 51 BGP viewers, giving us views of the Internet from 41 distinct ASs. All of the 41 BGP views (Table II) were taken from the same date (May 25th, 2001) at approximately the same time of day.

by counting prefixes whose length are less than or equal to 24 in BGP tables.

C. *Local vs. Non-local BGP View*

Given our dataset, we ask, "How well are the peering relationships maintained by a certain AS observed by *other* ASs?" For example, can AT&T's BGP routing tables discover UUNET's AS neighbors reasonably well? How well will a small ISP's BGP view predict the AT&T's AS neighbors? This question has very practical relevance to the goal of our paper since constructing global AS-level topology today has been predicated on collecting a small number of BGP views from the Internet. Therefore, for a majority of existing ASs, we essentially have to predict their AS neighbors from a few selected BGP views.

To answer the above question, we consider two kinds of BGP views: "local" and "non-local". From a certain AS's perspective, a BGP view originating from that AS is considered *local* and those originating from any other ASs are *non-local*. For example, for AT&T (AS7018), a BGP view from AS7018 is considered local, whereas a BGP view from AS701 is non-local. From the UUNET (AS701)'s perspective, the opposite holds. Therefore, each of our selected 41 ASs has one local view and 40 non-



(a) Teleglobe (b) ESnet

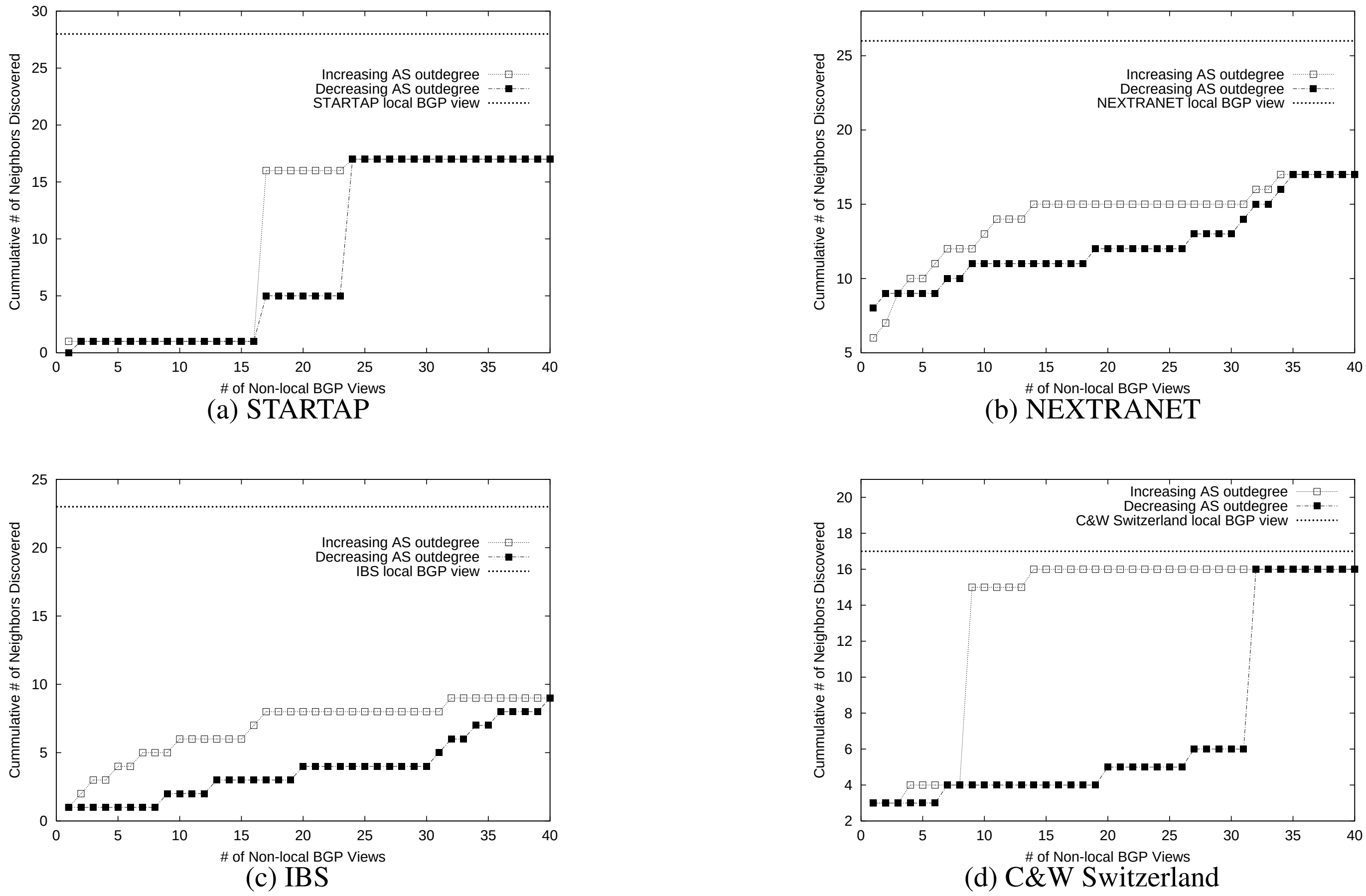
(c) Telstra (d) RIPE NCC

Fig. 3. Marginal Utility of Non-local BGP Views: Tier-2 ASs

local views. We consider that any kind of peering relationship maintained by AS X will be best observed by its local BGP view. Based on this, we compare—for each of the 41 ASs—AS X's vertex degree predicted by its 40 non-local views against that of its local view. By doing so, we will be able to quantify the completeness of *non-local* views.

In Figures 2 to 4, we look at the marginal utility of discovering AS neighbors by non-local views. That is, as we incorporate more non-local views (with respect to a certain AS), we look at how many more neighbors connecting to that AS are found. The total 40 non-local views are added in two different orders: A non-local view from the largest AS is added first and then a non-local view from the second largest AS is added, and so forth (noted as "decreasing AS outdegree" in the figures). The opposite order (the smallest AS first) is noted as "increasing AS outdegree". The horizontal dotted line in each figure represents the vertex degree predicted by the given AS's local BGP view. In the following, we present our findings for Tier-1 ASs and non-Tier-1 ASs separately since doing so provides us with more insights into non-local BGP views. For the purpose of this study, we roughly categorize our 41 ASs into three hierarchy levels (Tier-1, Tier-2 and Tier-3), based on its vertex degree derived from its own local BGP view. The vertex degrees of Tier-1 ASs (Tier-2 ASs) are one order of magnitude larger than those of Tier-2 ASs (Tier-3 ASs).

Tier-1 ASs. Figure 2 shows that a sufficient number of non-local BGP views manage to discover most of the neighbors connecting to Tier-1 ASs. Interestingly, however, individual non-local views are seen to contribute differently. More specifically, the neighbors of Tier-1 ASs tend to be better discovered by non-local views from smaller ASs than from larger ASs. In case of C&W, the smallest AS's non-local view discovers 760 out of 890 existing neighbors, whereas the non-local view from the largest AS is only able to discover 530 of them. This phenomenon can be intuitively explained by the *non-transitive* peer-to-peer relationship and the *transitive* provider-to-consumer relationship [12]. That is, the information regarding the pairwise peer-to-peer relationships maintained by a certain AS does not circulate among its peers themselves, but does propagate to its downstream customers. For this reason, the BGP views from smaller ASs are likely to better observe Tier-1 ASs' connections.



(a) STARTAP (b) NEXTRANET

(c) IBS (d) C&W Switzerland

Fig. 4. Marginal Utility of Non-local BGP Views: Tier-3 ASs

Non-Tier-1 ASs. An important observation about non-Tier-1 ASs (Figures 3 and 4) is that the combined, 40 non-local BGP views clearly fail to observe many existing peering relationships, though there are some exceptions (Figures 3(a), 4(a) and 4(d)). In these exceptional cases, one notices that a single AS (the 14th largest AS, the 17th smallest AS, and the 9th smallest AS respectively) is seen to find most of the given AS's neighbors. In the case of Figure 3(a), the 14th AS turns out to be the Teleglobe's European network. In the case of Figure 4(a), the 17th smallest AS belongs to RIPE NCC. The STARTAP network is an National Science Foundation-funded infrastructure connecting several international research sites. In the case of Figure 4(d), the 9th AS is a regional ISP in Switzerland. Conducting a `traceroute` probe to a router in that AS reveals that the ISP has a provider-customer relationship with C&W in Switzerland. Therefore, these cases re-confirm how well customer views can discover their providers' peering connections. For a majority of non-Tier-1 ASs whose customer views are not available to us, their connectivity information is not sufficiently approximated by 40 non-local BGP views.

Still, it came as a surprise to us that dozens of default-free BGP tables of different ASs are hardly sufficient to capture the major portion of the Internet connectivity as far as non-Tier-1 ASs are concerned. Given that a non-negligible number of neighbors of a certain AS can be concealed from other ASs, we seek to look more carefully at those missing neighbors.

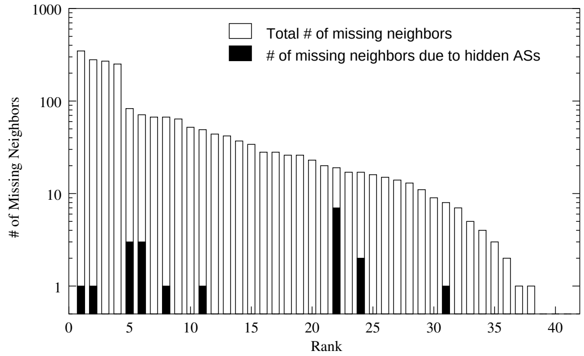
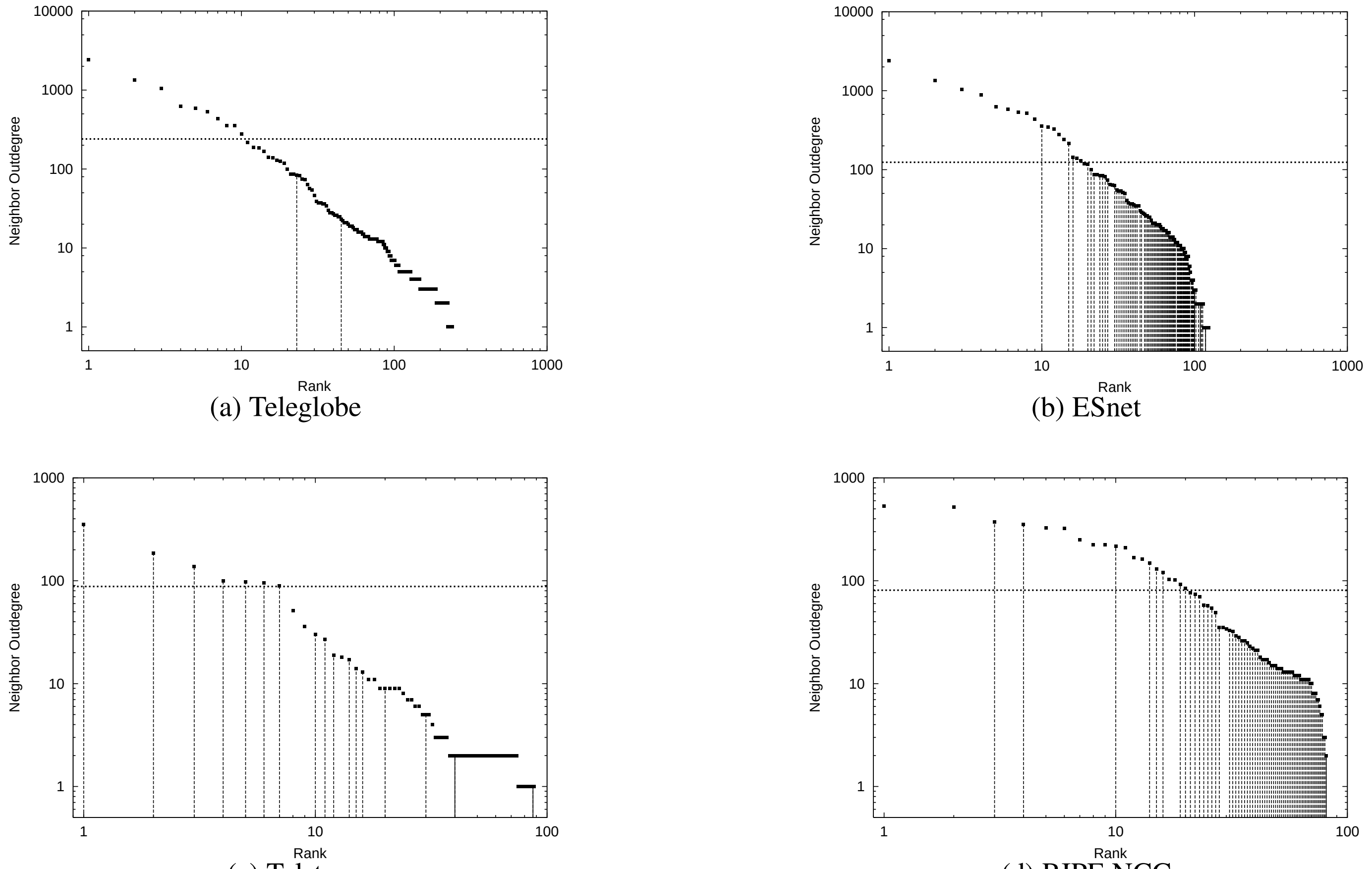


Fig. 5. Number of Missing Neighbors

First, are those missing neighbors caused by hidden *nodes* or hidden *edges*? A *node* is "hidden" if its AS num-



(a) Teleglobe (b) ESnet

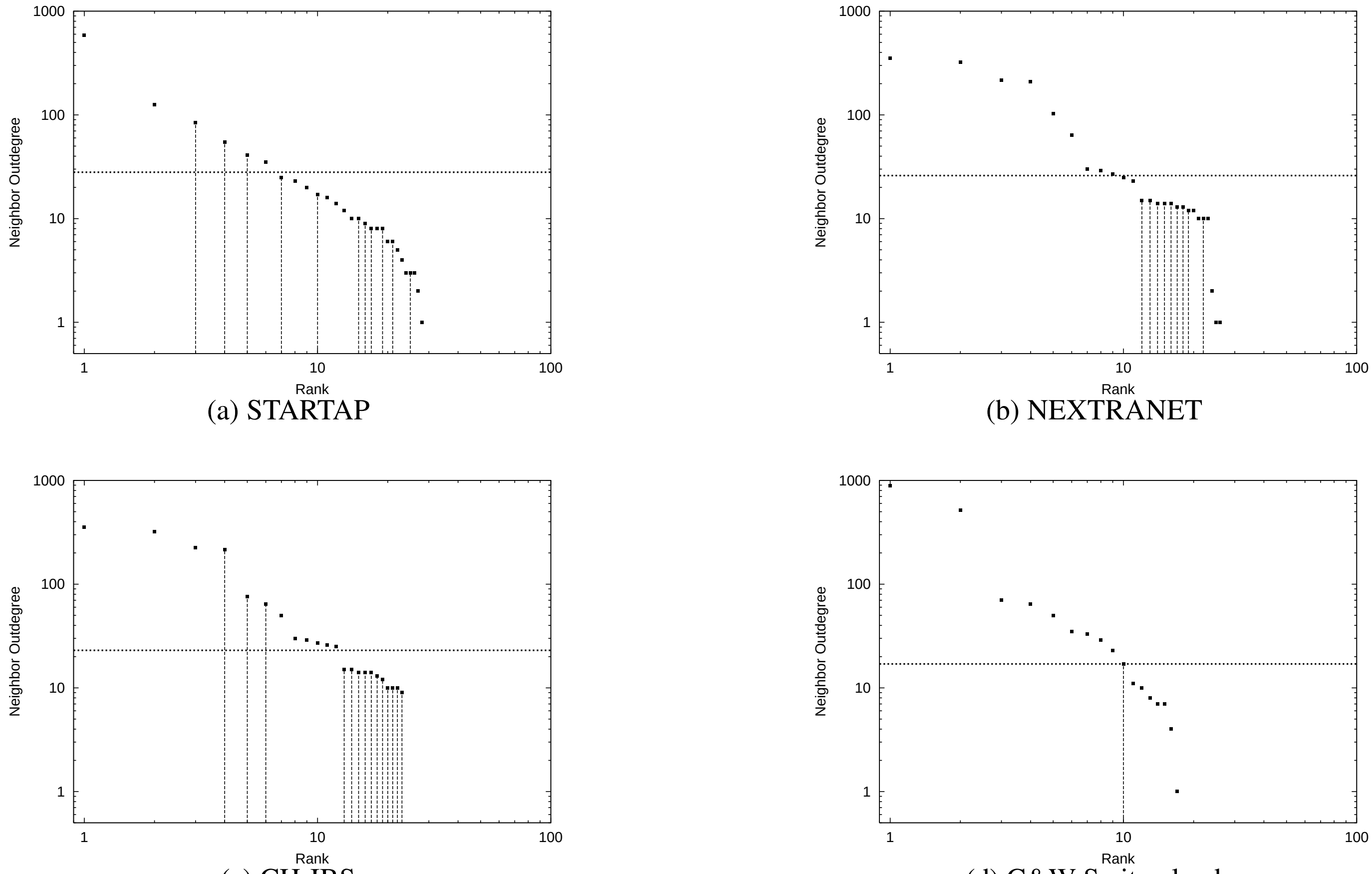
(c) Telstra (d) RIPE NCC

Fig. 6. Vertex degrees of 4 Tier-2 ASs' Neighbors

ber exists in the local AS's routing table but not in any of the other ASs' routing tables. When an AS is hidden, its address space may still be reachable to other ASs as part of a larger aggregated address space [13]. On the other hand, a hidden *edge* means that while the AS numbers of both end points of the edge are present in other ASs' routing tables, the neighbor peering relationship between the two end points is not listed in any AS path. Figure 5 lists, in decreasing order, for each AS X, the number of its neighbor peering relationships not detected in any of the other 40 ASs' routing tables. The solid component of each bar is the number of hidden ASs, i.e., ASs whose AS numbers are not present at all in non-local BGP views. The rest is caused by the non-local views not detecting the peering relationships between AS X and its neighbors. One can see that the number of missing neighbors from hidden ASs is negligibly small (the y-axis is in log-scale). Thus, the majority of missing neighbors are caused by hidden AS *links*, not hidden ASs.

Next, given that BGP views can fail to observe a non-negligible number of existing AS links, we want to know what differentiates those links from the visible ones. In the absence of detailed AS relationship and internal policy routing information, we base our analysis on studying AS vertex degree as follows. Since most missing neighbors occur with respect to non-Tier-1 ASs, we only consider non-Tier-1 ASs. Figures 6 and 7 show the vertex degrees of several non-Tier-1 ASs' neighbors, sorted in decreasing rank order.[3] The horizontal line in the figures indicate the vertex degree of the local AS (for example, Figure 6(a) shows that Teleglobe has a vertex degree of about 250). If a certain neighbor is hidden in all non-local views, then it is marked with a vertical dashed line. The frequency of vertical dashed lines indicates how many neighbors are missing in non-local views, and the y values of those lines point out the vertex degrees of the missing neighbors. Except for Figure 6(c), the first few largest neighbors (most likely upstream providers) of each AS are invariably well observed by other ASs. This is because those AS links connecting any AS and its upstream providers should eventually be visible in the BGP views of Tier-1 ASs that are located in the top hierarchy. Meanwhile, a majority of missing neighbors have vertex degrees comparable to or less than

[3] An AS vertex degree here is derived from the merged 41 BGP views.



(a) STARTAP (b) NEXTRANET

(c) CH-IBS (d) C&W Switzerland

Fig. 7. Vertex degrees of 4 Tier-3 ASs' Neighbors

that of given AS. Given the prevailing public peering connections occurring at NAPs and exchange points [14], we conjecture that many of those hidden links are based on peer-to-peer relationships.

D. Other Local BGP Views

To help troubleshoot Internet-wide routing problem, several ISPs have made available public but limited access to several of their selected BGP routers. These are called the *Looking Glass* sites. The looking glass sites also make available BGP summary information where a set of peers and aggregated BGP activity statistics are displayed. From this BGP summary information, we can elicit a set of neighbor ASs connecting to the local AS. Since those looking glass sites are supposed to be used interactively, we had to select some of those sites [15] and have our crawling script periodically collect their BGP summary information.[4] Once we obtain the number of each AS's neighbors from its corresponding BGP summary information, we compare it with the one predicted by our BGP views, much like the way we did previously. However, unlike default-free BGP tables, the BGP summary information from individual looking glass sites may not list all its existing neighbors since those routers originating the information may not be located in default-free zone. Therefore, the number of neighbors inferred from this BGP summary information should be regarded as the *minimum* number of actually existing neighbors.

Figure 8 summarizes the result obtained from using the Looking Glass sites in a plane, where the y-axis gives the vertex degree predicted by the combined non-local BGP views and where on the x-axis, we plot the vertex degree based on individual looking glass sites. On the same figure, we also plot the neighbor discovery result of the 41 ASs from the previous section (noted as "BGP-table-based local AS info"). Once again, it is clear that the vertex degrees predicted by non-local BGP views is misleading.

Contrary to conventional wisdom, our case study shows that a non-negligible number of existing AS connections can be commonly hidden in most BGP routing tables and that the observability of AS connections in the tables depends to a large extent on existing AS relationships. This

[4]This script starts to run at the same time as we collect BGP dumps from route servers.

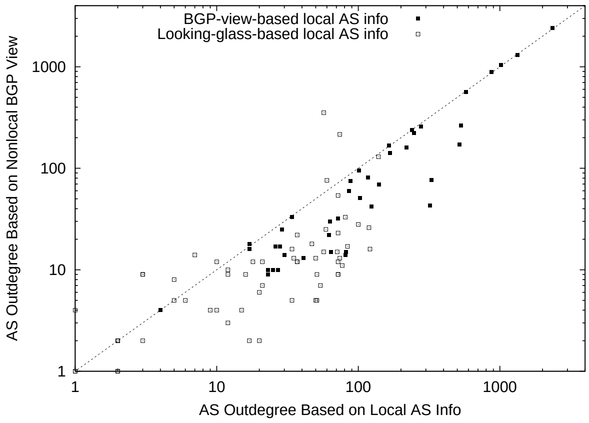


Fig. 8. The Completeness of Non-local BGP Views

in turn suggests that the Internet might maintain much richer connectivity than is currently being observed by relying on a handful of BGP routers.

III. THE AS GRAPH VERTEX DEGREE DISTRIBUTION REVISITED

The rather surprising result from our earlier BGP view analysis now casts a more fundamental question: "How would the *complete* Internet AS-level topology deviate from its Oregon-based counterpart, where the latter is considered to be an accurate snapshot of the actual AS-level topology?"

Our result implies that in order to best predict the AS-level topology, BGP views should be collected from end-customer ASs located in the *lowest* AS hierarchy. However, we do not know how many of them should be mobilized to discover most of existing upstream connections. Facing this obstacle, we resort to the Internet Routing Registry (IRR) [16] to further glean local AS neighborhood information. The IRR maintains individual ISPs' routing information in several public repositories in an attempt to coordinate global routing policy. The IRR's routing policy database expresses routing information at various levels (e.g., individual address prefixes or ASs, etc.). The following two imaginary database records exemplify how the routing information of the IRR is expressed in Routing Policy Specification Language (RPSL).

```
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 AS60003
           to AS3 announce ANY
changed:   admin@foo.com 20010313
source:    RADB
```

The first record states that "1.2.3.0/24" belongs to AS1 as of Mar. 13, 2001. The latter record, which expresses AS1's import and export routing policies, indicates that AS1 has two peering neighbors AS2 and AS3 with which it exchanges route reachability information of AS4. From this import and export policies specification, we can conveniently infer the neighboring ASs of AS1.

A. *On the Freshness of the Internet Routing Registry*

The question now is then the reliability of such manually-registered policy routing information. The motivation of the IRR is to minimize the negative impact of the growing number of ASs and the accompanying complexity of inter-AS connectivity on the Internet routing infrastructure [17]. However, being predicated on voluntary publication of routing policy, the IRR database may not be complete and some part of it can simply remain out-of-date.

It is known that an increasing number of ISPs rely on the IRR to filter route announcements at border routers [18]. Especially, the RIPE database of the IRR is actively used by most ISPs in Europe. Many European exchange points [19], [20], [21], [22] specify as a membership requirement that members register their routes and peering policy in the RIPE database. Based on these widespread practices in Europe, we consider the RIPE database a viable option when trying to infer AS-level topology.

We first set out to verify the freshness of the RIPE database. For this purpose, for each routing registry record we looked at its last update time and collated its routing information with the BGP tables[8], [23] corresponding to the *same* period as the update time. Registry entries recording specific route (address prefix) are checked for their route origin information; registry entries specifying the routing policies of ASs are checked against the ASs' neighbors lists.

We downloaded public IRR database files mirrored at [24] on May 25th 2001 and compared each of their records with the BGP table created on the record's last update time. A given registry information can be either (a) consistent, (b) inconsistent, or (c) not available in the tables. For example, a BGP table can correctly observe that 1.2.3.0/24 originates from AS X (a), or incorrectly tells that 1.2.3.0/24 belongs to AS Z (b), or simply does not observe the prefix at all (c). We consider a registry entry obsolete if its information used to be consistent[5] but has

[5] A certain record is considered consistent with a BGP table if the

become inconsistent with or not available in the more recent (May, 25th) BGP table.

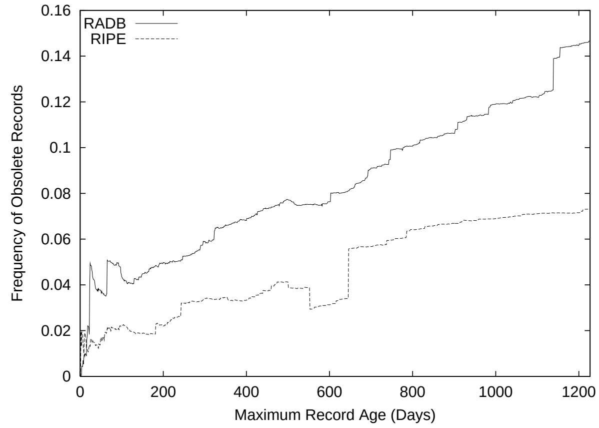


Fig. 9. Frequency of Records with Obsolete Route Info

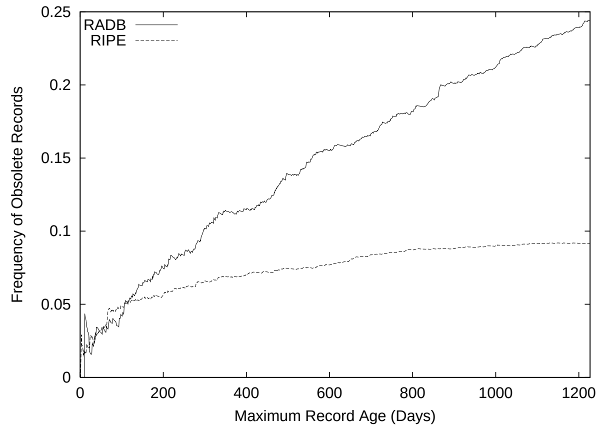


Fig. 10. Frequency of Records with Obsolete AS Policy Info

Figures 9 and 10 compare the two largest IRR databases (RADB & RIPE) in terms of freshness. They show the frequency (y-axis) of obsolete records among those that have been last updated within a certain number of days, where the number of days are given on the x-axis. The age of a certain record thus indicates how many days have passed since the record had been last updated till the entire database was retrieved on May 25th. Both figures clearly demonstrate that the RIPE database is maintained more up-to-date than the RADB database.

B. AS Graph Vertex Degree Distribution

Satisfied with the relative freshness of the RIPE database, we next extracted AS topology information from the RIPE database. As suggested by Figures 9 and 10, the reliability of individual database records tends to deteriorate as they age. Therefore, we only considered RIPE records that were at most one month old.

route origin AS specified in the record is correct according to the BGP table, or if there is no such AS neighbor that is found in the BGP table but not specified in the record.

TABLE III
AS GRAPH STATISTICS

Source	# nodes (%inc)	# edges (%inc)
Oregon	11,174	23,409
Combined 41 BGP views	11,201 (0.24%)	25,472 (8.9%)
Oregon+LG	11,234 (0.54%)	26,519 (13.3%)
Oregon+LG+RIPE(1w)	11,387 (1.9%)	28,836 (23.2%)
Oregon+LG+RIPE(2w)	11,498 (2.9%)	31,872 (36.2%)
Oregon+LG+RIPE(1m)	11,573 (3.6%)	34,581 (47.7%)

Table III summarizes six different AS graphs that result from a different number of sources of information (41 BGP views, Oregon route server, Looking glass sites and RIPE database). All the information was collected on the same day. '%inc' denotes the percentage of increase with respect to the Oregon-based AS graph. The looking glass information was retrieved from about 360 routers belonging to about 70 different ASs. The RIPE database information was selectively included based on the record age. For example, the length of period (1 week, 2 weeks or 1 month) in the parenthesis denotes the maximum age of those records that were included. We can easily notice that the AS graph becomes much less sparse as we incorporate more information.

Finally, we check how those increasingly denser AS graphs affect the power-law characteristics that have been identified by Faloutsos et al. [1] using Oregon-based AS graph. According to Figure 11, newly-added edges turn out to significantly increase the vertex degree of those nodes whose ranks are between 10 and 1000. Thus, the overall vertex-degree-rank distribution takes on a more pronounced convex shape. The frequency distribution of Figure 12 shows that newly-added edges have minimal effect on the frequency of nodes whose vertex degrees are less than 10. However, starting from vertex degree 10, the vertex degree frequency tends to deviate upward, forming an overall concave distribution curve. These results are consistent with our earlier findings in II-C suggesting that there may exist much richer connectivity among non-Tier-1 ASs than being observed by a handful of BGP routers.

Next, we look at whether such renewed AS vertex degree distribution qualitatively remain the same over time. We have been collecting the looking glass information and the IRR database once every week since Mar. 2001. Currently we have 12 instances of these datasets, each of which was collected on the same day of the week. In Figure 13, we plot the complementary CDF of AS vertex degree distribution for the 12 instances of

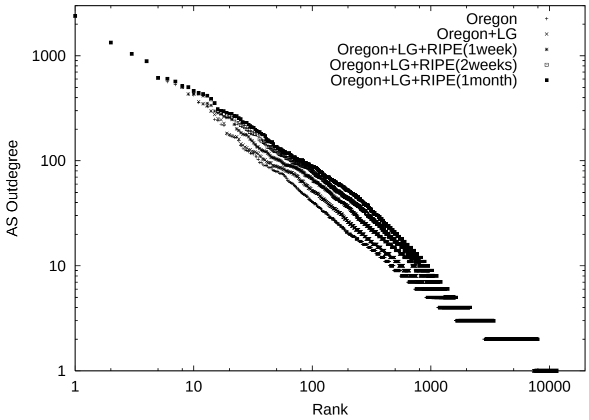


Fig. 11. Power-Law 1: AS Vertex Degree vs. Rank

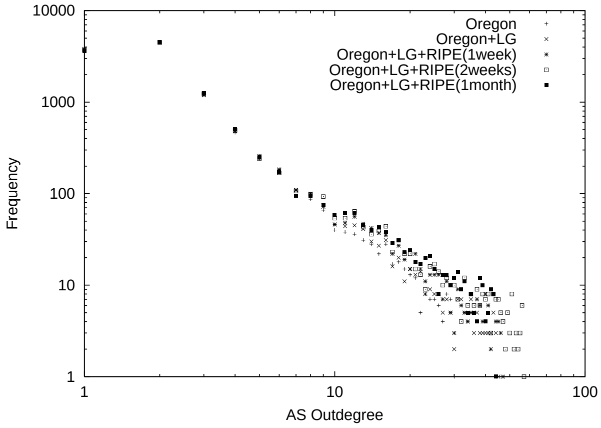


Fig. 12. Power-Law 2: Frequency Distribution of AS Vertex Degree

each of our 5 different AS graphs (Oregon, Oregon+LG, Oregon+LG+RR(1week), Oregon+LG+RR(2week), Oregon+LG+RR(1month)). We can see that the qualitative difference; i.e., pronounced deviations from the power law behavior of the Oregon-based AS maps, caused by relying on additional information remains over time.

IV. Conclusion

With the recent significant increase in research efforts focusing on Internet routing behavior, routing-related measurements have become a highly valuable commodity, and NLANR and the Oregon route server have been the leaders in supplying the research community with invaluable and relevant data, in particular with BGP measurements. Compared to other concurrent efforts that have focused, for example, on Internet workload characterization, where measurements have become more of a household commodity over time and where there has been a large community of traffic experts, the number of researchers with expertise in Internet routing is still relatively small, and until just a few years ago, BGP was known to, leave alone used by, only a handful of experts. As a result, the general research community can be expected to be more prone to using routing-related measurements such as the NLANR BGP measurements in ways for which the data are really not applicable, sufficient, or intended for in the first place. In this paper, we show that indeed, as far as past investigations into the Internet topology at the AS level are concerned, many of the findings that have been reported in the literature to date have used the publicly available BGP measurements without realizing the possible pitfalls associated with taking the data at face value and being unfamiliar with the specifics of BGP— in short, without examining whether or not the use of the data is justified for studying the problem of inferring the Internet AS connectivity. Our results confirm that while the actual connectivity of the Internet at the AS level is quite high, BGP measurements typically see only a portion of all existing AS connections. From BGP's perspective, this observation comes as no surprise, because it expresses the defining property of BGP; that is, BGP reflects AS relationships and not physical AS connectivity. In this sense, the main lesson learned from the case study presented in this paper is that since network-related measurements often reflect network protocol-specific features, it is in general wise to consult or work with experts who know the particular protocol(s) well enough to help decide whether or not the collected data can be used when arguing for the general validity of an empirical finding about the Internet.

Acknowledgments

We thank Tim Griffin for valuable discussions and for offering us the UUNET's BGP routing table.

References

[1] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proceedings of ACM SIGCOMM '99*, August 1999.

[2] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct 1999.

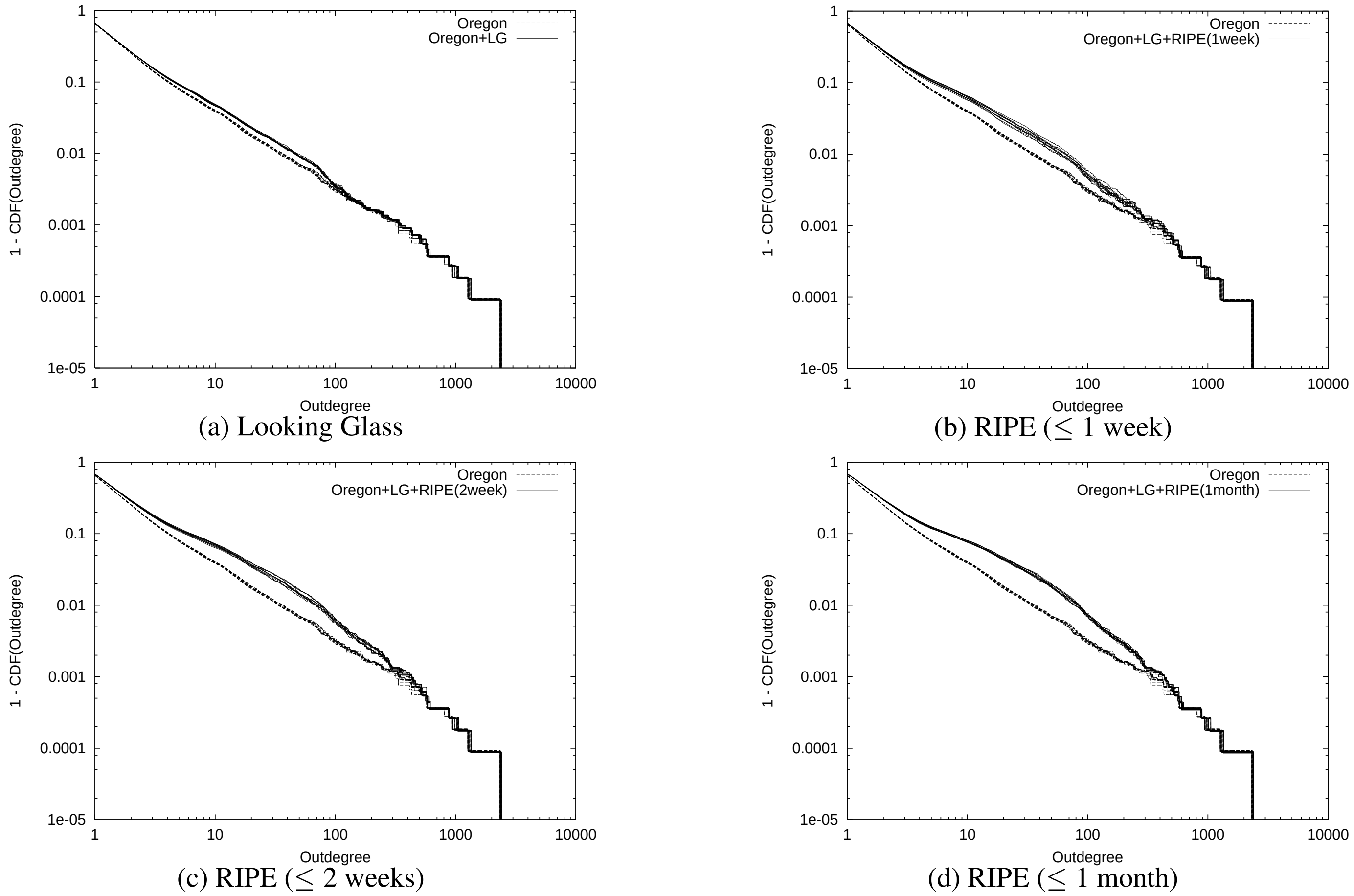
[3] R. Albert and A.-L. Barabasi, "Topology of evolving networks: Local events and universality," *Physical Review Letters*, vol. 85, pp. 5234–5237, 2000.

[4] A. Medina, I. Matta, and J. Byers, "Brite: A flexible generator of Internet topologies," Tech. Rep. BU-CS-TR-2000-005, Boston University, CS Dept., 2000, http://www.cs.bu.edu/brite/.

[5] C. R. Palmer and J. G. Steffan, "Generating network topologies that obey power laws," in *Proceedings of the Global Internet Symposium, Globecom2000*, November 2000.

[6] C. Jin, Q. Chen, and S. Jamin, "Inet: Internet topology generator," Tech. Rep. CSE-TR-433-00, University of Michigan, EECS Dept., 2000, http://topology.eecs.umich.edu/inet.

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



(a) Looking Glass (b) RIPE ($\leq$ 1 week)

(c) RIPE ($\leq$ 2 weeks) (d) RIPE ($\leq$ 1 month)

Fig. 13. Frequency Distribution of AS Vertex Degree Over Time

[8] National Laboratory for Applied Network Research, "NLANR RouteViews archive," http://moat.nlanr.net/Routing/rawdata.

[9] W. Willinger, R. Govindan, S. Jamin, V. Paxson, and S. Shenker, "Scaling phenomena in the Internet: Critically examining criticality," 2001, preprint.

[10] Swiss Network Operators Group, ," http://www.swinog.ch/tools.html.

[11] A. Broido and K. Claffy, "Analysis of BGP data from Oregon route views," *Workshop on Network-Related Data Management (NRDM 2001)*, May 2001.

[12] L. Gao, "On inferring autonomous system relationships on the Internet," in *Proc. IEEE Global Internet Symposium*, November 2000.

[13] E. Chen and J. Stewart, "RFC 2519: A framework for inter-domain route aggregation," February 1999.

[14] G. Huston, "Interconnection, peering and settlements," in *Proceedings of the 9th Annual Conference on the Internet Society*, 1999.

[15] Traceroute.org, "Public route server and looking glass list," http://www.traceroute.org/.

[16] Merit Network, "Internet routing registry," http://www.irr.net.

[17] R. Govindan, C. Alaettinoglu, G. Eddy, D. Kessens, S. Kumar, and W. Lee, "An architecture for stable, analyzable Internet routing," *IEEE Network Magazine*, January-Feburary 1999.

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

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

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

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

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

[23] S. McCreary and B. Woodcock, "PCH RouteViews archive," http://www.pch.net/documents/data/routing-tables.

[24] Merit Network, "Internet routing registry mirror site," ftp://ftp.radb.net/routing.arbiter/radb/dbase/.