

How are different user demographics actually experiencing the internet?

Henning Schulzrinne
Columbia University

December 2020

Traditional end-user (residential) measurements have been focused on abstract networks, largely assuming that networks, both across networks and within a single provider, were homogeneous. For practical reasons, larger-scale longitudinal measurements also focused on the largest providers. For example, the FCC *Measuring Broadband America* effort only captures performance across about a dozen very large service providers, largely serving urban and suburban areas. It omits smaller rural carriers and modalities such as fixed wireless access (FWA). Measurements by commercial providers such as Ookla are generally spot measurement, i.e., initiated by the user on demand, with no indication of the speed that the user subscribed to.

We know through a variety of non-engineering metrics, however, that the experience of the internet can differ significantly across geography and even demographics. Similarly, spot measurements such as those performed in Pennsylvania by M-Labs indicate high variability in the internet speed experienced. For example, it is plausible that rural areas and urban low-income areas suffer from lower reliability, e.g., due to lack of maintenance and longer loops, and lower effective speeds. Low-income households in MDUs may experience higher Wi-Fi congestion, reducing their effective bandwidth well below their carrier-provided bandwidth. However, all of these are speculation, with hints of data, even though they raise fundamental questions of digital equity. Average numbers or simple distributions tend to hide the underlying causes for different performance, which are often not strictly technical limitations, but deployment and operational choices by network providers.

Separately, we know relatively little about the adoption of broadband services across user groups. For example, do users who are relatively new to the internet differ significantly from more experienced users and does their behavior converge? Are there distinctions by age, education and family size? Do families with lower speeds suffer from temporary or longer-term network congestion, either because of limitations in their access link or issues related to Wi-Fi congestion or outdated equipment?

In some cases, we may even be able to deduce, in near real-time and while preserving user anonymity, changes in household behavior. For example, we can use network performance data to see if pandemic-induced stay-at-home orders are visible in network usage data [1], possibly separated by region or other demographic indicators.

As raw speed and other basic networking metrics, averaged over large user populations, are starting to be less relevant to how users experience the internet, we should structure data gathering efforts to address some of the digital equity concerns and help inform policy making in digital inclusion. This includes selecting measurement locations that reflect the demographics of internet users, and capturing those demographics. It would be helpful for the social science community and measurement community to collaborate on identifying which indicators are of primary concern (e.g., rurality, race, income, education, age of housing stock).

References

- [1] Jessica De Oliveira Moreira, Amey Praveen Pasarkar, Wenjun Chen, Wenkai Hu, Jan Janak, and Henning Schulzrinne. Social distancing and the internet: What can network performance measurements tell us? arXiv, 2020.