

InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure – Public Review

David Wetherall
Google & University of Washington
wetherall@google.com

Internet mapping is a vital tool for understanding the design and operation of the Internet because it can cut across decentralized structures. Yet despite the advances in mapping over the years, we have had no public map of the underlying physical infrastructure – the wires and fiber-optic cables that make up the Internet.

Until now: the authors compiled such a map for the fiber-optic conduits between US cities. They did so by drawing on a wide variety of public sources, and then extracting and integrating this data with considerable manual effort. The paper describes the method of construction and presents a study of the resulting map. It tells us how ISPs share conduits, how routes follow roads and other rights of way, and more. It provides quantitative data where all we had was anecdotes.

The reviewers were unanimous in their appreciation for this effort and the value they place on the map, which is available as a resource for the research community. We expect it to feed into new research efforts, and hope it will improve in coverage and quality over time. The authors stretch too far in claiming the validation of the map data. They undertake consistency checking as they are able, but true validation in the sense of determining correctness will take time. As for the study of the map, it is compelling as an initial characterization of routes and conduit sharing, but not convincing as an assessment of risks due to shared physical infrastructure. While the authors reach in this direction, it is difficult to assess high-level properties such as risk without combining the map with other kinds of data.

A separate issue for our community is how to evaluate papers when they have associated data that will be made public as an integral part of their contribution. Judging submissions solely on the value of the paper without regard for associated data is our custom today, and was the standard applied to this paper. It makes clearing the acceptance bar that much harder, which is not much of an incentive to do the work. So here's a personal opinion: we should allow key data to be treated as part of the paper submission when it is offered up by the authors, and evaluate the combination.