Paths on the Autonomous Systems (AS) graph of the Internet derived from the Border Gateway Protocol (BGP) can help researchers to understand the dynamics of Internet topology and to interpret how that topology may enable nations to enact censorship, surveillance, or other Internet control measures. Unfortunately datasets of these paths are not readily publicly available, and paths collected from measurements such as traceroute measurements or BGP probes tend to be incomplete. Simulation frameworks have been used to generate AS paths based on empirical AS graphs. We have reimplemented one such framework, BGPSim [1] in Python as an efficient, open source BGP path simulation tool called BGPSimPy. With this tool we produced publicly available routing tree datasets that will enable future research on AS level topology. Previous research endeavors in this direction have only identified chokepoints in single snapshots. We apply the path simulation and routing tree datasets to view snapshots of the Internet over multiple years in order to understand how chokepoints at the national level may evolve to facilitate censorship. As an application of the routing trees we look at multiple qualitative sources of Internet freedom and censorship in order to interpret the relationship between AS level topology and actual censorship activity.

References

[1] Phillipa Gill, Michael Schapira, and Sharon Goldberg. Modeling on quicksand: dealing with the scarcity of ground truth in interdomain routing data. *ACM SIGCOMM Computer Communication Review*, 42(1):40–46, 2012.