A Review on Border Gateway Protocol and Internet Routing Registry

[1]Rakesh Phanindra. Akula,
[1]Department of Computer Science & Engineering,
[1]Jyothishmathi Institute of Technology and Science, JNTUH, Karimnagar, AP, INDIA.

[2]Vaishali P Khobragade,
[2]Department of Computer Science & Engineering,
[2]Jyothishmathi Institute of Technology and Science, JNTUH, Karimnagar, AP, INDIA.

[3]G.Swamy
[3]Department of Computer Science & Information Technology,
[3]Jyothishmathi Institute of Technology & Sciences, JNTUH, Karimnagar, AP, INDIA.

Abstract

The strength of the Internet relies heavily on the strength of BGP routing. BGP is the glue that holds the Internet together: it is the common language of the routers (hat interconnects networks or Autonomous Systems (AS). The robustness of BGP and our ability to manage it effectively is hampered by the limited global knowledge and lack of coordination between Autonomous Systems. One of the few efforts to develop a globally analyzable and secure Internet is the creation of the Internet Routing Registries (IRRs). IRRs provide a voluntary detailed repository of BGP policy information. The IRR effort has not reached its full potential because of two reasons: a) extracting useful information is far from trivial, and h) its accuracy of the data is uncertain. In this paper, we develop a methodology and a tool (Nemecis) to extract and infer information from IRR and validate it again BGP muting tables. In addition. Using our tool, we quantify the accuracy of the information of IRR. We find that IRR has a lot of inaccuracies, hut also contains significant and unique information. Finally, we show that our tool can identify and extract the correct information from IRR discarding erroneous data. In conclusion, our meth0dolgy and tool close the gap in the IRR vision for an analyzable Internet repository at the BGP level.

Keywords: BGP Table, IRR, IXP (Internet Exchange point) **,** IRR(Internet Rate of Return).

I. INTRODUCTION

The overarching goal of this work is to model and improve the robustness of the Internet at the BGP level. The Border Gateway Protocol is the protocol that dictates routing between Autonomous Systems (AS). And implements their business policies. The importance of BGP has become clear in the network community over the last five years, and several efforts have improved our understanding of BGP [1] [2] [3]. However. We still have a long way to go Studies show that BGP operates in a far from robust state and many of its behaviors are not well understood. The need for a robust Internet has created efforts like the Internet Routing Registries (IRR), a distributed database, where ASes store their policies. However, IRR has not reached its potential nor fulfilled the initial vision [SI. Our work attempts to take the IRR to the next level. We provide a systematic approach and a tool. Nemecis, to extract and infer useful information from IRR, with the ultimate goal to use this information to model, manage and protect Internet routing. There exist a number of tools to measure actual BGP routing, like ping, trace route, looking glass. BGP table dumps. But there does not exist a tool to bridge the gap between intended policy (configuration) and actual routing. Internet Routing Registries (IRR) [4], contain the policy of a large number of networks, expressed in a high level language, RPSL [5] [6] . These registries are considered by a lot of

people to be useless and outdated, based primarily on empirical evidence. To the best of knowledge, there does not exist a tool that can analyze these policies, and check their validity or freshness. The registries are maintained manually and in a voluntary basis to a large extent, and the policies remain as simple text. Thus, analyzing IRR is not a trivial task. The difficulties lie in: a) RPSL is very flexible, so policies can be very complex. b) There can be many different ways to express the same policy. c) The registries can contain inaccurate and incomplete data. At the same time the information of IRR is important in order to understand and interpret Internet Routing, since Routing tables are not sufficient to understand the intended policies.

In this paper, we develop a methodology and a tool for addressing the issues we just described. We call our tool Nemecis, which stands for Network Management and Con-figuration System. Our goal is to provide a framework for the analysis of RPSL policies, which can be used during the configuration phase, or the operation phase. During the configuration phase we can check the registered policy for correctness. During the operation phase, we can check whether the intended policy matches the actual routing. This way, we can reduce the time it takes to discover and fix routing problems. Most importantly, we can start to monitor how Internet routing works. In fact, our tool is among the first public tools to analyze the IRR policies. RIPE, has as a long term goal to validate the policies that Autonomous Systems register, and thus increase the robustness of BGP. Our work here is the first step in reaching this ambitious goal.

BGP routing tables: We consider the AS edges derived from multiple BGP routing table dumps [5], and compare them to the Route view data (OBD). The question we try to answer is what the information that the new BGP tables bring is. We use the term BD to refer to the union data from all available BGP table Dumps. Table I lists the acronyms for our data sets.

IRR data: We systematically analyze the IRR data and identify topological information that seems trustworthy by Nemecis [7]. We follow a conservative approach, given that IRR may contain some outdated and/or erroneous information. We do not accept new edges from IRR, even after our first processing, unless they are confirmed by trace routes (using our RETRO tool). Over all, we find that IRR is a good source of hints for missing links. For example, we discover that more than 80% of the new edges found in the new tables (i.e., the AS edges in BD but not in OBD) already exist in IRR [14]. Even compared to BD, IRR has significantly more edges, which are validated by RETRO as we explain below.

II. A 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 [17], the RIPE/RIS route collectors [7] and public route servers. Several other route collectors were not Operational at the time that the data was collected and there- fore, we do not include the min this study. For each BGP routing table dump, we extract its “AS PATH” field and generate an AS topology graph. We then merge 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. The statistics of BD are similar to what was reported in [5]. Interestingly, BD has only 0.5% additional ASes, but 20.4% more AS edges as compared with OBD.

For comparison purposes, we pick the most widely used AS Graph OBD as our baseline graph. For each of the other BGP outing tables, we examine the number of additional AS edges hat do not appear in OBD, as classified by their business relationship. As shown in Table 1, from each of the BGP routing able that provides a significant number of new edges to OBD, Most of the newfound 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: (1) most edges in a BGP table are provider-customer; and (2) 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 42 643 AS edges and merge new AS edges derived from the BGP table dumps other than OBD, one table dump at a time, sorted by the number of new edges they provide.

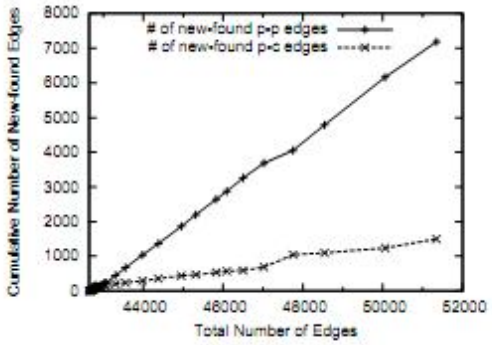

Fig1: Most new edges in BD but not in OBD are peer-to-peer edges.

At the end, when all the BGP table dumps in our data set are included, we obtain the graph BD; this has 51 345 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 curves (corresponding to the new found p-p edges and the p-c 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 33 841 AS edges, 2024 (5.98%) of which are peer-to-peer edges. On the other hand, the more complete AS graph OBD has 42 643 edges, and 5551 (13.0%) of these edges are peer-to-peer edges. The union 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 pay more attention to discovering peer-to-peer links.

Route collector or Router server name	# of Nodes	# of Edges	# of edges with type inferred			edges not in OBD	edges not in OBD w/ type		
			total	p-p	p-c		total	p-p	p-c
route-views(*OBD*)	19843	42643	42570	5551	36766	0	0	0	0
route-views2	19837	41274	41230	4464	36514	1029	1028	835	191
route-views.eqix	19650	34889	34876	1027	33640	674	674	530	143
route-views.linx	19655	37259	37246	3246	33765	2511	2511	2188	319
route-views.isc	19753	36152	36139	1915	34004	784	783	663	118
rrc00.ripe	19770	36479	36465	1641	34605	655	654	543	111
rrc01.ripe	19640	34193	34180	1121	32855	617	617	512	105
rrc03.ripe	19737	39147	39129	3850	35042	3233	3228	2609	616
rrc05.ripe	19765	32676	32659	1122	31324	1095	1091	658	432
rrc07.ripe	19618	32811	31797	1219	30394	804	803	724	79
rrc12.ripe	19628	33841	33827	2024	31606	1611	1610	1417	193
Total(*BD*)	19950	51345	51259	12734	38265	8702	8689	7183	1499

Table 1:A Collection Of BGP Dump Table

II. DISCOVERING 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 [7] as per the exporting policies of registered ISPs. The purpose of using Nemecis to filter the IRR is that, Nemecis 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 relation- ships: 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 custom 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, peer IRRnc and sibling IRRnc. From these edges, we define the set IRR dual to include the edges for which both adjacent ASes register matching relationships. (Contrarily, IRRnc includes edges for which only one AS registers a peering relation- ship while the other AS does not register at all.) Similarly, the IRR dual set can be decomposed by type of edge into three sets: pc IRR dual, peer IRR dual and sibling IRR dual.

III. IXPs AND DISAPPEARED LINKS

Note that, when two ASes are participants 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 [13].We defer a discussion of our method and tool on how to find the IXP participants to Section 5. However, even when we know the IXP participants, identifying the edges is still a challenge: not all participants connect with each other. In addition, the peering agreements among the IXP participants are not.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 IXP all the set of all edges that make up all of these cliques. Publicly known. IXP all contains 141,865 distinct AS edges.

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

Table 2: Many missing peer-to-peer links are at IXPs

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 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 [12]. We call this set of AS edges peer BD-OBD. Here we use the minus sign to denote the difference between two sets: A-B is the set of entities in set A but not in set B. Second, we look at the AS edges that appear in peer IRRnc but not in the graph BD. We call this set of links peer IRR inc BD. These AS links are the ones that are potentially missing from BD. We define the peer IRR dual links not in BD as peer IRR dual-BD.Having made this classification, we compare each class with the super set, IXP all, of edges that we constructed earlier. With our first comparison, we find that approximately 86% of the edges in peer BDOBD are in IXP all and hence, are potentially IXP edges. Next, we observe that 60% of the edges in peer IRR inc BD and 83% of the edges in peer IRR dual BD are in IXP all. Thus, if they exist, they could be IXP edges.

IV. OUTLINES OF THE PEER TO PEER EDGES

We study the properties exhibited by nodes that peer. Therefore, we examine the degrees, d1 and d2, of the two peering nodes that make up each peer-to-peer edge. Let us clarify that the degrees d1 and d2 include both peer-to-peer and provider-customer edges. One would expect that d1 and d2 would be “comparable”. Intuitively, one would expect that 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.

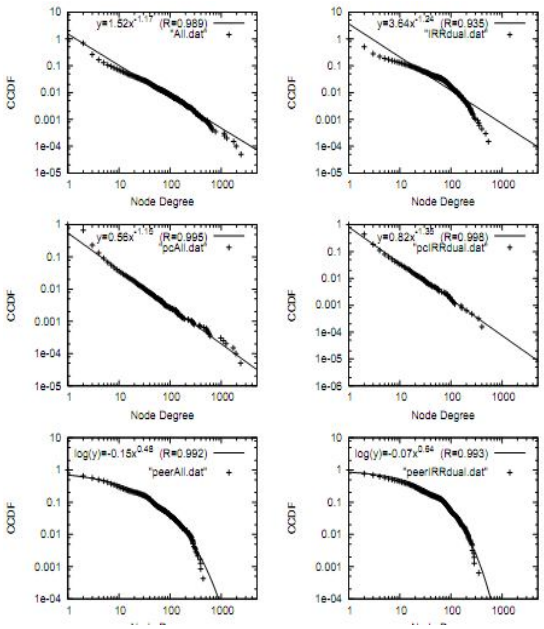

Figure 2: The degree distributions of ALL (left) and IR-Rdual (right) in the top row, their provider-customer degree distributions in the middle row, and their peer-to-peer degree distributions in the bottom row.

V. VALIDAING 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. However, as we have noted before, the peer-to-peer edges learned through the IRRs and IXP all are not guaranteed to exist. Therefore, in this section we focus on validating their existence to the extent possible. Note here that with the validation, we eliminate stale information that may still be present in the IRR and IXP data sources. The degree distributions of ALL (left) and IRR dual (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. Edges connect ASes whose degrees differ by a factor of 2. We plot the CDF of the distribution of the ratio min (d1, d2)/max (d1, d2) 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 might argue that this is an artifact of having peer-to-peer edges between low degree nodes, say d1 = 2 and d2 = 11, whose absolute degree difference is arguably small. This is why we examine the absolute difference of the degrees next. (2) 35% of the peer-to-peer edges have nodes with an absolute difference greater than 215. We plot the CDF of the distribution of the absolute value |d1 −d2|, where d1 and d2 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. We intend to investigate why quite a few high degree ASes establish peer relationship with low degree ASes in the future.

VI. IMPACT ON THE INTERNET TOPOLOGY

There has been a long debate on whether the degree distribution of the Internet at the AS level follows a power law [9] [10] [11] [3]. This debate is partly due to the absence of a definitive statistical test. 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 [11]. 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 IRR dual, which is a subset of the AS edges recorded in IRR filtered by Nemecis, the correlation coefficient is only 93.5%, see Fig. top right. 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 [8]. We show an indicative set of degree distribution plots for graph ALL on the left column and IRR dual on the right[14]. 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 middle row. Note that, although the combined degree distribution of IRR dual does not follow a power law (top row right), its provider-customer sub graph 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 3 in the bottom row. It is natural to ask why 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 [14]. 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. This is less obviously rue for peer-to-peer edges. If AS1 becomes a peer of AS2, AS1 does not benefit from the other peer to-peer edges of AS2: a peer will not transit traffic for a peer.

CONCLUSION

Drawing conclusions about BGP Table and IRR our work develops a methodical framework for the cross-validation and the synthesis of most available sources of topological information. We are able to find and confirm approximately 300% additional edges. additionally, we recognize that Internet Exchange Points (IXPs) hide major topology information and most of those new discovered peer-to-peer AS links are incident at IXPs. The reason for such a fact is probably because, most missing peer-to-peer links are likely to be at the middle or lower level of the Internet hierarchy, and peering at some IXP is a cost-efficient way for the ASes to setup peering associations with other ASes. We show that by adding these new AS links, some research results based on previous partial topology, such as routing decision and ISP profit/cost, change dramatically. Our study suggests that business-oriented studies of the Internet should make a point of taking into consideration as many peer-to-peer edges as possible.

REFERENCES

[1] T. Griffin. F. Shepherd. and G. Wilfong, "The stable path problem and Interdomain routing:" IEE/ACM Transactions on Networking. 2002.
[2] T. Griffin and G. Wilfong. "On the correctness of ibgp configuration," ACM Sigcomm, 2002.
[3] O. Maennel and A. Feldmann. "Realistic BGP traffic for test labs," ACM Sigcomm, 2002.
[4] "Internet Routing Registries." http://www.irr.nrr/.
[5] Alaettinoglu. C. Villamizar. E. Gerich. D.Kessens. D. Meyer, T.Bates, D.Karrenberg, and M.Terpstra. "Routing Policy Specification Language (RPS)", RFC2622.
[6] D.Meyer, J.Schmitz. C.Orange, M.Prior, and C.Alaettinglu, "using RPSL. In practice," RFC 2560.
[7] G. Siganos and M. Faloutsos. Analyzing BGP Policies: Methodology and Tool. In IEEE Infocom, 2004.
[8] H. Chang, S. Jamin, and W. Willinger. To Peer or not to Peer: Modeling the Evolution of the Internet's AS Topol-ogy. In IEEE Infocom, 2006.
[9] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On Power-law Relationships of the Internet Topology. In SIG-COMM, 1999.
[10] A. Medina, I. Matta, and J. Byers. On the origin of pow-erlaws in Internet topologies. CCR, 30(2):18–34, April 2000.
[11] 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.
[12] L. Gao and F. Wang. The extent of AS path inflation by
routing policies. In IEEE Global Internet, 2000.
[13] N. Spring, R. Mahajan, and T. Anderson. Quantifying the
causes of path inflation. In ACM Sigcomm, 2003.
[14] A. Ganesh, L. Massoulie, and D. Towsley. The Effect of Network Topology on the Spread of Epidemics. In IEEE
infocom, 2005.