A Collaborative Crowdsourced User-Carrier-App Ecosystem to Enable Next Generation Wireless Research

Zhi-Li Zhang, Feng Qian, Arvind Narayanan, Z. Morley Mao[§]
University of Minnesota, [§]University of Michigan

Many past research studies have developed tools to measure and characterize the deployed cellular and WiFi networks in order to further our understanding for such deployed networks to innovate at various layers (e.g., application, network, transport, MAC). However, with the recent rollout of 5G, it poses significantly more unique challenges and needs more measurement support. To address this challenge, we propose a collaborative crowdsourced user-carrier-app ecosystem that would enable researchers to get more fine-grained control of the measurement methodologies and collect relevant context information that impact performance. Ideally, we would incorporate some limited collaboration mechanism from the cellular/network operators such that the resulting measurement-based experimentation can be interpreted more meaningfully. We argue that providing the right interface to incentivize provider's collaboration is the key to overcome the barrier to collaboration with commercial cellular ISPs.

For cellular system research, there remain two challenges: (1) a lack of visibility into commercial 5G's lower layers, and (2) a lack of cross-layer, large-scale 5G network data for research. These challenges were known in the 3G/4G era due to the closed cellular ecosystem (compared to WiFi), the complexity of 3GPP protocols, the monetary cost of conducting cellular networks, to name a few. In 5G, these challenges have become even more prominent due to 5G's high complexity, protocol heterogeneity and deployment strategies (e.g., low-band vs. mid-band vs. mmWave-band 5G, standalone vs. non-standalone 5G), and diverse performance range. For example, mmWave 5G is highly sensitive to the physical environment, mobility, and even weather; this raises the bar of collecting data that comprehensively captures the 5G performance in diverse contexts. It is also important to note that 5G is supposed to have a more programmable core: understanding how to leverage that programmability to enable new network services can be an important goal for measurement studies.

To summarize, we propose to discuss the following questions at the workshop and the associated issues from the goals of "Workshop Overview": How to design a sustainable crowdsourced user-carrier-app ecosystem that incorporates optional support from ISPs to enable the next-generation wireless research, for the short-term, 5G research activities across layers?

- What data is needed? The data needs for 5G requires more significant details on the context of the device (e.g., orientation, line of sight, weather, types of 5G services available: low/mid/high-band 5G, SA/NSA 5G, etc.). A deep understanding of the cellular operator's infrastructure and network policies is also critical as traffic appears to be differentiated based on the traffic types.
- What infrastructure is needed to collect such data? Beyond just the data collection from devices with cellular connectivity, software radios with lower-layer access are essential to help interpret the lower layer control channel data that is not exposed by commercial smartphones. A distributed infrastructure to control multiple devices to perform measurement concurrently is essential to understand concerns such as interference.
- Should the community collectively tackle the issues of equitable and sustainable data sharing and curation? Yes, this is essential, as we would like to support an evolvable platform that will support future generations of wireless technologies, not just for 5G.
- How can we develop best practices to facilitate cooperation or collaboration with commercial service providers as we collect data? Carriers are incentivized to support such measurement efforts if they can benefit from a better understanding of their users' performance experiences, and they also hope their networks to be characterized in a representative manner. We want to take advantage of such incentives by offering an interface that allows different degrees of collaboration. For example, we can provide the lowest support through some binary questions about whether there was some maintenance event that explained the observed performance issues in contrast to more detailed input on when maintenance events occurred.