Available online at www.internationjournalofsocialscience.com
INTERNATIONAL JOURNAL OF SOCIAL SCIENCE AND HUMANITY RESEARCH
E ISSN: 2689 3525
Volume:07; Issue:01 (2024)

The deployment of the Internet Autonomous System Classification of the Topology of Internet Autonomous Systems

CAO YUANQING[1], DR. MIDHUNCHAKKARAVARTHY[2a]
[1]PhD. Research Scholar in Information Technology, Lincoln University College Malaysia
[2]Professor in Lincoln University College Malaysia
Contact Details: [a]midhun@lincoln.edu.my

Abstract

Although the Internet has been studied from its origin, academics did not pay considerable attention to its development and evolution until the mid-1990s, when it became economically feasible in the United States. To begin, Govindan and Reddy [16] first defined an AS level graph. The graph was shown as a collection of nodes and linkages, with each node representing an Internet domain. They discovered that, despite the fact that the Internet has grown significantly, the degree and route dispersion have remained consistent with pre-boom levels. Faloutsos et al., however, coined the term "Internet topology" in their landmark research [17]. If the data from the BGP monitor was trustworthy and complete enough, the researchers concluded that it allowed for the first comprehensive look at AS topology. They found that the distribution of AS degrees followed a simple power formula. Academic interest in Internet topology and data gathering methods, graph creation, and graph methodologies and analysis has grown since the publication of this landmark work [18].

Keyword: RIS, PCH, AS Topology, BGP, Internet topology

INTRODUCTION

Many autonomous systems are linked together over the Internet, including tens of thousands of assessments operated by various administrative bodies. BGP determines how ASes communicate with one another (BGP). Using BGP, each autonomous system (AS) can select the routes it imports and exports from its neighbors. AS relationships, a sort of business arrangement between ASes, drive these policies, which are set by network administrators. P2P and P2C partnerships are the two most popular types of AS connections (p2p). It is the customer's duty to pay for the service provider's role in moving communications from the Internet to other networks. Two ASes can easily share traffic with their clients, but not with their providers. As a logical result of this economic paradigm, an AS typically does not export its provider and peer routes to those providers or peers. Understanding the Internet's structure, inter-domain routing dynamics, and evolution necessitates an understanding of the business relationships between ASes. Because business relationships are a closely guarded secret within organizations, Internet researchers find it challenging to appropriately identify AS Relationships [1].

The collecting of BGP path measurements is a significant driving force behind the inference of AS links. As a result of these efforts, two initiatives emerged: Route Views by the University of

Available online at www.internationjournalofsocialscience.com
INTERNATIONAL JOURNAL OF SOCIAL SCIENCE AND HUMANITY RESEARCH
E ISSN: 2689 3525
Volume:07; Issue:01 (2024)

Oregon in the United States and Routing Information Service (RIS) by RIPE in Europe, both of which serve as major examples of how these efforts have been employed. Route collectors located across the world (VPs) maintain BGP peering sessions with other ASes, also known as Vantage Points. VP AS numbers occur first in the AS paths obtained by collectors who initially peer with the VP. Every day, Route Views and RIPE route collectors listen to these VPs' BGP routing table entries and archive them. They use them every day. Peering sessions between the route collector and the VP can be established either directly or through an Internet Exchange Point (IXP). An IXP is a centralized site where Assess can communicate and exchange information. Passive measures do not require new traffic to be added to the network [2, 3].

LITERATURE REVIEW

The BGP routing table entries are the most important when attempting to ascertain an ASN's topology. There are AS paths connecting the nearby VP's routing table to a prefix block of IP addresses. To keep things simple, we'll refer to this graph as AS Graph. These nodes and edges represent the connections between a group of autonomous systems (AS). Adding AS connections such as p2c and p2p to the edges of an Annotated AS graph results in a new type of AS graph [4].

BGP routes can be obtained from various sources. Packet Clearing House (PCH) [3], a key proponent of IXPs, monitors IXPs from all over the world. Links from an Internet Routing Registry (IRR) [4], a collection of routing policy databases intended to provide a comprehensive view of the Internet, are less common. These databases, also known as Routing Arbiter Databases (RADb), are operated by organizations such as RIPE [5] and Merit Network [6]. AS links can be extracted from a set of IP addresses. Traceroute and tracert are popular network troubleshooting commands. In response to the command, a list of IPv4 addresses associated with routers responding via the Internet Control Message Protocol (ICMP) is returned. Traceroute measurements. CAIDA's Archipelago (Ark) measurement equipment is used to generate the IPv4 Routed /24 AS Connections Dataset [7] by translating IP addresses to matching ASes and constructing AS linkages. [7]. Adding traceroute measures to the network expands the quantity of traffic that is monitored. To summarize, three methods exist for determining an AS link: BGP routes, traceroute, and IRR.

STATEMENT OF THE PROBLEM

More than half of previous research has looked at Internet topology in terms of node degree distribution, betweenness, and average hop count at the Autonomous System (AS) level [10]. In contrast, macroscopic metrics do not capture the local properties of AS connections. Most older models, on the other hand, are simple abstractions that do not account for the many types of nodes and linkages. As a result, these assessments fail to account for regional variances in the evolution of the Internet as influenced by economic, political, and business issues. What is the cause of this growth in size? How globally networked is the Internet? How will it change by region? Finally, the present techniques make no practical recommendations on how to Allocate resources. Even CDN managers may find it challenging to cache data in many locations, create new peering relationships, and anticipate performance/cost trade-offs. [11], as an example To assess the consequences of their actions, they must first understand the Internet's architecture and dynamics.

The study aims

To create efficient internet infrastructure to meet increasing traffic demands.

Research Questions:

• How has the Internet evolved over time and across locations?

RESEARCH DESIGN:

The UCLA data repository will be utilized to generate the basic AS graph, which will only comprise AS nodes and AS links. Because graph topologies change every day, a monthly breakdown of daily data will be required. The UCLA dataset has some flaws, such as erroneously promoted links due to routing table issues, path poisoning, or router failures. These ephemeral events are limited to a few hours. As a result, we eliminated AS connections that existed only once each month in order to eliminate any potential misleading paths.

To determine the AS relationship type, we used CAIDA data from the Topology data set. As part of the inference technique in [23], we didn't filter the CAIDA dataset because It removes incorrect information. Although the CAIDA dataset did not include information on the connection type for 10% of the linkages in the basic AS graph, we had to guess it.

DATA ANALYSIS

In addition to SVM (146), K-Nearest Neighbor (147), and Random Forest (148), we test numerous more models. Among these models, the Random Forest classifier is the most effective. The Receiver-Operator Curve (ROC) represents the tradeoff between sensitivity and specificity. As can be observed, the False Positive and True Positive rates remain below 10% and above 80%, respectively. This classification task yielded an area under the ROC curve of 0.98. The Precision-recall curve for this model will show how the two metrics were compromised. To assess the Random Forest classifier's precision and recall, we intend to divide the samples 70% to 30% for training and testing. We will repeat this technique exactly one thousand times.

CONCLUSION

We described a machine learning strategy for inferring edge types in AS graphs built from open-source data. The Gentle AdaBoost machine learning approach and the five node attributes collected from the AS graph were used to train a classifier for p2p and p2c edges. We use our method to categorize three AS graphs: a BGP network, a traceroute graph, and an IRR graph. Each classifier is evaluated using two datasets. The first test set is based on the BGP dataset, and the second on the CAIDA AS connection inference dataset. Combining the three independent AS graphs enables the computation of edge types in an AS graph. We analyze three. Different graphs and one composite graph were examined to find their distinguishing features. All three graphs contain a large number of distinct p2p and c2c edges. Each. Integrating the three graphs provides a far more complete picture of the Internet's peer-to-peer and peer-to-consumer ecosystems.

References

[1] D. Achlioptas, A. Clauset, D. Kempe, and C. Moore. On the bias of tracer- oute sampling. In STOC, ACM, volume 1581139608, page 0005. Citeseer, 2005.

[2] M. B. Akgun and M. H. Gunes. Bipartite internet topology at the subnet- level. In Network Science Workshop (NSW), 2013 IEEE 2nd, pages 94–97. IEEE, 2013.

[3] R. Albert and A.-L. Barab´asi. Topology of evolving networks: local events and universality. Physical review letters, 85(24):5234, 2000.

[4] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani. k-core de- composition of internet graphs: hierarchies, self-similarity and measurement biases. arXiv preprint cs/0511007, 2005.

[5] B. Augustin, B. Krishnamurthy, and W. Willinger. Ixps: mapped? In Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference, pages 336–349. ACM, 2009.

[6] A.-L. Barab´asi and R. Albert. Emergence of scaling in random networks. science, 286(5439):509–512, 1999.

[7] A. Baumann and B. Fabian. Who runs the internet?-classifying autonomous systems into industries. In WEBIST (1), pages 361–368, 2014.

[8] M. Baur, U. Brandes, M. Gaertler, and D. Wagner. Drawing the as graph in 2.5 dimensions. In Graph Drawing, pages 43–48. Springer, 2005.

[9] K. Boitmanis, U. Brandes, and C. Pich. Visualizing internet evolution on the autonomous systems level. In Graph Drawing, pages 365–376. Springer, 2008.

[10] J. Brodkin. Why YouTube buffers: The se- cret deals that make-and break-online video. http://arstechnica.com/information-technology/2013/07/why-youtube-buffers-the-secret-deals-that-make-and-break-online-video/, July 2013.

[11] T. Bu, N. Duffield, F. L. Presti, and D. Towsley. Network tomography on general topologies. In ACM SIGMETRICS Performance Evaluation Review, 2002.

[12] T. Bu and D. Towsley. On distinguishing between internet power law topol- ogy generators. In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, volume 2, pages 638–647. IEEE, 2002.

[13] Caida ark project. http://www.caida.org/projects/ark/.

[14] Caida datasets. http://data.caida.org/datasets/.

[15] Caida. Ipv4 and ipv6 as core: Visualizing ipv4 and ipv6 internet topology at a macroscopic scale in 2014. http://www.caida.org/research/topology/ as_core_network/2014/, 2014.

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

[17] H. Chang, S. Jamin, and W. Willinger. Internet connectivity at the as- level: an optimization-driven modeling approach. In Proceedings of the ACM SIGCOMM workshop on Models, methods and tools for reproducible network research, pages 33–46. ACM, 2003.

[18] H. Chang, S. Jamin, and W. Willinger. To peer or not to peer: Modeling the evolution of the internets as-level topology. Ann Arbor, 1001:48109–2122, 2006.

[19] H. Chang and W. Willinger. Difficulties measuring the internet's as-level ecosystem. In Information Sciences and Systems, 2006 40th Annual Confer- ence on, pages 1479–1483. IEEE, 2006.

[20] K. Chen, D. R. Choffnes, R. Potharaju, Y. Chen, F. E. Bustamante, D. Pei, and Y. Zhao. Where the sidewalk ends: Extending the internet as graph using traceroutes from p2p users. In Proceedings of the 5th international conference on Emerging networking experiments and technologies, pages 217–228. ACM, 2009.

[21] Q. Chen, H. Chang, R. Govindan, and S. Jamin. The origin of power laws in internet topologies revisited. In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceed- ings. IEEE, volume 2, pages 608–617. IEEE, 2002.

[22] R. Cohen and D. Raz. The internet dark matter-on the missing links in the as connectivity map. In INFOCOM, 2006.

[23] A. Dhamdhere and C. Dovrolis. Ten years in the evolution of the internet ecosystem. In Proceedings of the 8th ACM SIGCOMM conference on Internet measurement, pages 183–196. ACM, 2008.

[24] A. Dhamdhere and C. Dovrolis. The internet is flat: modeling the transi- tion from a transit hierarchy to a peering mesh. In Proceedings of the 6th International COnference, page 21. ACM, 2010.

[25] A. Dhamdhere and C. Dovrolis. Twelve years in the evolution of the inter- net ecosystem. IEEE/ACM Transactions on Networking (ToN), 19(5):1420– 1433, 2011.

[26] G. Di Battista, M. Patrignani, and M. Pizzonia. Computing the types of the relationships between autonomous systems. In INFOCOM 2003. Twenty- Second Annual Joint Conference of the IEEE Computer and Communica- tions. IEEE Societies, volume 1, pages 156–165. IEEE, 2003.

[27] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, G. Ri- ley, et al. As relationships: Inference and validation. ACM SIGCOMM Computer Communication Review, 37(1):29–40, 2007.

[28] X. Dimitropoulos, D. Krioukov, G. Riley, and K. Claffy. Classifying the types of autonomous systems in the internet. SIGCOMM Poster, 151, 2005.

[29] B. Donnet and T. Friedman. Internet topology discovery: a survey. Commu- nications Surveys & Tutorials, IEEE, 9(4):56–69, 2007.

[30] S. N. Dorogovtsev and J. F. Mendes. Evolution of networks: From biological nets to the Internet and WWW. Oxford University Press, 2013.

[31] N. Economides. The economics of the internet backbone. NYU, Law and Economics Research Paper, (04-033):04–23, 2005.

[32] B. Edwards, S. Hofmeyr, G. Stelle, and S. Forrest. Internet topology over time. arXiv preprint arXiv:1202.3993, 2012.

[33] Eu telecom’s rules. https://ec.europa.eu/digital-agenda/en/ telecoms-rules/.