Modeling Internet Topology Dynamics

Hamed Haddadi[*]
University College London

Steve Uhlig
Delft University of Technology

Andrew Moore
University of Cambridge

Richard Mortier
Vipadia Ltd

Miguel Rio
University College London

ABSTRACT

Despite the large number of papers on network topology modeling and inference, there still exists ambiguity about the real nature of the Internet AS and router level topology. While recent findings have illustrated the inaccuracies in maps inferred from BGP peering and traceroute measurements, existing topology models still produce static topologies, using simplistic assumptions about power law observations and preferential attachment.

Today, topology generators are tightly bound to the observed data used to validate them. Given that the actual properties of the Internet topology are not known, topology generators should strive to reproduce the variability that characterizes the evolution of the Internet topology over time. Future topology generators should be able to express the variations in local connectivity that makes today's Internet: peering relationships, internal AS topology and routing policies each changing over time due to failures, maintenance, upgrades and business strategies of the network. Topology generators should capture those dimensions, by allowing a certain level of randomness in the outcome, rather than enforcing structural assumptions as the truths about Internet's evolving structure, which may never be discovered.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Network topology; I.6.4 [**Simulation and Modeling**]: Model Validation and Analysis

General Terms

Topology, Models, Measurement

Keywords

Internet, Topology generation

1. INTRODUCTION

The spatial analysis of the Internet is necessary for research on network planning, optimal routing techniques and failure detection measures. An important step towards in-depth analysis of the Internet performance is the availability of network topologies at different levels of granularity. In the topology generation literature, current research focuses on distribution-driven methods, which rule out the randomly generated graphs and aim at attaching meta-data information (metrics) to the links and routers generated in a graph. Information about directionality of links, the delay and bandwidth, would also be of significant value to researchers using those graphs.

Generation of the topology of the Internet calls for a model that achieves a good balance between keeping global (structural) characteristics and more local properties like node degrees and local interconnection structure. Figure 1 displays the end result of a small portion of the Internet that might be produced with an end-to-end topology generator.

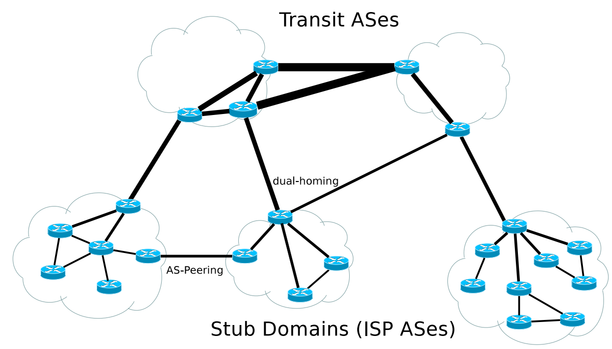

Figure 1: An abstract part of the Internet, link widths represent relative bandwidth.

An area which is of importance for simulation work is focus on tools that produce a representative view of the Internet topology, from the Autonomous System (AS) level and their relationships, down to the Point of Presence (PoP) level and eventually the links and router level topology. An example is enterprise networks spanning the globe, with nodes and routers being part of multiple ASes. Such a tool would provide researchers with a testbed for looking at problems such as wide-area failure detection, propagation of network epidemics, fault isolation or resilience routing. The main goal of our work is to build a representative, end-to-end network topology generator, with the ability to incorporate multiple scales and dynamic evolution of the networks.

Current topology generators are highly bound to the observed data used to validate them. In closely related works, the importance of temporal is shown by using traffic matrices, which represent the traffic flow over a given topology. Soule *et al.* [16] demonstrate that traffic matrices represent

[*]This work was done while the author was visiting the Computer Laboratory, University of Cambridge.

the spatial and temporal characteristics of the network and their models are validated against measured traffic matrices. We believe that for thorough analysis of the dynamics and evolution of networks, not only traffic matrices, but all aspects of topologies must be analyzed over a period of time. Topology generators are not concerned with the volume of data across links. Rather, they focus on the structure of networks. They produce static graphs of the Internet which then need to be validated against a static map of the Internet. Given that the actual properties of the Internet topology are not known, topology generators should strive at reproducing the variability that characterizes the evolution of the Internet topology over time. Future topology generators should be able to express the variations in local connectivity that makes today's Internet: peering relationships, internal AS topology, number of ASes and links are constantly evolving due to failures, maintenance, and upgrade of the ISPs. Realistic topology generators should capture those dimensions.

2. INTERNET TOPOLOGY MODELS

In the topology generation research, there is need for a generator that provides a complete network from transit AS level to router level topology. Obtaining meta-data from observed topologies is difficult. Some attempts have been made to label AS-level links with inferred peering relationships, part of them being validated by ISPs [4]. Other works have tried to infer the delay and bandwidth of ISP topologies [17]. However, data originated from different sampling processes are difficult to combine. For instance, combining the intradomain map of an ISP with BGP data might not give sufficient details about the router-level topology or the interconnections between ASes that are not observed by the measurements. If incorrect AS interconnections or policies are used for simulation purposes, then the resulting routes might be far from realistic [10, 11].

We are aware of a single attempt towards an end-to-end topology generator, IGen [15], that generates router-level topologies using network design heuristics and geographic constraints. In IGen, routers are generated in continents, are grouped into Points of Presence (POP) according to the geographic proximity of the routers, and the links between routers are created using network design heuristics. ISP topologies are indeed not instances of random graphs, but are the result of careful design [2, 5, 1] as well as geographic, technical [8] and business constraints. IGen aims to generate realistic ISP topologies, and allowing those topologies to be interconnected like real ASes might be. However, IGen does not solve the problem of choosing among the diverse possible ISP topologies nor how they should be interconnected. Generating topologies of the size of the Internet remains challenging, as a proper understanding of the importance of different structural aspects of the Internet is missing today.

Recently, we carried out an extensive comparative study of different topology generators and their underlying mathematical models. From this comparison of available topology generators, we identified their similarities and differences in replicating a number of observed topologies. This comparison was done by using the topology generators which are available for router and AS level models of the Internet, e.g., Inet [20], PFP [21], IGen [15], GT-ITM [3]. Given a known input, such as certain number of nodes, the output from these generators and the underlying models can be directly compared, using topological metrics. Comparisons have shown that the underlying assumptions in these generators, such as preferential attachment and rich-club connectivity, are largely biased. For years, researchers have modeled the Internet AS topology using topologies collected from BGP and traceroute data. These data are shown to be inaccurate, missing many peering links and underestimating the effects of these peering links on the AS topology structure [13]. The shortcomings in current topology generators are partly stemmed at the very observations used to design the generators, as not all parts of the Internet are equally visible from given vantage points [12]. Such issues call for understanding the nature of the Internet AS and IP level topology, and accordingly, consider representative topologies for simulations.

3. DYNAMIC NETWORK MODELS

The Internet is not a static network: links and routers are continuously added and removed. Routing failures happen, new prefixes are advertised and withdrawn. Both the physical graph of the network, and the visibility we have of it through routing, change. There has been little work done in topology research in order to incorporate discrete events such as addition of nodes and links or link failures into a topology generator. Such research is becoming increasingly important for enterprise network operators who try to achieve security and resilience by segmentation of networks into various operational domains using VLANs, private AS numbers, global routers and firewalls. As an important stage towards identifying dynamics of topologies, building up on previous efforts in the community in analysis of link failures in IP networks [6] and the evolution of AS level topologies [14], we are developing a methodology for embedding several aspects of the evolution of the Internet in our generated topologies.

3.1 AS topology evolution

At the AS level, there is a constant growth in the number of peering links between ISPs [13]. Also, due to policy routing and hot potato routing, the changes at the IP level affect the AS level [19]. Simulation for applications such as routing protocols and analysis such as studies in prefix hijacking would benefit from topologies which take into account the changes of the network over time, similar to real network behavior. We believe that using static topologies does not fully exploit the potential scenarios that one should consider in simulations. Another important aspect of the networks that is not captured by current models is the move of the Internet AS topology towards having a meshed core of tier-1 ISPs, alongside multiple peering relationships between edge ASes, and an atypical connection models of some ASes such as the content providers which form many peering connections with as many ASes as possible in order to avoid high transit charges [12]. In addition to information about peering links, the availability of models of growth and evolution of networks will enable us to include dynamic models for generating synthetic AS topologies.

3.2 IP layer dynamics

The control plane at the router level has different characteristics to those seen at the AS level. At the router level, the dynamics are more frequent and tend to have a shorter durations. Regular maintenance works, router and link failures, traffic engineering, firewall misbehaviors and other factors all effect the routing at the IP layer. Operators do not disclose information about routing changes and link failures. This has made it difficult to model the behavior of router level Internet topology. We are building up on previous work [6], by inferring the characteristics of router level topology of major ISPs, looking at short term and long terms trends, while considering the effects of the ISP network. This will also allow us to build a model for dynamic topology generation at the router level.

4. VALIDATION

Validating generated topologies leads to several challenges. First, topology generators aim at generating classes of graphs, whose properties are similar to those of the Internet. Defining such classes depend on both how and what topological properties should be reproduced [9]. Deciding about the boundary between realism, i.e. how closely the generated graphs match reality, and genericity, i.e. how well the topology generator samples the chosen class of graphs, has not been studied to our knowledge. A single study [9] did investigate how well topological properties of the AS-level Internet are approached, when degree correlations are increasingly constrained to match those of observed AS topologies.

Second, current topology generators rely on strong structural assumptions about AS-level interconnections, which are questionable. For instance, [18] has shown that enforcing a hierarchy is not necessary to properly reproduce the global properties of the Internet. Rather, by letting the node degree distribution create a loose hierarchy produces better results. This raises the question of which properties, if any, should be enforced by a topology generator to create a realistic topology. Too much realism is not desirable, as it limits the graph space that the generated topologies span.

Third, the topological properties of the Internet have, in our opinion, not been sufficiently studied. The private nature of ISPs topologies, as well as the importance of business relationships for ASes, has impeded on a more thorough knowledge of the Internet structure. The sampling made by tools like traceroute or by BGP data is not properly understood, as it requires reverse-engineering the often complex path choice made by the routing protocols. Understanding the visibility of routing paths is key to better evaluation of the impact of missing links from the Internet measurements [12].

Finally, the properties that may be considered as relevant for correctly reproducing the Internet graph depend much on the existing topological metrics. Topological metrics have been introduced to measure peculiar aspects of graphs, which might not be relevant for the Internet. Studying a large set of topological metrics on different types of real-world networks has shown that many metrics are largely redundant [7]. Deciding about a set of topological metrics is critical to judge the realism of a generated topology. It is, however, a subjective decision on set of properties that should be reproduced, and how well the topological metrics of generated topologies should match those of reference (observed) topologies. These challenges might depend on the usage of the generated topologies.

5. CONCLUSIONS AND FUTURE WORK

Accuracy of simulations on Internet protocols and applications is strongly dependent on the correct use of synthetic topologies at large scales. The Internet should not simply be seen as a power law network of transit and stub ASes. The core is a dense mesh of tier-1 ISPs interconnected by peering relationships. At the edge on the other hand, costumer ISPs aim at increasing their peering relationships with other established ASes, through Internet eXchange Points (IXP). Content distribution networks are also increasingly adding peering links with as many as they can afford. The availability of rich topology information from numerous vantage points (at least at the AS-level) calls for an end to blind use of random networks or power law models for simulations. In addition, findings on growth, birth and death rates of ASes, alongside failure models of routers and links, should allow to build a concrete topology generator which models enough of the realism of the network to simulate the end-to-end behavior of the Internet. Pursuing this goal, we aim to form a collaboration with network operators, alongside topology generator designers, to provide a representative dynamic topology generator to the research community.

Acknowledgment

This work is conducted as part of the EPSRC UKLIGHT MASTS project (Grants GR/T10503/01 and GR/T10510/03).

6. REFERENCES

[1] D. Alderson, J. Doyle, R. Govindan, and W. Willinger. Toward an optimization-driven framework for designing and generating realistic Internet topologies. *SIGCOMM Computer Communications Review*, 33(1):41–46, 2003.

[2] D. Alderson, L. Li, W. Willinger, and J. C. Doyle. Understanding Internet topology: principles, models, and validation. *IEEE/ACM Transactions on Networking (TON)*, 13(6):1205–1218, 2005.

[3] K. L. Calvert, M. B. Doar, and E. W. Zegura. Modeling Internet topology. *IEEE Communications Magazine*, 35(6):160–163, 1997.

[4] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, kc claffy, and G. Riley. AS relationships: inference and validation. *SIGCOMM Computer Communications Review*, 37(1):29–40, 2007.

[5] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou. Heuristically optimized trade-offs: A new paradigm for power laws in the Internet. In *ICALP '02: Proceedings of the 29th International Colloquium on Automata, Languages and Programming*, pages 110–122, London, UK, 2002. Springer-Verlag.

[6] G. Iannaccone, C. nee Chuah, R. Mortier, S. Bhattacharyya, and C. Diot. Analysis of link failures in an IP backbone. In *IMW '02: Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, Marseille, France, 2002.

[7] A. Jamakovic, S. Uhlig, and I. Theisler. On the relationships between topological metrics in real-world networks. In *European Conference on Complex Systems*, Dresden, Germany, 2007.

[8] L. Li, D. Alderson, W. Willinger, and J. Doyle. A first-principles approach to understanding the

Internet's router-level topology. In *Proceedings of ACM SIGCOMM 2004*, pages 3–14, Portland, OR, 2004.

[9] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat. Systematic topology analysis and generation using degree correlations. In *Proceedings of ACM SIGCOMM 2006*, pages 135–146, Pisa, Italy, 2006.

[10] W. Muhlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig. Building an AS-topology model that captures route diversity. In *Proceedings of ACM SIGCOMM 2006*, pages 195–206, Pisa, Italy, Sept. 2006.

[11] W. Muhlbauer, S. Uhlig, B. Fu, M. Meulle, and O. Maennel. In search for an appropriate granularity to model routing policy. In *Proceedings of ACM SIGCOMM 2007*, Kyoto, Japan, Aug. 2007.

[12] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. In Search of the elusive Ground Truth: The Internet's AS-level Connectivity Structure. In *ACM SIGMETRICS*, Annapolis, USA, June 2008.

[13] R. Oliveira, B. Zhang, and L. Zhang. Observing the Evolution of Internet AS Topology. In *Proceedings of ACM SIGCOMM 2007*, Kyoto, Japan, Aug. 2007.

[14] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, 2004.

[15] B. Quoitin. Topology generation based on network design heuristics. In *CoNEXT'05: Proceedings of the 2005 ACM conference on Emerging network experiment and technology*, pages 278–279, Toulouse, France, 2005.

[16] A. Soule, A. Lakhina, N. Taft, K. Papagiannaki, K. Salamatian, A. Nucci, M. Crovella, and C. Diot. Traffic matrices: balancing measurements, inference and modeling. In *Sigmetrics*, pages 362–373, New York, NY, USA, 2005. ACM Press.

[17] N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP topologies with rocketfuel. In *Proceedings of ACM SIGCOMM 2002*, pages 133–145, 2002.

[18] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topology generators: degree-based vs. structural. In *Proceedings of ACM SIGCOMM 2002*, pages 147–159, Pittsburgh, PA, 2002.

[19] R. Teixeira, A. Shaikh, T. Griffin, and J. Rexford. Dynamics of hot-potato routing in ip networks. *SIGMETRICS Perform. Eval. Rev.*, 32(1):307–319, 2004.

[20] J. Winick and S. Jamin. Inet-3.0: Internet topology generator. Technical Report CSE-TR-456-02, University of Michigan Technical Report CSE-TR-456-02, 2002.

[21] S. Zhou. Characterising and modelling the Internet topology, the rich-club phenomenon and the PFP model. *BT Technology Journal*, 24, 2006.