Putting the Links Together: Measurements and Impact

Yihua He yhe@cs.ucr.edu
Georgos Siganos siganos@cs.ucr.edu
Michalis Faloutsos michalis@cs.ucr.edu
Srikanth Krishnamurthy krish@cs.ucr.edu
Department of Computer Science and Engineering
University of California, Riverside

***Abstract*— The lack of an accurate model of the Internet topology at the Autonomous System (AS) level is a limiting factor in many design, simulation, and modeling efforts. In this paper, we tackle this problem and provide two main contributions: (a) we develop methodologies and tools to identify missing AS links, and (b) we quantify the effect of the new links on the Internet topology model. First, we conduct a large-scale comprehensive synthesis of most available sources of information. We cross-validate and compare BGP routing tables, Internet Routing Registries, and traceroute data, while we extract significant new information from the less-studied Internet Exchange Points (IXPs). We identify *40% more edges* and *approximately 300% more peer-to-peer edges* compared to commonly used datasets. Second, we identify patterns of the new edges and quantify their effect on important topological properties. Given the new peer-to-peer edges, we find that for some ASes more than 50% of their paths stop going through their ISP providers assuming No Valley Customer Prefer routing. A surprisingly observation is that the degree of a node may not be a good "judge" of which ASes it will peer with: the two degrees differ by a factor of four or more in 50% of the peer-to-peer links. Finally, we attempt to provide estimates of how many more edges we may be missing.**

I. Introduction

An accurate model of the Internet topology is important for simulating, analyzing, and designing protocols effectively [1]. With an accurate topology, first, we can design and analyze new interdomain routing protocols, such as HLP [2], that take advantage of the properties of the Internet AS-level topology. Second, we can create more accurate models for simulation purposes[3]. Third, we can analyze and estimate phenomena, such as the spread of viruses [4][5], more accurately.

Developing an accurate model of the Internet topology at the AS level remains as a challenge, despite the recent flurry of studies[6][7][8][9][10][11][12][13]. Several sources of topological information exist: (a) archives of BGP routing tables, (b) Internet Routing Registries, and (c) archives of traceroute data. Each of the sources has its own advantages and limitations, as we will discuss in the next section. Several studies attempt to find a complete Internet topology [6][7][8][9] but, as we explain in section II-C, our study synthesizes more sources of information in a more exhaustive way. Relatively few efforts study the Internet Exchange Points (IXPs) [14][15][16], but they have different goals from ours. Finally, some studies [17][18] document the limitations of the sources of topological information, but without necessarily attempting to identify a more complete topology.

The goal of this work is to answer two related questions:
a. How many more AS links can we find?
b. Does finding these missing links matter?

For the first question, although we have many sources of information, they have not been exploited to their fullest. Each of these sources provides an incomplete, sometimes inaccurate and complementary to each other view of the topology. In addition, most previous studies examine these sources in isolation (typically using the BGP tables). Very few studies have looked at pairwise combinations of these datasets. Furthermore, IXPs have received little attention and we argue that they hide significant topological information. The second question asks whether finding new links will change our understanding of the Internet in a non trivial way. If the properties of the Internet model do not change, then we can be content with the current topologies we have.

The contribution of this work is a massive in-depth study of missing links and their importance on our understanding of the Internet at the AS level. The work consists of two parts: (a) we discover a significant number of new edges, and (b) we show that their effect on topology is non-trivial. In addition, we identify properties of the missing links, which can help us understand why we missed them and how we can possibly find more.

a. Identifying missing links: The contribution here is dual: (a) we develop methodologies and tools, and (b) we produce a more accurate Internet topology.

First, we identify and validate a significant number of AS links by a careful cross-reference and synthesis of most known sources of information: BGP tables, traceroute, and IRR. Second, we extract significant new information from Internet Exchange Points (IXPs), which are typically not used in topological studies. In our effort to do so, we improve significantly the current state of the art method for identifying the AS participants in an IXP.

Note that we set a highly conservative standard in our search: we only accept new edges when we are confident that they exist. In other words, we do not provide a union of the existing sources of information, but a critical synthesis. For example, we use our traceroute-based tool, RETRO, to confirm the existence of edges which we suspect, exist.

We arrive at several interesting observations:

(i) We find a significant number of new edges: We discover *40% more edges* (15%) and *approximately 300% more peer-to-peer edges* (65%) compared to the widely used Oregon Routeviews dataset (all available BGP routing tables).

(ii) Most of the newly discovered edges are peer-to-peer edges: The current topological models have a bias by under-representing peer-to-peer edges.

(iii) IXPs and IRR are good sources of potential new edges: Nearly 95% of the new peer-to-peer links are incident at IXPs, while nearly 80% of the new edges existed in the set of potential edges, which we extracted from IRR. Note that we refer to carefully filtered IRR data using the state of the art tools [19] for this purpose.

b. The properties and the impact of the new links: The new edges change our view of the Internet AS topology in a significant way. In addition, we identify interesting patterns of the new edges, which are mostly peer-to-peer edges. Our study can be summarized in the following main points.

(i) The new edges affect the routing decisions of ASes significantly: The new edges affect to a larger exent the ASes with degrees in the 10 to 300 range, which we refer to as "middle-class" ASes for convenience. We attempt to quantify the financial benefit ASes have from the new peer-to-peer edges assuming *No Valley Prefer Customer (NVPC)* routing. Note that the No Valley Prefer Customer routing is an abstraction of the actual Internet routing. It is extremely difficult, if not impossible, to model Internet routing in full detail at such a large scale. At the same time NVPC routing is much more realistic than shortest path routing. For some ASes, more than 50% of their paths stop going through a provider. We conclude that business-oriented and routing studies should consider all peer-to-peer edges for accurate results.

(ii) The degrees of the nodes of a peer-to-peer link can vary a lot: We find that 50% of the peer-to-peer edges are between nodes whose degrees differ by a factor of more than 4 or by a degree difference of 144. This has direct implications on how we think about and model peer-to-peer edges. For instance, this observation suggests that researchers need to use caution when using the degree as an indication of whether two ASes could have a peer-to-peer relationship. Our results can provide guidelines to AS policy inference algorithms, which partly rely on the node degree [20][21][22].

(iii) We find that provider-customer and peer-to-peer edges have significantly different properties and they should be modeled separately: We propose a divide-and-conquer approach for modeling the edge distribution in the Internet topology. We find that the distribution of the provider-customer *only* edges can be accurately described by a power-law (with correlation coefficient higher than 99%) in all the topological instances that we examine. In contrast, the distribution of peer-to-peer edges is better described by a Weibull distribution with correlation coefficient higher than 99%.

(iv) More peer-to-peer edges may exist: We estimate that we may be missing roughly 35% peer-to-peer edges compared to the peer-to-peer edges we know at the end of this study. Our estimate is based on the following rationale. Given our conservative approach, we only add possible edges when they can be verified by RETRO. Currently, we can only verify a a percentage of these possible edges. Not being able to verify an edge could mean that we do not have a traceroute server appropriately located. Thus, our estimate is an educated guess on how many more possible edges we could verify, if we had more traceroute servers.

The rest of this paper is organized as follows. We review the data sources and previous work in Section II. In Section III, we synthesize most known data sources to find the missing AS links. In Section IV, we quantify the impact of our new found AS links. We introduce our methods to identify the IXP participants in Section V. In Section VI, we summarize our work.

II. Background

In this section, we describe the most popular data sources and their pros and cons in terms of the data quality. Then, we present the data that we use in this study. Finally, we review related work and explain how our work differs.

A. Data Sources and Their Limitations

A. BGP routing information. BGP routing table dumps are probably the most widely used resources that provide information on the AS Internet topology. Each table entry contains an AS path, which corresponds to a set of AS edges. Several sites collect tables from multiple BGP routers, such as Routeview[23] and RIPE/RIS[24].

A.1 (Pro) : Trustworthiness. An advantage of the BGP routing tables is that their link information is considered reliable. If an AS link appears in a BGP routing table dump, it is almost certain that the link exists.

A.2 (Con): Incompleteness. Completeness is a limiting factor that hinges on the number of points of observation. A single BGP routing table has the union of "shortest" or, more accurately, preferred paths from the point of observation. As a result, such a collection will not see edges that are not on the preferred path for this point of observation. Several theoretical and experimental efforts explore the limitations of such measurements [17][25].

A.3 (Con): Statistical bias per link-type. Some types of AS links are more likely to be missing from BGP routing table dumps or traceroute data. Thus, apart from being incomplete, the measured graph may not fairly represent the different types of links. First, peer-to-peer links are likely to be missing due to the selective exporting rules of BGP. Typically, *a peer-to-peer link can only be seen in a BGP routing table of these two peering ASes or their customers.* A recent work [13] discusses in depth this limitation. Second, it is likely to miss alternative and back-up paths. By definition, a router advertises only the best path to each destination, namely an IP prefix. Therefore, the back-up paths will not show up in distant ASes unless the primary link breaks. To address the problem, a recent effort suggests the need for actively probing backup links [12].

B. Traceroute-based information. Using traceroute, one can explore IP paths and then translate the IP addresses to AS numbers, thus obtaining AS paths.

B.1 (Pro): The traceroute path information is considered reliable, since it represents the path[1] that the packets actually traverse.

B.2 (Con): IP-to-AS inaccuracies. The challenge with the traceroute data is the mapping of an IP path to an AS path. The problem is far from trivial, and it has been the focus of several recent efforts [15][16].

B.3 (Con): Traceroute shares A.2 and A.3 limitations. A traceroute server explores the routing paths from its location towards the rest of the world, and thus, the collected data has the same limitations as BGP data in terms of completeness and link bias.

C. Internet Routing Registries (IRR). Internet Routing Registry (IRR)[26] is the union of a growing number of world-wide routing policy databases that use the Routing Policy Specification Language (RPSL). In principle, each AS should register routes to all its neighbors (that reflect the AS links between the AS and its neighbors) with this registry.

C.1 (Con): IRR information is manually maintained and there is no stringent requirement for updating it. Therefore, without any processing, AS links derived from IRR are prone to human errors, could be outdated or incomplete. The up-to-date IRR entries provide a wealth of information that could not be obtained from any other source. A recent effort [19] shows that, with careful processing of the data, we can extract a non-trivial amount of correct and useful information.

B. The data sources used in this study

BGP table information. We download and archive information from a number of research initiatives [23][24] and ISPs, which make their BGP routing tables publicly accessible[27].

Traceroute data. We use traceroute data collected from our tool RETRO, which we present later, and from Skitter [28]. *Skitter* is one of the most popular traceroute-based archives. Skitter continuously traces IP paths from about two dozen monitors to almost all IP blocks and makes the results publicly available.

IRR data. We are very cautious with the use of IRR data. First, we carefully process the data with the state of the art tool [19] to obtain a set of possible edges. Second, we validate these edges with traceroutes using our RETRO tool.

IXP data. We first obtain the home pages and IP blocks of IXPs from PCH[29]. We then developed a tool to identify IXP participants from the retrieved information. (Section V) At last we generate a set of potential IXP edges, which give us hints on the location of possible missing AS links.

C. Related Work and Comparison

There has been a large number of measurements studies, with different goals, at different times, and using different sources of information. To the best of our knowledge no previous work combined the sources of information as we do here, for the purpose of identifying missing AS links. In fact, most previous efforts use a single source of information. In our overview below, we limit the presentation to the most relevant and recent studies.

The most relevant previous work is done by Chang et al. [6]. They identify new edges by looking at several sources of topological information including BGP tables and IRR. They estimate that 25%-50% AS links were missing from Oregon Routeview BGP table, the most commonly used data set for AS topology studies. Their work was an excellent first step towards a more complete topology.

Our work has the following characteristics that distinguish it from the effort above and most previous other efforts: (a) we use many more sources of information, and in addition, we use the Internet Exchange Points in a significant way to identify more edges, (b) we use a more sophisticated and comprehensive tool [19] to filter the IRR data, which previous studies did not have, (c) we employ a "guess-and-verify" approach for finding more edges by identifying potential edges and validating them through targeted traceroutes, (d) we accept new edges conservatively and only when they appear in a BGP table or a traceroute.

NetDimes [8] is an ingenious effort to collect a massive amount of host-based traceroute information. The key here is to increase the number of traceroute points by turning cooperative end hosts into observation points. The challenge now becomes the measurement noise removal, the collection, and processing of the information [30]. Our approach and NetDimes could complement and leverage each other towards a more complete and accurate topology.

Several other interesting measurement studies exist. Donnet et al. [31] proposes efficient algorithms for large-scale topology discovery by traceroute probes. Rocketfuel [32] explores ISP topologies using traceroutes. In [9], the authors examine the information contained in BGP updates. In [13], a macroscopic estimate of the missing edges is developed, but without as much effort in identifying where these missing edges are. Finally, a recent effort explores the information that can be extracted from IRR [19].

There are several efforts that study the topology and they would benefit from an accurate and complete topology. A plethora of efforts attempts to model the topology and to generate realistic topologies [33][34][35][36][37]. A recent study [11] models the evolution of the Internet topology by investigating the process of AS peerings.

The exhaustive identification of IXP *participants* has received limited attention. Most previous work focuses on identifying the existence of IXPs. For example, recent work [15] [16] proposes methods to identify whether a given path passes through an IXP in an effort to provide accurate AS-traceroute capabilities. Theoretically, if we had all possible traceroute paths, and the detection of IXPs could be done accurately, we could identify all the IXP participants. But this

[1]An exception is when the route changes while a path is being explored by a traceroute.

TABLE I

THE TOPOLOGICAL DATA SETS OF OUR STUDY. WE BOLD THE MAJOR SETS WHICH HAVE WELL VALIDATED EDGES.

OBD	The Oregon routeviews BGP table Dump
BD	OBD and other additional BGP table Dumps
IRRnc	IRR edges processed by Nemecis with non-conflicting policy declarations
IRRdual	IRRnc edges correctly declared by both adjacent ASes
BD+IRR	BD and the edges of IRRdual confirmed by RETRO
IXPall	Union of cliques of IXP participants
ALL	BD+IRR and the potential IXP edges that are confirmed by RETRO

TABLE II

THE STATISTICS OF THE TOPOLOGIES GENERATED FROM STUDY.

Name	Nodes	Edges	p-c	p-p
OBD	19.8k	42.6k	36.7k	5.5k
BD	19.9k	51.3k	38.2k	12.7k
BD+IRR	19.9k	56.9k	38.2k	18.3k
ALL	19.9k	59.5k	38.2k	20.9k

would require a really large number of traceroutes from many different vantage points. To avoid this problem, Xu et al. [14] develop what appears to be the first systematic method for identifying IXP participants. Inspired by their work, our approach subsumes their method, and thus, it provides more complete and accurate results, as we discuss in section V.

III. COMBINING TOPOLOGICAL SOURCES

In this section, we analyze and compare the information provided by the different sources on the AS level topology. The different sources have complementary information of variable accuracy. Thus, we cannot just simply take the union of all the edges. A careful synthesis and cross-validation is required. At the same time, we are interested in identifying the properties of the missing AS links.

In a nutshell, our study arrives at three major observations regarding the properties of the missing AS links: (1) most of the missing AS edges are of the peer-to-peer type, (2) most of the missing AS edges from BGP tables appear in IRR, and (3) most of the missing AS edges are incident at IXPs. At different stages of the research, these three observations direct us to discover even more edges, some of which do not appear in any other source of information currently.

We present an overview of our work in order to provide the motivation for the different steps that we take. We start from the data set from Oregon routeviews BGP table Dump (*OBD*)[23], the BGP table dumps collected at route-views.oregon-ix.net, which is by far the most widely used data archive. The work consists of four main steps.

A. BGP routing tables: We consider the AS edges derived from multiple BGP routing table dumps[7], and compare them to the Routeview data (OBD). The question we try to answer is what is the information that the new BGP tables bring. We use the term *BD* to refer to the combined data from all available BGP table Dumps. Table I lists the acronyms for our datasets.

B. IRR data: We systematically analyze the IRR data and identify topological information that seems trustworthy

TABLE III

A COLLECTION OF BGP TABLE DUMPS

Route collector or Router server name	Location or AS Num	# of Nodes	# of Edges	# of edges with type inferred			edges not in	edges not in OBD w/ type		
				total	p-p	p-c	OBD	total	p-p	p-c
Oregon Routeview route collectors										
route-views(*OBD*)	Oregon	19843	42643	42570	5551	36766	0	0	0	0
route-views2	Oregon	19837	41274	41230	4464	36514	1029	1028	835	191
route-views.eqix	Virginia	19650	34889	34876	1027	33640	674	674	530	143
route-views.linx	London	19655	37259	37246	3246	33765	2511	2511	2188	319
route-views.isc	Palo Alto	19753	36152	36139	1915	34004	784	783	663	118
route-views.wide	Tokyo	19649	26974	26963	636	26183	35	35	26	8
RIPE RIS route collectors										
rrc00.ripe	Amsterdam	19770	36479	36465	1641	34605	655	654	543	111
rrc01.ripe	London	19640	34193	34180	1121	32855	617	617	512	105
rrc02.ripe	Paris	3966	4260	4256	27	4204	3	3	0	3
rrc03.ripe	Amsterdam	19737	39147	39129	3850	35042	3233	3228	2609	616
rrc04.ripe	Geneva	19694	29840	29827	660	28997	452	451	366	85
rrc05.ripe	Vienna	19765	32676	32659	1122	31324	1095	1091	658	432
rrc07.ripe	Stockholm	19618	32811	31797	1219	30394	804	803	724	79
rrc10.ripe	Milan	19578	29225	29213	342	28704	122	122	102	20
rrc11.ripe	New York	19592	29784	29772	716	28892	441	441	377	64
rrc12.ripe	Frankfurt	19628	33841	33827	2024	31606	1611	1610	1417	193
Public route servers										
belwue.de	AS553	19637	25208	25195	194	24836	93	92	21	69
on.bb.telus.com	AS852	19551	24893	24883	129	24610	30	30	24	5
ab.bb.telus.com	AS852	19551	24894	24884	129	24610	30	30	24	5
ip.tiscali.net	AS3257	19652	24808	24798	168	24498	51	51	1	48
gblx.net	AS3549	19263	24878	24866	57	24680	54	54	2	51
eu.gblx.net	AS3549	19267	26066	26054	60	25859	55	55	2	52
savvis.net	AS3561	19210	24630	24620	31	24464	3	3	1	1
as5388.net	AS5388	19526	25049	25039	133	24785	58	58	2	55
gt.ca	AS6539	19472	25822	25981	187	25477	62	62	1	60
as6667.net	AS6667	19459	25073	25063	179	24754	83	83	15	67
sunrise.ch	AS6730	15414	19637	19627	198	19307	32	32	20	11
he.net	AS6939	19396	25661	25650	207	25314	70	70	9	60
ip.att.net	AS7018	19572	26822	26809	83	26581	67	67	0	66
optus.net.au	AS7474	19606	30045	30032	507	29363	338	337	223	113
wcg.net	AS7911	19420	25566	25556	89	25325	63	63	0	62
colt.net	AS8220	19261	27988	27975	237	27580	128	128	30	180
bmcag.net	AS9132	19461	24639	24628	128	24362	17	17	8	8
rhein-main-saar.net	AS15837	19540	26003	25991	492	25359	312	311	223	87
Total Combined (*BD*)		19950	51345	51259	12734	38265	8702	8689	7183	1499

by Nemecis[19]. We follow a conservative approach, given that IRR may contain some out-dated and/or erroneous information. We do not accept new edges from IRR, even after our first processing, unless they are confirmed by traceroutes (using our RETRO tool). Overall, we find that IRR is a good source of missing links. For example, we discover that more than 80% of the new edges that are found in the new tables (i.e., the AS edges in BD but not in OBD) already exist in IRR [26]. Even compared to BD, IRR has significantly more edges, which are validated by RETRO as we explain below.

C. IXPs and potential edges: We identify a set of potential IXP edges by using our work on IXP participants from Section V. We find that many of the peer-to-peer edges missing from the different datasets could be IXP edges.

D. Validation using RETRO: We use our traceroute tool, RETRO, to verify potential edges from IRR and IXPs. First, we confirm the existence of many potential edges we identified in the previous steps. We find that more than 94% of the RETRO verified IRR AS edges indeed go through IXPs. We also discover edges that were not previously seen in either the BGP table dumps or IRR. In total, we have validated 300% more peer-to-peer links than those in the

OBD dataset from Routeviews.

The statistics of the topologies generated from the different datasets in our study are listed in Table. II

Note that in this section we focus on identifying new edges, while we examine the significance of these new edges in modeling the Internet topology in the next section.

A. The new edges from a BGP table dump

We collect multiple BGP routing table dumps from various locations in the world, and compare them with OBD. On May 12, 2005, we collected 34 BGP routing table dumps from the Oregon route collectors [23], the RIPE/RIS route collectors [24] and public route servers [27]. Several other route collectors were not operational at the time that the data was collected and therefore, we do not include them in this study. For each BGP routing table dump, we extract its "AS_PATH" field and generate an AS topology graph. We then combine these 34 graphs into a single graph and delete duplicate AS edges if any. The resulting graph, which is named as *BD (BGP Dumps)*, has 19,950 ASes and 51,345 edges that interconnect these ASes. The statistics of *BD* are similar to what was reported in [7]. Interestingly, *BD* has only 0.5% additional ASes, but 20.4% more AS edges as compared with *OBD*.

To study the business relationships of these edges, we use the state-of-the-art algorithm PTE [21], which seems to outperform most previous such approaches. Specifically, it significantly increases the accuracy (over 90%) of inferring peer-to-peer AS links. Note that PTE is not able to infer the relationship of a small fraction (0.17%) of the edges, which could very well be due to routing anomalies. The rest of the AS edges are classified into three basic types on the basis of business relationships: provider-customer, peer-to-peer and sibling-to-sibling. Among them, sibling-to-sibling links only account for a very small (0.12%) portion of the total AS edges and we do not consider them in this study. We count the number of peer-to-peer (or "p-p" for short) and provider-customer (or "p-c" for short) AS links for each BGP routing table and the statistics are summarized in Table III.

For comparison purposes, we pick the most widely used AS graph *OBD* as our baseline graph. For each of the other BGP routing tables, we examine the number of additional AS edges that do not appear in *OBD*, as classified by their business relationship. As shown in Table III, from each of the BGP routing tables that provides a significant number of new edges to *OBD*, most of the new-found edges are of the peer-to-peer type.

BGP table biases: underestimating the peer-to-peer edges. A closer look at the data reveals an interesting dichotomy:

a. *Most edges in a BGP table are provider-customer.*

b. *Given a set of BGP tables, most new edges in an additional BGP table are peer-to-peer type.*

We can see this by plotting the types of new edges as we add the new tables. In Fig. 1, we plot the cumulative number of new found peer-to-peer edges and provider-customer edges versus the total number of edges. To generate this plot, we start with *OBD* with 42643 AS edges and combine new AS edges derived from the BGP table dumps other than *OBD*, one table dump at a time[2]. At the end, when all the BGP table dumps in our dataset are included, we obtain the graph *BD*; this has 51345 AS edges in total. Among these edges, there are 7183 peer-to-peer edges and 1499 provider-customer edges that do not exist in the baseline graph *OBD*. Clearly, Fig 1 demonstrates that we discover more peer-to-peer AS edges than provider-customer edges when we increase the number of vantage points. Furthermore, the ratio of the number of new found peer-to-peer edges to the number of new found provider-customer edges is almost constant given that the two plots (corresponding to the new found p-p edges and the p-c edges) in Fig. 1 are almost straight lines.

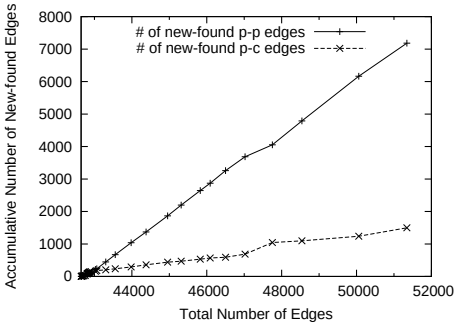

Fig. 1. Most new edges in BD but not in OBD are peer-to-peer edges.

The percentage of peer-to-peer edges increases with the number of BGP tables. A complementary observation is that for a BGP-table-based graph, the more complete it is (in number of edges), the higher the percentage of peer-to-peer links. For example, the AS graph derived from rrc12.ripe.net has 33841 AS edges, 2024 (5.98%) of which are peer-to-peer edges. On the other hand, the more complete AS graph *OBD* has 42643 edges, and 5551 (13.0%) of these edges are peer-to-peer edges. The combined graph *BD* has an even higher percentage (24.8%) of peer-to-peer links.

The above observations strongly suggest that in order to obtain a more complete Internet topology, one should consider peer-to-peer links than any other type of primary AS links.

B. Exploring IRR

We carefully process the IRR information to identify potential new edges. Recall that we do not add any edges until we verify them with RETRO later in this section.

We extract AS links from IRR on May 12, 2005 and classify their business relationships using Nemecis [19] as per the exporting policies of registered ISPs. Nemecis is the the state of the art tool for parsing and analyzing IRR data

[2]The table dumps are added in the order of the number of new edges they provide. The table dump with the least number of new edges is added first.

TABLE IV

AS EDGES IN IRR (MAY 12, 2005) WITHOUT RELATIONSHIP CONFLICT

Name of Graphs	# of non-0 degree Nodes	# of Edges	Avg Degree	Perc. of total IRR edges	Perc. of IRR edges without conflict
IRRnc	16952	89540	10.56	92.6%	100.0%
peerIRRnc	6619	49411	14.93	51.1%	55.2%
pcIRRnc	15277	37619	4.925	38.9%	42.0%
siblingIRRnc	2277	2510	2.204	2.6%	2.8%
peerIRRdual	1561	18453	23.64	19.1%	20.6%
pcIRRdual	6298	8748	2.778	9.1%	9.8%
siblingIRRdual	226	143	1.265	0.1%	0.1%

and has attracted significant interest from the NANOG/RIPE community (the network administrators)[38][39]. One of its functionalities is of particular interest here. It can successfully eliminate most badly defined or inconsistent edges and, it can infer with fair accuracy the business relationships of the edges.

There are 96,654 AS links in total and they are classified into three basic types in terms of their relationships: peer-to-peer, customer-provider and sibling-to-sibling. Sometimes two ASes register conflicting policies with each other. For example, AS_A may register AS_B as a customer while AS_B registers AS_A as a peer. There are 7,114 or 7.4% of such AS links and we exclude them in our data analysis. We call the remaining edges *non-conflicting IRR edges* or *IRRnc*. Considering the different types of policies, this set can be decomposed into three self-explanatory sets: *pcIRRnc*, *peerIRRnc* and *siblingIRRnc*. From these edges, we define the set *IRRdual* to include the edges for which both adjacent ASes register matching relationships. Similarly, the IRRdual set can be decomposed by type of edge into three sets: *pcIRRdual*, *peerIRRdual* and *siblingIRRdual*.

The statistics of these datasets are summarized in Table IV. Two major components, peer-to-peer edges and provider-customer edges account for 90.0% (51.1% + 38.9%) and 97.2% (55.2% + 42.0%) of the number of total AS edges and non-conflicting AS edges, respectively, in the IRR. The number of edges in the more reliably defined *IRRdual* set is significantly less than that of the *IRRnc*. In other words, AS edges in *IRRdual* and its subsets (*peerIRRdual*, *pcIRRdual* and *siblingIRRdual*) are fewer but we are more confident about: (a) their existence, and (b) their business relationships.

We make the following two observations:

a. IRR is a good source of hints for missing edges. We perform the following thought experiment. *If we only knew of OBD data set, would IRR be a good source of potential edges?* We compare the edges in graph *BD* but not in graph *OBD* with the edges in IRR. We find that 83.3% of these edges exist in IRR: 7251 from a total of 8702 new edges. This high percentage suggests that the IRR can potentially be a source for finding new edges. We also notice that from among these 7251 edges, 6302 are classified in terms of their business relationships by Nemecis[19]. From among these classified edges, 5303 edges are of the peer-to-peer type and only 832 are of the provider-customer type. This confirms the result[3] shown in Fig. 1, where most new found AS edges are of the peer-to-peer type.

TABLE V

PERCENTAGE OF IRR EDGES MISSING FROM *BD*

Name	# of edges	# of edges NOT in *BD*	# of edges Missing Perc.
IRRnc	89,540	63,660	71.1%
peerIRRnc	49,411	39,894	80.7%
pcIRRnc	37,619	22,466	59.7%
siblingIRRnc	2,510	1,300	51.8%

b. IRR has many more edges compared to our most complete BGP-table graph (BD). Motivated by the observation above, we examine the number of AS edges in IRR that are not included in *BD*. Table V summarizes the number and the type of IRR AS edges that do not appear in *BD*. From among the IRR AS edges inferred as non-conflicting types, 71.1% are missing from *BD*. The percentage is especially high for peer-to-peer edges: 80.7% of the peer-to-peer AS edges in IRR are missing from *BD*. This suggests that there may be many IRR links that exist but are yet to be verified. We also notice that 59.7% of the provider-customer AS edges are missing. At this point, we can only speculate that most of these missing provider-customer AS edges represent backup links.

C. IXPs and missing links

Note that, when two ASes areparticipants at the same IXP, it does not necessarily mean that there is an AS edge between them. If two participating ASes agree to exchange traffic through an IXP, this constitutes an AS edge, which we call an *IXP edge*. Many IXP edges are of peer-to-peer type, although customer-provider edges are also established.

Identifying IXP edges requires two steps: (a) we need to find the IXP participants, and (b) we need to identify which edges exist between the participants. The first part is addressed in Section V. Identifying the edges is also a challenge. First, not all participants connect with each other. Second, the peering agreements among the IXP participants are not publicly known.

We start with a superset of the real IXP edges that contains all possible IXP edges: we initially assume that the participants of each IXP form a clique. We denote by *IXPall* the set of all edges that make up all of these cliques. *IXPall* contains 141,865 distinct[4] AS edges.

Potential missing edges and IXP edges. We revisit the previous sets of edges we have identified and check to see if they could be IXP edges. First, we look at the peer-to-peer

[3] Recall that, for Fig. 1, the business relationships are inferred by the PTE algorithm[21], instead of Nemecis[19] which we use here. Both algorithms seem to give similar results which provides high credibility to both the data and the interpretations.

[4]Sometimes two ASes have the ability to peer at multiple IXPs. We count such IXP edges as one AS edge, since at the AS level topology, these edges appear as a single AS edge.

TABLE VI

MANY MISSING PEER-TO-PEER LINKS ARE AT IXPS

Name	# of Edges	$\bigcap$ *IXPall*	Perc.
peerBD-OBD	7183	6197	86%
peerIRRnc-BD	39894	23979	60%
peerIRRdual-BD	13905	11477	83%
BD-OBD	8702	6910	79%

AS edges that appear in *BD* but not in *OBD*. These are the peer-to-peer AS edges missing from *OBD* but are discovered with *BD*. We call this set of AS edges *peerBD-OBD* [5]. Second, we look at the AS edges that appear in *peerIRRnc* but not in the graph *BD*. We call this set of links *peerIRRnc-BD*. These AS links are the ones that are potentially missing from *BD*. We define the *peerIRRdual* links not in *BD* as *peerIRRdual-BD*.

Having made this classification, we compare each class with the super set, *IXPall*, of edges that we constructed earlier. The statistics are shown in Table VI. With our first comparison, we find that approximately 86% of the edges in *peerBD-OBD* are in *IXPall* and hence, are potentially IXP edges. Next, we observe that 60% of the edges in *peerIRRnc-BD* and 83% of the edges in *peerIRRdual-BD* are in *IXPall*. Thus, if they exist, they could be IXP edges.

In summary, the analysis here seems to suggest that, most of the peer-to-peer AS links missing from the BGP dumps but present in IRR are potentially IXP edges.

D. Validating links with RETRO

With the work so far, we have identified sets of edges and obtained hints on where to look for new edges: (1) most missing links are expected to be the peer-to-peer type, (2) IRR seems to be a good source of information, (3) many missing edges are expected to be IXP edges.

Here, we focus on finding and validating the peer-to-peer links missing from the BGP table dumps. *Note here that with this method, we eliminate stale information that may still be present in some of the considered data sources.*

To verify potential edges, we use the RETRO tool, which we describe briefly.

RETRO: a large-scale traceroute capability. We develop an essential tool for detecting and verifying AS edges. By employing public traceroute servers collected in [40] and [41], we construct RETRO (REverse TraceROute), a convenient tool that can collect traceroute server configurations, send out traceroute requests, and collect/parse traceroute results dynamically. Currently, we have a total of 404 reverse traceroute servers which contain more than 1200 distinct and working vantage points. These vantages points cover 348 different ASes and 55 different countries.

One challenge in using public traceroute servers is that the interface configuration of the traceroute servers keeps changing. Examples of such changes are the inclusion of additional required fields or the upgrading/renaming of the back-end server program. Although the rate at which such changes occur is small, when the number of traceroute servers is large, changes are seen quite often. To overcome this, we construct a smart self-adjustable front end that can parse traceroute server interfaces at the time when traceroutes are triggered. The front end employs a number of heuristics that captures traceroute servers' essential configurations dynamically on the fly. Using our tool, we can automatically identify the interface and parse over 90% of the traceroute servers.

To verify the existence of the edges in *peerIRRnc-BD*, we would like to witness these edges on traceroute paths. Typically, when a traceroute probe passes through an IXP edge between AS A and AS B, it will contain the following sequence of IP addresses: $[IP_{AS\_A}, IP_{IXP}, IP_{AS\_B}]$. If such a pattern is observed with our traceroute probes, it is almost certain that an IXP edge between AS A and AS B exists.

Using Skitter: We first used the Skitter[28] traces as our verification source; however, we found that it was not suitable for our purposes. Between May 8 and May 12 in 2005, we collected a full cycle of traces from each of the active Skitter monitors. Despite a total number of 21,363,562 individual traceroute probes in the dataset, we were only able to confirm 399 IXP edges in *peerIRRnc-BD*. The reason could be that the monitors were not in the "right" place to discover these edges: the monitors should be at the AS adjacent to that edge, or at one of the customers of those two ASes. With the limited number of monitors (approximately two dozen active ones) in Skitter, it is difficult to witness and validate many of the peer-to-peer AS edges.

Using RETRO: With the RETRO tool, we conduct the following procedure to verify AS edges in the *peerIRRnc-BD* set. For each edge in *peerIRRnc-BD*, we find out if there are any RETRO monitors in at least one of the two ASes incident on the edge. For about 2/3 of the edges in *peerIRRnc-BD*, we do not have a monitor in either of the two ASes on the edge. If there is at least onemonitor, we try to traceroute from that monitor to an IP that belongs to the other AS on the edge. There are two problems in finding the right IP address to traceroute to. First, some ASes do not announce or can not be associated with any IP prefixes and thus, we are not able to traceroute to these ASes. Second, most of the rest of the ASes announce a large range (equal to or more than 256 i.e., a full /24 block) of IP addresses. To maximize our chances of performing a successful traceroute, we choose a destination from the list of IP addresses that has been shown to be reachable by at least one of the Skitter monitors. We then trigger RETRO to generate a traceroute from the selected monitor to the destination IP address that we choose. To avoid overloading of the public traceroute servers, we enforce a minimum of 900 seconds between any two traceroutes from the same monitor.

Most missing peer-to-peer links are incident at IXPs. We define a *candidate* to be a potential edge between two ASes, which satisfy the following two conditions: (a) we

[5]Here we use the minus sign to denote the difference between two sets. For example, *A-B* is the set of entities that appears in set *A* but not in set *B*.

TABLE VII

RETRO VERIFIES PEER-TO-PEER LINKS IN IRR MISSING FROM BD

Name	# of Edges	# of RETRO candidates	# of confirmed peering		
			total	via IXP	direct
peerIRRnc-BD	39894	8791	5646	5317	329
peerIRRdual-BD	13905	4487	3529	3351	178

TABLE VIII

RETRO VERIFIES AS EDGES NOT IN *BD* AND *IRRnc*

Name	# of Edges	# of RETRO candidates	# of confirmed peering		
			total	via IXP	direct
IXPall-BD-IRR	100,076	17,640	2,603	2,407	196

have a RETRO monitor located in one of the two ASes, and (b) there is at least one IP address from the other AS is reachable by the traceroute probe performed from the RETRO monitor. We have 8791 such "candidates" for the potential AS edges in *peerIRRnc-BD*. By appropriately performing traceroutes on candidates, we get traceroute paths.In these paths, we search for two patterns for each candidate (AS_A, AS_B): (a) [IP_{AS_A}, IP_{AS_B}]. , and (b) [IP_{AS_A}, IP_{IXP}, IP_{AS_B}]. If either of the two patterns appears, it is almost certain that the AS edge between AS_A and AS_B exists either as (a) a direct edge or, (b) as an IXP edge, respectively. The results that we obtain at the end of the above process are summarized in Table VII.

Among 8791 candidates in *peerIRRnc-BD*, RETRO is able to confirm that a total of 5646 edges indeed exist. Note that this method can only confirm the existence, but not prove the absence of an edge. It could very well be that that the traceroute does not pass through the right path.[6] The most interesting result is, from among the 5646 verified edges, 5317 or 94.2% of them are IXP edges. The result suggests that most of the missing peer-to-peer links from BGP tables are in fact incident at IXPs. Our result strongly suggests that in order to look for missing peer-to-peer links from BGP tables, we should examine IXPs more carefully.

Missing edges not observed anywhere else. From the results so far, we suspect that the missing edges are often IXP edges. Following this pattern, we identify and confirm edges that previously had not been observed in any other data source.

We consider those AS edges in *IXPall* that are neither in *BD* nor in *IRRnc*, and call them *IXPall-BD-IRR*. We then attempt to trace these edges by using RETRO. The results from our experiments are summarized in Table VIII.

We find 2,603 new AS edges from out of 17,640 RETRO candidate paths. The percentage of confirmed new AS edges is 14.8%. This is much lower than what we see with *peerIRRnc-BD*. This is due to the fact that IXPall is an overly aggressive estimate. In addition, we have already identified that many edges from IXPall are in the previous sets (*BD* and *peerIRRnc-BD*).

[6]The number of traceroute sources, and our inability to direct traceroute probes (e.g. unsupported loose-source routing, hot-potato routing) constrain the capabilities of our effort.

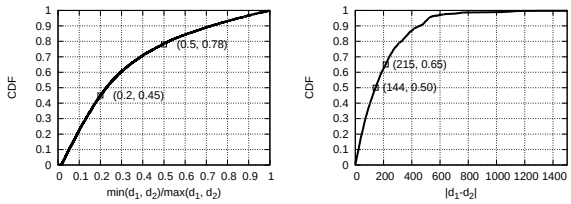

Fig. 2. Degree ratio distribution(left) and degree difference distribution (right) of all peer-to-peer AS links in the Internet.

We also notice that there is a small number of confirmed edges that are shown to exhibit direct peering instead of peering at some IXP. A closer look reveals that many of such cases are because of the fact that a small number of routers do not respond with an ICMP message with the incoming interface, and therefore, the IXP IP address, which is supposed to be returned by the traceroute, is "skipped". Note that this phenomenon does not stop us from identifying the edge. It just makes us underestimate the percentage of IXP edges from among the confirmed edges.

IV. PROPERTIES AND SIGNIFICANCE OF THE NEW EDGES

In this section, we identify properties of the new edges. Then, we examine the impact of the new edges on the topological properties of the Internet. Finally, we attempt to extrapolate and estimate how many edges we may still be missing.

A. Patterns of the peer-to-peer edges

We want to examine in more detail, which nodes end up peering with each other. Therefore, we examine the degrees, d_1 and d_2, of the two peering nodes that make up each peer-to-peer edge. Let us clarify that the degrees d_1 and d_2 include both peer-to-peer and provider-customer edges. One would expect that d_1 and d_2 would be "comparable". Intuitively, the degree of an AS is *loosely* related to the importance and its place in the AS hierarchy; we expect ASes to peer with ASes at the same level.

The node degree of the nodes spanning a peer-to-peer link can differ significantly. We compare the two degrees using their: (a) ratio, (b) absolute difference. Note that these two metrics provide complementary view of difference, which leads to the following two findings:

(1) Close to 78% of the peer-to-peer edges connect ASes whose degrees differ by a factor of 2. In Fig. 2 (left), we plot the CDF of the distribution of the ratio $min(d_1, d_2)/max(d_1, d_2)$ of the peer-to-peer edges. Another observation is that 45% of the peer-to-peer edges connect nodes whose degrees differ by a factor of 5. This is a surprisingly large difference. One could argue that if this could be an artifact of having peer-to-peer edges between low degree nodes, say $d_1 = 2$ and $d_2 = 11$, whose absolute degree difference is arguably small. This is why we examine the absolute difference of the degrees next.

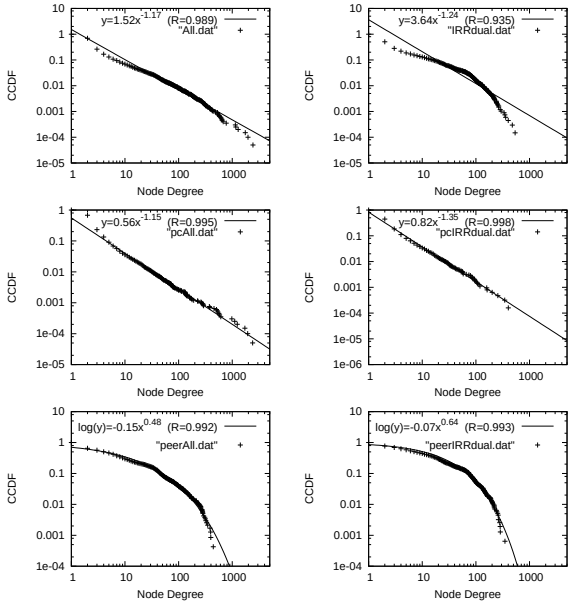

Fig. 3. The degree distributions of *ALL* (left) and *IRRdual* (right) in the top row, their provider-customer degree distributions in the middle row, and their peer-to-peer degree distributions in the bottom row.

(2) 35% of the peer-to-peer edges have nodes with an absolute difference greater than 215. In Fig. 2 (right), we plot the CDF of the distribution of the absolute value $|d_1 - d_2|$, where d_1 and d_2 remain as defined earlier. Another interesting observation is that approximately half of the peer-to-peer edges have a degree difference larger than 144. Differences of 144 and 215 are fairly large if we consider that roughly 70% of the nodes have a degree less than 4.

B. Impact on modeling the Internet topology

We study the effect of the peer-to-peer edges on some commonly used Internet properties. Among all the properties that we examined, we show the ones that lead to the most interesting observations.

1) The degree distribution: There has been a long debate on whether the degree distribution of the Internet at the AS level follows a power-law[42][43][44][6]. This debate is partly due to the absence of a definitive statistical test. For example, in Fig. 3 top left, we plot the complementary cumulative distribution functions (CCDF), on a log-log scale, of the graph *ALL* defined earlier in Table I The distribution is highly skewed, and the correlation coefficient of a least square errors fitting is 98.9%. However, one could still use different statistical metrics and argue against the accuracy of the approximation [44].

Furthermore, the answer could vary depending on which source we think is more complete and accurate, and the purpose or the required level of statistical confidence of a study. For example, if we go with *IRRdual*, which is a subset of the AS edges recorded in IRR filtered by Nemesis, the correlation coefficient is only 93.5%, see Fig. 3 top right.

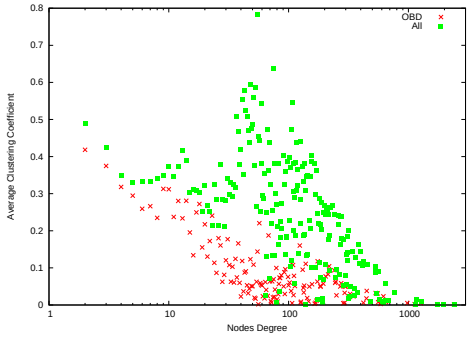

Fig. 4. The per-degree average clustering coefficient $\overline{\gamma_d}$ versus the degree for graphs *ALL* and *OBD*.

To settle the debate, we propose a reconciliatory divide-and-conquer approach. We propose to model separately the degree distribution according to the type of the edges: provider-customer and peer-to-peer. We argue that this would be a more constructive approach for modeling purposes. This decomposition seems to echo the distinct properties of the two edge types, as discussed in a recent study of the evolution on the Internet topology [11].

In Fig. 3, we show an indicative set of degree distribution plots for graph *ALL on the left column and* IRRdual on the right. We show the distributions for the whole graph (top row), the provide-customer edges only (middle row), and the peer-to-peer edges only (bottom row). We display the power-law approximation in the first two rows of plots and the Weibull approximation in the bottom row of plots.

We observe the following two properties:

(a) The provider-customer-only degree distribution can be accurately approximated by a power-law. The correlation coefficient is 99.5% or higher in the plots of Fig.3 in the middle row.

Note that, although the combined degree distribution of *IRRdual* does not follow a power law (top row right), its provider-customer subgraph follows a strict power law (middle row right).

(b) The peer-to-peer-only degree distribution can be accurately approximated by a Weibull distribution. The correlation coefficient is 99.2% or higher in the plots of Fig.3 in the bottom row.

Why do the two distributions differ? We suggest the following explanation. Power-laws are related to the rich-get-richer behavior: low degree nodes "want" to connect to high degree nodes. For provider-customer edges, this makes sense: an AS wants to connect to a high-degree provider, since that provider would likely provide shorter paths to other ASes. For peer-to-peer edges, things are different. If AS1 becomes a peer of AS2, AS1 does not benefit from the peer-to-peer edges of AS2: a peer will not transit traffic for a peer. Therefore, high peer-to-peer degree does not make a node more attractive as a peer-to-peer neighbor. We intend to investigate the validity of this explanation in the future.

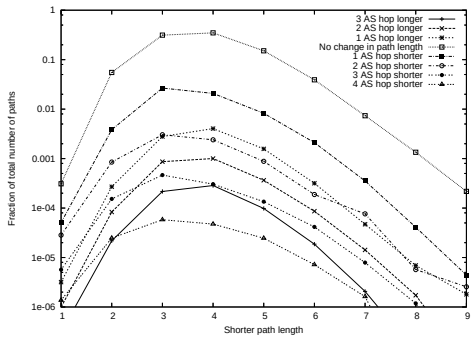

Fig. 5. The effect of the new links on the path length: fraction of the number of paths that change length versus the length of the shorter path, with each line representing a length change.

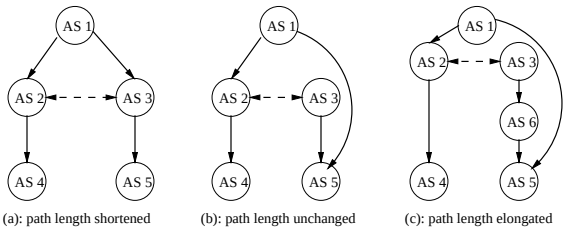

Fig. 6. The effect of adding a peer-to-peer link between AS 2 and AS 3 on the path from AS 4 to AS 5. The arrow points from the provider to the customer.

2) Clustering coefficient: We expect that the *ALL* graph will be more clustered since we add edges. To quantify this, we use the *clustering coefficient* which has been used to characterize and compare generated and real topologies [33]. Intuitively, the clustering coefficient captures how tightly connected is the one hop neighborhood of a node. For a node v_i with $n_i > 1$ neighbors, the clustering coefficient of v_i is $\gamma_i = \frac{m}{m_{max}}$, where $m_{max} = \frac{n_i(n_i-1)}{2}$, and m the number of edges between these neighbors. A clustering coefficient of exactly one means that the neighborhood is a clique. The average clustering coefficient of *OBD* is 0.25 and it increases to 0.31 in *ALL*.

The neighborhoods of "middle-class" nodes become more clustered. We find that the density increase is not homogeneous. We use $\overline{\gamma_d}$ to denote the *average clustering coefficient* of all nodes with degree d. In Fig. 4, we plot $\overline{\gamma_d}$ versus the node degree d, for two graphs: *ALL* and *OBD*. The thing to note is that the clustering coefficient increase is larger for nodes with degrees in the 10 to 300 range.

Note that this property characterizes the new edges, and could help us identify more missing edges in future studies.

3) AS path length: We study the effect of the new edges on the AS path lengths with policy-aware routing. The routing policy is a consequence of the business practices driven by contracts, agreements, and ultimately profit. We use the *No Valley Prefer Customer (NVPC)* routing, whose definition we quote [45] [46]:

No Valley: Peers and customers do not generally provide transit. Without paying customers, there is no incentive for an ISP to carry traffic.

Prefer Customer: Paths through customers are preferred to those through peers, which are in turn preferred to those via providers. Paths through customers generally represent income; through providers, an expense; and through peers, no (or little) cost.

We have approximately 20,000 ASes present in the Internet topology and examine all possible pairs of ASes. For each AS pair, we compare the AS path lengths with *OBD* and with *ALL*. We group those AS paths with the same shorter path length, and show their path length changes in Fig. 5. We find that approximately 10 million of the paths change in length. While we note that this is a small fraction of the total number of paths, it is still a significant number in terms of its absolute value. In addition, *no change in the length does not mean that the path did not change* as we discuss later.

One really interesting observation is that, by discovering the new edges, some of the new paths become in fact longer than before! This would never happen if the routing policy was based on the minimum hop criterion. Here, the length increase is due to the routing policy, which is based on the business relationship types. In other words, an AS will prefer a longer path through a peer than a shorter path through its provider. [7] In Fig. 6, we show all possible changes that a new peer-to-peer edge (AS 2 - AS 3) can cause to a path (from AS 4 to AS 5): (a) shorten the path, (b) change the path but maintain the same length, and (c) elongate the path. In Fig. 6 (c), AS 2 prefers to route through its peer AS 3 rather than through its provider AS 1 according the No Valley Prefer Customer routing.

In the next subsection, we study how many paths changed independently of whether they changed in length.

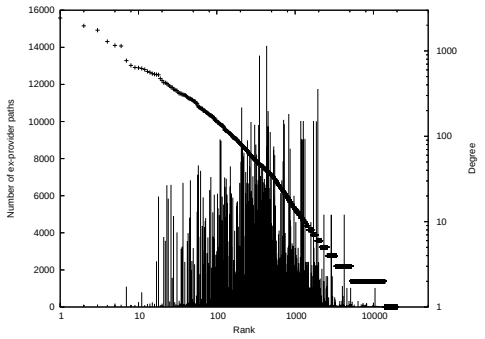

Fig. 7. The number of *ex-provider paths* (shown as impulses on the left y-axis) of each node in order decreasing node degree (shown as a semi diagonal line corresponding to the right y-axis). The x-axis shows the rank of the nodes in the order of descending degree.

[7] In practice, this is done by setting higher *local_pref* value to peer links than to provider AS links. When multiple BGP paths to a prefix are available, BGP will first choose the route with the highest *local_pref* value. [47]

C. The effect of the new edges on ISP revenue

We examine the effect of the new edges on the routing decisions of individual ASes. We consider a common business practice: an AS pays for sending traffic through its provider. We also consider the use of NVPC routing. How much do the new edges affect these decisions and thus, ISP income?

For each AS, we count how many of its paths stop going through one of its providers once the new edges are added. We refer to these paths as *ex-provider paths*. The number of ex-provider paths is an indication, of the financial gains for that AS. Clearly, there are other considerations, such as prefix-based traffic engineering and performance issues, that our analysis cannot possibly capture. However, our results are a good first indication of the effect of the new peer-to-peer links.

The significant financial benefits of the new peer-to-peer edges. We plot the number of the *ex-provider* paths for each node in Fig. 7. The x-axis represents the rank of the nodes on a log scale in order of decreasing degree; The y-axis at the left represents the number of ex-provider paths. In addition, we plot the node degrees (on the right y-axis) against their ranks as a semi diagonal line. We see that the effect on the routing on individual ASes is dramatic: there are many ASes, for each of which, several thousands out of the total 20K paths (to all other ASes) stop going through a provider. For some ASes, more than 50% of their paths stop going through their providers (10K out of 20K possible paths per AS).

The rise of the "middle class" ASes. Another interesting observation is that the nodes which seem to benefit the most from these changes have degrees in the range from 10 to 300 (right y-axis). Top tier nodes (top 20 ranked) almost do not benefit at all; this is expected, since they do not have any providers anyway. Nodes with really low node degree do not benefit much either, since nodes with very low degrees are less likely to have a peer-to-peer edge.

D. Are we missing a lot more peer edges?

Currently, the *ALL* graph has approximately 20.9K peer-to-peer edges. However, we were very conservative in adding edges from *IRRnc*: we required that the edges are verified by RETRO. So, a natural question is, how many more edges could we verify from *IRRnc* if we had more RETRO servers? We attempt to provide an estimate by extrapolating the success of our method in finding new edges. First, we provide a conservative estimation and later, a more liberal estimation, below.

Conservative extrapolation using IRRdual: We find 35% more peer-to-peer edges compared to ALL. We revisit the *IRRdual* graph and examine if we can include more edges than the ones we validate with RETRO. Recall from Table VII that we find that there are 13905 edges in the *peerIRRdual-BD*, and from these, only 4487 are "verifiable" candidates. From the verifiable edges, we actually verify 3529 or 78.6% of the verifiable edges. We generalize this percentage: we assume that if we had more RETRO monitors, we could verify 78.6% of the *peerIRRdual-BD*. This leads to an estimated 7.4K (10.9K−3.5K) peer-to-peer edges not in ALL, which has 20.9K peer-to-peer edges.

Liberal Extrapolation using IRRnc: We find 95% more peer-to-peer edges compared to ALL. In a similar way, we estimate how many edges we could verify from *peerIRRnc-BD*, which is a more "inclusive" set. Here, the total number of peer-to-peer edges is 39,894, the verifiable edges 8,791, and the verified edges 5,646. This gives rise to an estimate of $39894 \times 5646/8791$ = 25.6K peer-to-peer edges out of which 5.6K are already in ALL.

V. Identifying IXP Participants

In this section, we present a method for identifying the *participants* at Internet Exchange Points (IXPs). Our goal is to find all the participants at each IXP, and this is a non trivial problem [8]. We find that knowing the IXP participants is key for identifying many missing AS edges as explained in section III.

Our proposed method is arguably the best method to date. We show that our method outperforms the previous state of the art method, which we refer to as **XDZC** after the initials of the authors [14] (see Table IX). We not only combine previous methods, but we also introduce important improvements and new techniques.

Our approach consists of two complementary mechanisms: (A) a technique to infer IXP participants using the IXP's IP addresses, and (B) an automated tool to parse and retrieve public archival information.

A. Inferring participants from the IP addresses of IXPs

This part of our approach uses two techniques to infer IXP participants from IXP IP addresses: 1) **path-based inference**, where we perform a careful processing of collected traceroute data, and 2) **name-based inference**, where, we analyze the name and the related information with regard to IXPs from the DNS and/or WHOIS databases.

In both inference methods, we start with the IP address blocks allocated to the IXPs, which we call *IXP IP addresses*. We obtain this information from the Packet Clearing House (PCH) [29]. In terms of traceroute data, we use a full cycle of Skitter traceroute data between May 1, 2005 and May 12, 2005, and our RETRO traceroute data in May 2005.

1) **Path-based inference:** The high level overview of the method is deceptively simple. First, for each IXP IP address IP_{ixp} that we obtain from PCH, we search for the IP address that appears immediately after IP_{ixp} in each of the obtained traceroute paths. Second, if we find more than one such IP address for the particular IP_{ixp}, we select the one that appears most to be IP_{next}. We call the above procedure the *majority selection process*. Third, we find the AS ASx that owns the IP address, IP_{next} and consider that ASx to be a

[8] Efforts in improving IP-to-AS mapping try to identify IXP IPs, rather than the participant ASes. Their goal is different: they only need to find whether an observed IP address belongs to an IXP or not.[15][16].

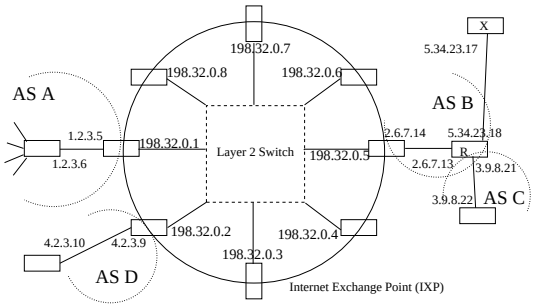

Fig. 8. A conceptual model of a typical IXP

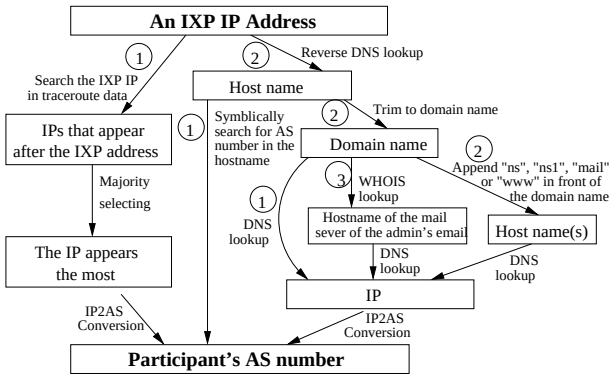

Fig. 9. The flow chart of our path-based method to infer IXP participants from IXP IP addresses. Starting from the top, the numbers in the circle indicate the priority (lowest number with highest priority) at a branching point.

participant at the IXP. Furthermore, we consider that IP_{ixp} is the IP interface via which ASx accesses the IXP.

To illustrate this with an example let us consider Fig. 8. A typical traceroute from AS A to router X yields the following sequence of IP addresses: [1.2.3.5, 198.32.0.5, 2.6.7.13, 5.34.23.17]. Since the address "2.6.7.13", which belongs to AS B, appears immediately after IXP IP address "198.32.0.5", we infer that, AS B is a participant AS, and that 198.32.0.5 is the interface that is assigned to AS B. Note from Fig. 8 that, irrespective of the location of the traceroute source and its destination, if an IXP address (the address 198.32.0.5 in our example) appears in a traceroute, the IP address that appears immediately after (the address 2.6.7.13 in our example) is owned by the AS (in our example AS B) that uses the IXP address (e.g. 198.32.0.5) to access the IXP as long as two conditions hold. These are: (1) each IXP interface address is assigned to a single AS, and (2) routers *always* respond to a traceroute probe with the address that corresponds to the incoming IP interface[9]. While the first condition largely holds, the second condition does not. There is a small chance that a router could respond to a traceroute probe with an alternate (not the incoming) interface[48][15]. In our example, router R could respond to a traceroute probe from AS A to router X with an alternate interface (e.g. 3.9.8.21), which makes the traceroute path appear as [1.2.3.5, 198.32.0.5, 3.9.8.21, 5.34.23.17]. Since 3.9.8.21 could be within the IP space of AS C, one could incorrectly infer that AS C is an IXP participant. We overcome this limitation with our *majority-selection process*; the basis is the assumption that *in the majority of the cases*, routers will repond to a traceroute probe with the incoming interface. This assumption has been shown to hold by numerous prior efforts [48][15][49].

The previously proposed method in [14] does not have the majority selection process. Furthermore the method does not associate the specific IXP IP interface addresses with their respective participating ASes. Our majority selection process eliminates measurement noise and thus, ensures a lower "false positive" rate. We map the discoverd AS participants to their assigned IXP IP addresses, and using this, exclude the addresses in the name-based inference process that we describe below. This practice reduces the number of total IXP IP addresses that are subject to the name-based inference procedures which are inherently less reliable, and thus reduces the possible errors overall.

[9]The *incoming interface* of a traceroute probe is the IP interface via which the probe enters the router.

TABLE IX

IXP PARTICIPANTS INFERRING COMPARISON

Name of IXP	Actual partici-pants	XDZC Approach [14]				Our Approach			
		correctly inferred	total inferred	$\mathcal{R}$	$\mathcal{P}$	correctly inferred	total inferred	$\mathcal{R}$	$\mathcal{P}$
MSK-IX	**154**	**90**	115	68%	90%	**136**	156	88%	87%
JPIX	**110**	**58**	82	53%	71%	**107**	128	97%	84%
FREEIX	**101**	**38**	39	38%	97%	**64**	65	63%	98%
AMS-IX	**211**	**177**	220	84%	80%	**182**	200	86%	91%
LINX	**175**	**164**	242	94%	68%	**168**	193	96%	87%
DE-CIX	**144**	**111**	124	77%	90%	**137**	142	95%	96%

2) **Named-based IXP participants inference.:** The basic name-based IXP participants inference method, which was proposed in [14], works in three main steps: (a) for every IP address in each IXP prefix space, we do a reverse DNS look up, and we find the host name for that IXP IP address, (b) we take the domain name part (*company.*{*com,net,org, etc.*}) from the host name, and do a DNS look up, which leads to a new IP address, and (c) we find the AS that owns this address, and this AS is considered a participant of that IXP. For example, IXP DE-CIX has the IP address 80.81.192.186. If we do a reverse DNS lookup, we get the host name "GigabitEthernet3-2.core1.ftf1.level3.net". A DNS lookup of the domain name "level3.net" yields an IP address of 209.245.19.41. An IP address to AS number conversion reveals that the IP address belongs to AS3356 (Level3). Therefore, AS3356 is considered a participant at DE-CIX.

Although this method has been used successfully by previous studies [14], it has two limitations: (a) sometimes it can return incorrect AS numbers for IXP participants, and (b) it does not always work: the DNS or the reverse DNS lookup may not return an answer.

We address the first limitation by excluding the IXP addresses that have been mapped on to AS participants by

our path-based inference method. This greatly reduces the number of IXP addresses that are to be examined by the named-based inference method and therefore reduces the possible number of erroneous results.

We address the second limitation by proposing three new methods to improve the successful rate of name-based inference:

a. Examining host names containing AS numbers. Sometimes, the DNS name of an IXP IP address contains the AS number of an IXP participant. For example, 195.66.224.71 is an IP address at the London Internet Exchange (LINX), which has a DNS name fe-3-4-cr2.sov.as9153.net. From that, we can infer that AS9153 is a participant at the LINX IXP.

b. Examining common naming practices. We can increase the success rate of DNS lookups by including common host names with the inferred domain names. For example, although *company.net* may fail to be resolved, the DNS look up may succeed with *ns.company.net*. In fact, there are several common hostnames such as "ns", "ns1", "mail" and "www". Hosts with these names *usually*[10] belong to the same AS. For example, 195.66.226.104 is an IP address at IXP LINX at London, England. The host name of that IP address is "linx-gw4.vbc.net" and the DNS lookup for the domain name "vbc.net" is unsuccessful. However, the DNS lookup for ns.vbc.net returns the address 194.207.0.129, which belongs to AS8785 (Astra/Eu-X and VBCnet GB).

c. Using the administrating personnel information. A WHOIS lookup for a domain name often has an administrative/technical contact person's e-mail address. The mail server is often within the same AS that corresponds to the domain name. For example, for "decix-gw.f.de.bcc-ip.net", all DNS lookups described previously, fail. However, if we look at the WHOIS lookup for domain "bcc-ip.net", we will find the contact email server is "bcc.de", which has an IP address of 212.68.64.114, and it belongs to AS9066 (BCC GmbH).

3) **Putting the two techniques together:** We integrate both the path-based and named-based techniques, into a tool for inferring IXP participants from IXP addresses. We start with the path-based technique, and for every IP address in the IP block of an IXP, we try to find it in a traceroute path. If this works, then we do not reexamine this IP address. Otherwise, we use the name-based inference and we utilize the three mechanisms that we proposed above. For completeness, we show the flow chart of the inference method in Fig. 9.

4) **Evaluating our inference approach:** We show the effectiveness of our inference approach using IXPs for which we know the participants. We also show that our approach outperforms the best method so far (XDZC) [14].

We use two complementary metrics: **Recall** $\mathcal{R}$ and **Precision** $\mathcal{P}$, which are widely used in the data mining literature for similar tasks. They are defined as follows: $\mathcal{R} = \frac{\mathcal{N}_{correct}}{\mathcal{N}_{actual}}$ and $\mathcal{P} = \frac{\mathcal{N}_{correct}}{\mathcal{N}_{inferred}}$ where $\mathcal{N}_{correct}$ is the number of correctly inferred participants from among those inferred, $\mathcal{N}_{actual}$ is the actual number of participants, and $\mathcal{N}_{inferred}$ is the total number of inferred participants. Note that the Precision metric, $\mathcal{P}$, has not been used in previous studies although it is critical for detecting false positives. Otherwise, we favor overly aggressive inference methods that suggest a large number of correct and incorrect participants.

For the comparison and for lack of a better criterion, we select the six largest IXPs (in terms of number of participants) for which we know the participants through the EURO-IX site [50] or the IXPs' own web sites. In Table IX, for each IXP, we list its actual number of participants, the number of ASes that our algorithm inferred, and the number of ASes that our algorithm inferred *correctly*. We also show the Recall and Precision metrics.

It is easy to see that: (a) our approach is very effective in determining most of the participants in these IXPs, and (b) our approach identifies correctly more participants than XDZC and almost always with better Precision. For the case of MSK-IX, we only have slightly lower Precision (by 3%) but a significantly higher Recall (by 20%).

B. Web-based archival inference

We notice there are some limitations on inferring IXP participants by the IXP IP addresses alone. For example, some IXPs do not have globally routable IP addresses and some IP addresses are either invisible by traceroute or appear as "*"s in responses to traceroute probes.

To overcome these limitations, we include an additional source of information by retrieving IXP participant information from the web sites. We have developed a tool that automatically downloads and parses the web pages, and outputs the AS numbers of the participants periodically. We use the European Internet Exchanges Association (EURO-IX)[50] which maintains a database with 35 IXPs and their participants. We are also able to collect information from the web pages of 31 other IXPs. Naturally, as any manually-maintained data, these archives can also contain inaccuracies. However, we did not find any major inconsistencies with our measured data.

C. The combined results

We applied our methods to infer the participants at various IXPs on May 12, 2005. We first use our web-based archival inference. For the rest of the IXPs, we collect information with regard to their IP address blocks from Packet Clearing House [29], and infer their participants from their IXP IP addresses by using our inferring heuristics. We identify 2348 distinct participants at 110 IXPs. Some ASes actively participate in multiple IXPs. For example, AS 8220 (Colt Telecom) is inferred as a participant in 22 different IXPs in 15 different countries. In this study, we have used the combined results as our source of IXP data.

VI. Conclusion

In nutshell, our work makes a stand for the underrepresented peer-to-peer edges in current Internet models. We find

[10]Some web hosting and caching centers are exceptions.

approximately 300% more such edges (20.9K from 5.5K in the commonly used OBD). We also show that these edges change several properties of the Internet model in a non trivial way.

As our first contribution, we develop several methods and tools that could be of independent interest in measurement studies: (a) we develop an improved method to identify IXP participants, and the edges that appear there, (b) we conduct a cross-validation and synthesis of most available sources of topological information including IRR and IXPs, (c) we develop the RETRO tool, and use this to verify edges before accepting them in our Internet topology.

Our main results are summarized below:

Patterns of the missing links. We arrive at three major observations, which can guide us in the search for more missing AS links: (a) most of the missing AS edges are of the peer-to-peer type, (b) many of the AS edges missing from the BGP tables appear in IRR, and (c) most of the missing AS edges are incident at IXP.

The impact of the new edges. The connectivity of the topology changes significantly especially for nodes in the middle-class (with 10 - 300 degree). More importantly, the peer-to-peer edges affect dramatically the routing decision of middle-class ASes (for some ASes, more than 50% of paths stop going through their providers) in policy based routing (NVPC). Therefore, business-oriented studies of the Internet should make a point of taking into consideration as many as possible peer-to-peer edges.

More peer-to-peer edges may exist. We find that a non trivial number of peer-to-peer edges may still be hiding. In conservative estimates, these edges could be as high as 35% in additional to what we already have in our most complete graph. In this study, the limiting factor is the availability of vantage points, such as traceroute servers and BGP tables.

Finally, our methodology can be seen as a ready-to-use tool: given more data and monitors, it can find more missing edges.

REFERENCES

[1] S. Floyd and V. Paxson. Difficulties in simulating the Internet. *IEEE Transaction on Networking*, Aug 2001.

[2] L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica. Hlp: A next-generation interdomain routing protocol. In *ACM Sigcomm*, 2005.

[3] O. Maennel and A. Feldmann. Realistic bgp traffic for test labs. In *ACM Sigcomm*, 2002.

[4] K. Park and H. Lee. On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law Internets. In *ACM Sigcomm*, Aug 2001.

[5] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *IEEE infocom*, 2005.

[6] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Towards capturing representative as-level internet topologies. *Computer Networks*, 44(6):737–755, 2004.

[7] B. Zhang, R. Liu, D. Massey, and L. Zhang. Collecting the internet as-level topology. *ACM SIGCOMM Computer Communication Review(CCR)*, January 2005.

[8] Yuval Shavitt and Eran Shir. Dimes: Let the internet measure itself. *ACM SIGCOMM Computer Communication Review (CCR)*, October 2005.

[9] X. Dimitropoulos, D. Krioukov, and G. Riley. Revisiting internet as-level topology discovery. In *PAM*, 2005.

[10] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, X. Dimitropoulos, kc claffy, and A. Vahdat. The internet as-level topology: Three data sources and one definitive metric. *ACM SIGCOMM Computer Communication Review (CCR)*, January 2006.

[11] H. Chang, S. Jamin, and W. Willinger. To peer or not to peer: Modeling the evolution of the internet's as topology. In *IEEE Infocom*, 2006.

[12] L. Colitti, G. Di Battista, M. Patrignani, M. Pissonia, and M. Rimondini. Active bgp probing. Technical report, University of Rome 3, 2005.

[13] D. Raz and R. Cohen. The internet dark matter –on the missing links in the as connectivity map. Technical Report technical report LCCN-2004-01, Technion, Israel Institute of Technology, 2004.

[14] K. Xu, Z. Duan, Z. Zhang, and J. Chandrashekar. On properties of internet exchange points and their impact on as tolology and relationship. In *Networking*, 2004.

[15] Z. Mao, J. Rexford, J. Wang, and R. Katz. Towards an accurate as-level traceroute tool. In *Sigcomm*, 2003.

[16] Z. Mao, D. Johnson, J. Rexford, J. Wang, and R. Katz. Scalable and accurate identification of as-level forwarding paths. In *Infocom*, 2004.

[17] M. Crovella A. Lakhina, J. W. Byers and I. Matta. Sampling biases in ip topology measurements. In *IEEE Infocom*, 2003.

[18] Shudong Jin and Azer Bestavros. An empirical study of inherent routing bias in variable-degree networks. Technical report, 2003.

[19] G. Siganos and M. Faloutsos. Analyzing bgp policies: Methodology and tool. In *IEEE Infocom*, 2004.

[20] L. Gao. On inferring autonomous system relationships in the internet. In *IEEE Global Internet*, November 2000.

[21] J. Xia and L. Gao. On the evaluation of as relationship inferences. In *IEEE Globecom*, November 2004.

[22] X. Dimitropoulos, D. Krioukov, B. Huffaker, k. claffy, and G. Riley. Inferring as relationships: Dead end or lively beginning. In *WEA*, April 2005.

[23] Oregon routeview project, http://www.routeviews.org.

[24] Ripe route information service, http://www.ripe.net/ris.

[25] D. Achlioptas, A. Clauset, D. Kempe, and C. Moore. On the bias of traceroute sampling, or power-law degree distributions in regular graphs. In *STOC*, 2005.

[26] Internet routing registry, http://www.irr.net.

[27] http://www.cs.ucr.edu/bgp.

[28] Skitter, http://www.caida.org/tools/measurement/skitter/.

[29] Package cleaning house, http://www.pch.net.

[30] Eran Shir. Personal communication via emails, December 2005.

[31] Benoit Donnet, Philippe Raoult, Timur Friedman, and Mark Crovella. Efficient algorithms for large-scale topology discovery. In *Proceedings of ACM SIGMETRICS*, June 2005.

[32] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson. Measuring isp topologies with rocketfuel. *IEEE/ACM Trans. Netw.*, 12(1):2–16, 2004.

[33] S. Jaiswal, A. Rosenberg, and D. Towsley. Comparing the structure of power law graphs and the internet as graph. In *ICNP*, 2004.

[34] Tian Bu and Don Towsley. On distinguish between internet power law topology generators. In *IEEE Infocom*, 2002.

[35] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topology generators: Degree based vs. structural. In *ACM Sigcomm*, 2002.

[36] A. Medina, A. Lakhina, I. Matta, and J. Byers. Brite:an approach to universal topology generation. *MASCOTS*, 2001.

[37] E. W. Zegura, K. L. Calvert, and M. J. Donahoo. A quantitative comparison of graph-based models for internetworks. *Transactions on Networking*, 5(6):770–783, December 1997.

[38] G. Siganos. Nemecis: A tool to analyze the IRR registries. *Nanog 30, Miami*, 2004.

[39] G. Siganos. Nemecis: A tool to analyze the Internet Routing R egistries. *RIPE 48, Amsterdam*, 2004.

[40] http://www.traceroute.org.

[41] http://www.bgp4.net.

[42] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM*, pages 251–262, 1999.

[43] A. Medina, I. Matta, and J. Byers. On the origin of powerlaws in Internet topologies. *CCR*, 30(2):18–34, April 2000.

[44] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. The origin of power laws in internet topologies revisited. In *Infocom*, 2002.
[45] L. Gao and F. Wang. The extent of as path inflation by routing policies. In *IEEE Global Internet*, 2000.
[46] N. Spring, R. Mahajan, and T. Anderson. Quantifying the causes of path inflation. In *ACM Sigcomm*, 2003.
[47] M. Caesar and J. Rexford. Bgp policies in isp networks. *IEEE Network Magazine*, Nov/Dec 2005.
[48] L. Amimi, A. Shaikh, and H. Schulzrinne. Issues with inferring internet topological attributes. In *SPIE*, July 2002.
[49] Y. Hyun, A. Broido, and K. Claffy. Traceroute and bgp as path incongruities. In *CAIDA*, 2003.
[50] European internet exchange association, http://www.euro-ix.net.