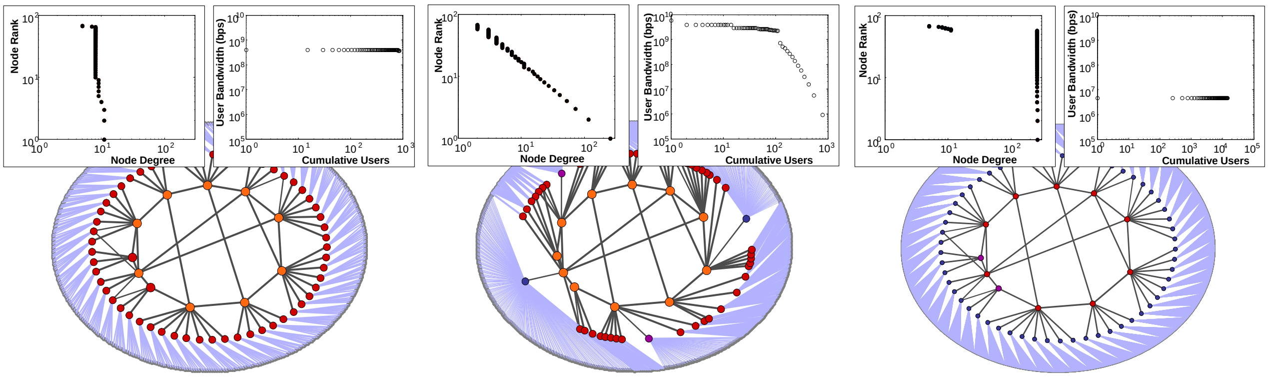

(a) (b) (c)

Figure 9: Distribution of node degree and end-user bandwidths for several topologies having the same core structure: (a) uniformly high bandwidth end users, (b) highly variable bandwidth end users, (c) uniformly low bandwidth end users.

is both simpler and more revealing. While there is nothing about our first-principles approach that precludes the incorporation of robustness, doing so would require carefully addressing the network-specific issues related to the design of the Internet. For example, robustness should be defined in terms of impact on network performance, it should be consistent with the various economic and technological constraints at work, and it should explicitly include the network-specific features that yield robustness in the real Internet (e.g., component redundancy and feedback control in IP routing). Simplistic graph theoretic notions of connected clusters [4] or resilience [42], while perhaps interesting, are inadequate in addressing the features that matter for the real network.

These findings seem to suggest that the proposed first-principles approach together with its implications is so immediate, especially from a networking perspective, that it is not worth documenting. But why then is the networking literature on generating, validating, and understanding network designs dominated by generative models that favor randomness over design and "discover" structures that should be fully expected to arise from these probabilistic models in the first place, requiring no special explanation? We believe the answer to this question lies in the absence of a concrete methodological approach for understanding and evaluating structures like the Internet's router-level topology. Building on [12, 48], this work presents such an approach and illustrates it with alternate models that represent a clear paradigm shift in terms of identifying and explaining the cause-effect relationships present in large-scale, engineered graph structures.

Another criticism that can be leveled against the approach presented in this paper is the almost exclusive use of toy models and only a very limited reliance on actual router-level graphs (e.g., based on, say, Mercator-, Skitter-, or Rocketfuel-derived data). However, as illustrated, our toy models are sufficiently rich to bring out some of the key aspects of our first-principles approach. Despite their cartoon nature, they support a very clear message, namely that efforts to develop better degree-based network generators are suspect, mainly because of their inherent inability to populate the upper-left corner in the likelihood-performance plane, where Internet-like router-level models have to reside in order to achieve an acceptable level of performance. At the same time, the considered toy models are sufficiently simple to visually depict their "non-generic" design, enable a direct comparison with their random counterparts, and explain the all-important tradeoff between likelihood and performance. While experimenting with actual router-level graphs will be an important aspect of future work, inferring accurate router-level graphs and annotating them with actual link and node capacities defines a research topic in itself, despite the significant progress that has recently been made in this area by projects such as Rocketfuel, Skitter, or Mercator.

Any work on Internet topology generation and evaluation runs the danger of being viewed as incomplete and/or too preliminary if it does not deliver the "ultimate" product, i.e., a topology generator. In this respect, our work is not different, but for a good reason. As a methodology paper, it opens up a new line of research in identifying causal forces that are either currently at work in shaping large-scale network properties or could play a critical role in determining the lay-out of future networks. This aspect of the work requires close collaboration with and feedback from network engineers, for whom the whole approach seems obvious. At the same time, the paper outlines an approach that is largely orthogonal to the existing literature and can only benefit from constructive feedback from the research community. In either case, we hope it forms the basis for a fruitful dialogue between networking researchers and practitioners, after which the development of a radically different topology generator looms as an important open research problem.

Finally, we do not claim that the results obtained for the router-level topology of (parts of) the Internet pertain to logical or virtual networks defined on top of the physical infrastructure at higher layers of the protocol stack where physical constraints tend to play less of a role, or no role at all (e.g., AS graph, Web graph, P2P networks). Nor do we suggest that they apply directly to networks constructed from fundamentally different technologies (e.g., sensor networks). However, even for these cases, we believe that methodologies that explicitly account for relevant technological, economic, or other key aspects can provide similar insight into what matters when designing, understanding, or evaluating the corresponding topologies.

Acknowledgments

Reiko Tanaka assisted in the preparation of an early version of this work. We are indebted to Matt Roughan, Ramesh Govindan, and Steven Low for detailed discussions of router-level topologies and traffic modeling, Stanislav Shalunov for data on the Abilene network, and Heather Sherman for help with the CENIC backbone.

7. REFERENCES

[1] Abilene Network. Detailed information about the objectives, organization, and development of the Abilene network are available from http://www.internet2.edu/abilene.

[2] W. Aiello, F. Chung, and L. Lu. A Random Graph Model for Massive Graphs. *Proc. STOC 2000.*