Analysis of the Inter-Domain Routing Implied by the Internet Routing Registries and Building of AS-Topologies that Captures Route Diversity
Start date of the project: January 2004
Duration: 48 months
Project Coordinator: Prof. Dr. math. Friedhelm Meyer auf der Heide
Heinz Nixdorf Institute, University of Paderborn, Germany

Due date of deliverable: December 2006
Actual submission date: December 2006
Dissemination level: PU – public

Work Package 1.1: Efficient discovery of network topology and relationships
Participants: Università degli Studi di Roma “La Sapienza” (UDRLS)
Technische Universität München (TUM)

Authors of deliverable: Wolfgang Mühlbauer (wolfgang@net.in.tum.de)
Maurizio Pizzonia (pizzonia@dia.uniroma3.it)
1 Introduction

Knowing the topological structure of the Internet is essential for a wide range of networking tasks, e.g., making decisions about peering relationships, choice of upstream providers, inter-domain traffic engineering, etc. However, obtaining a good model of the Internet is not an easy task as shown by the number of works published in leading international conferences on this research subject.

We contribute to the development of methodologies and tools for building reliable Internet models in two ways.

- We provide a methodology, and an on-line service, to extract peering information from the Internet Routing Registry. Both the method and the service are based on: a consistency manager for integrating information across different registries, an RPSL analyzer that extracts peering specifications from RPSL objects, and a peering classifier that aims at understanding to what extent such peering specifications actually contribute to fully determine a peering. A peering graph is built with different levels of confidence. We compare the effectiveness of our method with the state of the art. The comparison puts in evidence the quality of the proposed method.

- We introduce a model of the Internet that allows more than one router/policy per AS, and arbitrary relationships between ASes. The idea is to build a topology and policy model that is consistent with the routing as observed by BGP observation points (e.g., RouteViews [routeview], RIPE [ripb]) and use this model to predict routes/AS paths. To build and evaluate our model, we exploit BGP observations from more than 1,300 observation points. We separate these into two datasets: a training set, and a validation dataset. The training set is used to build a topology and policies completely consistent with observed routing. We do so using a set of simulation-based iterative refinement heuristics. Comparing with the information in the validation set, then allows us to measure the predictive capabilities of our model. We find that in 93% of the cases (see Section 3.3) the observed paths are learned by the corresponding router in our model.

2 How to Extract BGP Peering Information from the Internet Routing Registry

The Internet Routing Registry (IRR) [ripc, irrb] is a large distributed repository of information, containing the routing policies of many of the networks that compose the Internet. The IRR was born about ten years ago with the main purpose to promote stability, consistency, and security of the global Internet routing. It consists of several registries that are maintained on a voluntary basis. The routing policies are expressed in the Routing Policy Specification Language (RPSL) [AVG+99, MSO+99, BDPR05]. The IRR can be used by operators to look up peering agreements, to study optimal policies, and to (possibly automatically) configure routers.

There is a wide discussion about the current role of the IRR [SF04]. Some people consider it outdated and almost useless. Others have put in evidence its importance to understand the Internet routing and that it contains unique and significant information. Anyway, it is undeniable that the IRR keeps on being fed by many operators, that useful tools have been developed to deal with the IRR (see, e.g., IRRToolSet [irra]), and that several research issues on the Internet routing are, at least partially, based on the content of the IRR. However, as pointed out in [SF04], extracting information from the IRR is far from trivial: the policies written in RPSL can be quite complex, the level of accuracy of the descriptions largely varies, and, also because of its distributed nature, the IRR contains many inconsistencies [Ker02].
2.1 State of the art

The RIPE offers an IRR consistency check service (RRCC) [SGK+01, rrc] that aims at detecting unregistered peerings. It verifies whether a peering that can be inferred from operational routing data is also described, in some form, into the IRR. Currently the RIPE service extracts peerings from the IRR in a way that is much less accurate than ours. Actually, the need of a better analysis of the content of the IRR is pointed out by the RIPE itself that considers this as a long term goal [rrc].

On the research side, Mahadevan et al. [MKF+06] presented a comparison of several characteristics of the AS-level topologies built on the basis of different data sources, including the IRR. They also proposed a metric to characterize such topologies. Zhang et al. [ZLMZ05] derived an AS-level topology combining IRR data with BGP routing information collected from multiple sources, such as RouteViews [rou], looking glasses, and route servers. They showed that the data from the RIPE registry reveal topology information which cannot be found in other sources. Siganos et al. [SF04] developed a tool, called Nemecis [nem], that checks the correctness of IRR data and their consistency with respect to BGP routing table information. They argued that 28% of ASes have both correct and consistent policies and that RIPE is by far the most accurate registry. Carmignani et al. [CBD+02] presented a service for the visualization of IRR data. We shall compare the level of accuracy of the methods for extracting peerings from the IRR used in the above papers with respect to ours.

2.2 Results

We have defined a method and a on-line service to extract peering relationships from the IRR. Both the method and the service are based on: a consistency manager for integrating information across different registries, an RPSL analyzer that extracts peering specifications from RPSL objects, and a peering classifier that aims at understanding to what extent such peering specifications actually contribute to fully determine a peering. A peering graph is built with different levels of confidence. The figure 1 describes the architecture of the service.

Using the method (and the system) described, we analyzed the data of 68 registries, downloaded from [ripa, rad]. As a result, we extracted about double of peerings than the state of the art.

We have implemented our method as an on-line service, available on the web1. The service produces, on a daily basis:

1http://tocai.dia.uniroma3.it/~irr_analysis
• General statistics on the IRR (number of objects defined in each registry, amount of overlapping information between registries, etc.).

• A set of pairs of ASes, corresponding to peering relationships extracted from the IRR. Each pair is labeled with information about the context where it has been found, like the type of policy and the registry.

The service provides also on-demand plots about:

**Timestamp distribution**: distributions of the last update timestamps of RPSL objects. The figure 2 shows an example of timestamp distribution plot.

**Size variation**: evolution over time of the size of the IRR in terms of registered RPSL objects. The figure 3 shows an example of size variation plot.

**Peerings over time**: evolution over time of the number of peerings extracted from the IRR. The figure 4 shows an example of peerings over time plot.

As a side effect, our study highlights how the different RPSL constructions are actually used to specify peerings.

Further details can be found in the DELIS-TR-0388 [BRR06].

### 3 Building an AS-Topology Model that Captures Route Diversity

In this section we provide a brief overview of a model of the Internet that allows more than one router/policy per AS, and arbitrary relationships between ASes. The full work has been published in [MFM+06].
Figure 3: Size variation

Figure 4: Peerings over time
3.1 State of the Art and Weaknesses of Previous Approaches

In the past, large-scale models of routing in the Internet were frequently based on two assumptions. First, they assume that each AS consists of a single router, e.g., [MQWZ05]. However, analyzing our datasets – observed AS paths from RIPE, RouteViews [ripb, routeview] and other sources – reveals that for more than 30% of originating and observation AS pairs (over all prefixes advertised by the origin), we see more than one AS-path. Additionally, we compute AS-level paths on the AS-level graph obtained from our datasets using the BGP simulator C-BGP [Quo03] and compare those paths with the observed AS-paths. For slightly less than 50% of the prefix/observation point combinations the observing AS does not even learn the “correct” AS-path. All this shows that one router per AS is obviously not sufficient to capture the full diversity caused by intra-domain routing.

Second, previous models frequently relied on inferring customer-provider relationships utilizing the valley-free assumption[Gao00, SARK02, BPP03]. Unfortunately, such techniques seem to show bad results for inter-domain route prediction. We use a simple heuristic to infer economic AS relationships similar to [Gao00, SARK02, BPP03] and then set up a BGP simulation with C-BGP accordingly. The results are fairly discouraging with only 12.5% of the observed paths being correctly predicted in the simulation. Altogether, this indicates a low accuracy for AS-path prediction, if an AS-routing model is solely based on AS-level inference.

3.2 Approach

The goal is to capture the outcome of routing policies and the internal structure of ASes from observed BGP data to an extent, which allows to predict Internet path choices for previously unobserved AS-paths with high accuracy. We choose an approach that introduces multiple routers within an AS, so-called quasi-routers, and routing policies that are agnostic about inferred relationships such as customer-provider. Quasi-routers do not represent the physical router topology of a network. Each quasi-router can be thought of as a group of routers within an AS all making the same choice about best route.

We rely on BGP data from more than 1,300 BGP observation points including those provided by RIPE [ripb], Routeviews [routeview], Geant [ID] and Abilene [Abi]. We separate these into two datasets: a training dataset used to build the model and a validation dataset to evaluate the predictive capabilities of our model afterwards.

Our approach uses a simulation-based iterative refinement heuristic to build an AS-topology model that is consistent with observed routing. We start from the simplest AS-model possible. It consists of one quasi-router per AS and contains one edge between any two connected ASes of the AS-level graph obtained. Then, we execute a C-BGP [Quo03] simulation for each prefix and compare the predicted AS-paths with the actual observed AS-paths of the training set. Mismatches can be due to two reasons: First, the model prefers the shortest AS-path in the absence of policies. Second, the quasi-routers inside an AS do not suffice to capture the required route diversity. Therefore, we alter the model by adding routing policies and/or quasi-routers accordingly. For more details, the reader is referred to [MFM*06]. We do not correct the discrepancies between an observed path and the path predicted by our model in one step. Rather, we move from the origin of the observed route towards the observation points. If a mismatch is found, we ensure that the desired route is propagated one AS further towards the observation point in each iteration. After applying the changes, we always restart simulations.

Note, that only paths of the training set are considered during the iterative refinement of our model. For a fair evaluation of the predictive capabilities of the refined model, we use a separate dataset, called validation set. To measure how accurately the model predicts those “previously” unobserved paths, metrics are introduced. After the simulation one has access to the routing information base (RIB) for a certain prefix at any quasi-router. Therefore we can compare for each
AS the AS-path that is recorded in the BGP data to those AS-paths chosen in the simulation. The degree of mismatch is measured by the following two metrics:

- **RIB-In match**: The observed route at an observation point is contained in the corresponding simulated RIB-In for at least one quasi-router in the observed AS. This means that the observed route has at least been learned by a router in the simulation, but not necessarily selected as best route.

- **RIB-Out match**: At least one quasi-router in the AS has selected the route with the observed AS-path as its best route and propagates it to its neighbors.

### 3.3 Results

Now we derive an AS-routing model for a set of training data and evaluate the effectiveness of our refinement heuristic using separate validation data. First, an AS graph is constructed from all AS-path information available in our datasets. Of the 1,300 BGP observation points of our dataset, we randomly assign 2/3 to the training set and the remainder ones to the validation set. We sub-select 1,000 originating ASes and the corresponding AS-paths from both the training and the validation set. Our refinement heuristic (see Section 3.2) now incrementally refines the model with the goal of achieving an exact match between the AS-paths predicted by the model and the training set.

Figure 5(a) shows the progress of the heuristic with each iteration as measured in terms of RIB-In matches and RIB-Out matches. The length of the longest AS-path is 10, and 11 iterations happen to suffice our goal of perfect RIB-Out matches. Notice that the early progress of the heuristic is excellent. Just one iteration more than doubles the percentage of RIB-Out matches from 24.5% to 59.3%, and increases the potential RIB-Out matches and RIB-In matches to more than 70% and 85% respectively. Given that the average length of the AS-paths is about 4.3, it is not surprising that we achieve RIB-Out matches for all but 5% of the AS-paths after five iterations.

Given an AS-routing model we can now evaluate its predictive capabilities for our example question for a different set of observation points. We find, based on the subset of validation paths, that we improve our prediction capabilities from 25.5% (without routing policies) to 63% for RIB-Out matches (see Figure 5(b)). Let us point out that the major improvements happen during the first six iterations. To judge the qualitative improvement of our results vs. those reported by Mao et al. [MQWZ05] we point out that our results hold across more than 300 observation points rather than 3 ASes and are significantly better. In terms of RIB-In matches which correspond to matches we have 93% vs. their 82%, 64%, and 16%.
The main reason for the effectiveness of our refinement heuristic is that an AS can, if necessary, consist of multiple routers. Figure 6 shows a histogram of the number of quasi-routers per AS in the refined model (for those ASes with more than one quasi-router). For almost all ASes (14,305) one quasi-router suffices. For 71 we need two. Yet, there are 138 ASes that need more than 9 quasi-routers. Not surprisingly we find that among the ASes with many quasi-router are predominantly well-known tier-1 ASes. After all, tier-1 ASes offer peerings at quite a number of peering locations, peer with and provide service to many other ASes, and have a sizable backbone network.

A more detailed discussion of the results can be found in the DELIS-TR-0401 [MFM+06].

References


[irra] Internet Routing Registry Toolset (IRRToolSet). 


