

Internet Science Starts at Home: Integrating Residential Network Data into the Internet Measurement Space

Amy Csizmar Dalal, Carleton College

While telemedicine, remote work, and distance learning existed prior to 2020, the global COVID-19 pandemic drove a mass, immediate movement towards these services. Our continued reliance on networking within the home highlights inequities in last mile connections — whether a household can connect at all, or connect at the broadband speeds necessary for remote work and distance learning — as well as the precariousness of networked application performance at large scale. Home networks are highly heterogeneous, from the last mile connection, to the network topology within the home, to the mix of devices and applications, to diurnal usage patterns. As we continue to rely heavily on home connectivity, a trend likely to persist well after the pandemic ends, understanding the complex ways in which these areas of heterogeneity interact to affect both application performance and network performance, inside and outside the home, is crucial.

The critical research questions for the community are these: How do we assess network, and networked application, performance within the home environment? How do we integrate this assessment with Internet performance data to achieve a *holistic understanding* of network performance and of end users' quality of experience (QoE)?

Any solution requires sharing data collected from heterogeneous environments that has been anonymized and aggregated to preserve user privacy, and integrating this data with broader Internet measurements to identify data patterns associated with specific QoE degradations. The key initial issues are:

- **Data identification.** We lack a uniform definition of QoE for networked applications and an agreed-upon way to measure QoE. How do we define performance and set performance levels? Are these levels user-specific, context-specific, application-specific, time-of-day-specific, or some combination of these?
- **Data collection infrastructure.** Extreme heterogeneity precludes a one-size-fits-all approach. Most routers, applications, and devices do not provide an open interface for data collection. Deploying infrastructure within the home requires the cooperation of the residents as well as the ISP, and cannot rely on residents having a particular level of technical expertise.
- **Data sharing.** Home network data reveals sensitive aspects of users' lives. Privacy concerns need to be front and center, both for data collection within a single residence and for the eventual sharing of data. Even assuming data can be shared, determining how to utilize this data to determine the root cause of degraded QoE is a wicked problem.

The home network is now the primary “user interface” to the Internet. From a usability standpoint, it fails spectacularly: it provides little to no actionable or understandable feedback; it requires expert technical knowledge to navigate; its outputs cannot be easily generalized across home environments. The Internet similarly lacks a “user interface” into the home space, as home network performance is opaque to the broader Internet. Performance transparency, vital to the growth and evolution of present and next-generation applications and protocols, requires the full cooperation and participation of the Internet research community.