Quality Measurement over Quality Data Reuse in Cellular Networks

Zizheng Liu, Chunyi Peng
Purdue University

Cellular network measurement is important and hard, particularly for academic community. The well-known barrier is no open data collected from operational networks at scale. Despite efforts and attempts from the government (e.g, FCC's mobile broadband speedtest [1] and Bureau of Cyber Statistics [2]), the broad academic community can not obtain large-scale datasets from network operators (say, AT&T, Verizon, T-Mobile), measurement companies (e.g., Ookla, OpenSignal) and mobile app providers (e.g., Google, Apple, Amazon, Facebook). Unfortunately, there is little that we (academia) can do to change it in the foreseen future [3].

Question. Our question is what we can do to improve cellular network measurement, particularly overcome the data barrier. We focus on *technical* aspects as the answer can easily turn into actions to take.

We do have data but data is hard to reuse. A number of research groups have conducted a wide variety of measurements over cellular networks; They publish papers and most of them release (are willing to release) their datasets and even tools for data collection. To encourage research reproducibility and data reuse, the community have created a badging system to assess the quality of these artifacts (three badges assigned by SIGCOMM Artifacts Evaluation Committee [4]). Despite these efforts, data is still hard to reuse with two technical challenges.

First, the reuse barrier is rooted in the nature of measurement research. Data is collected for one specific research objective A and is not easy to reuse for another objective B. Distinctive objectives often call for different experiments and datasets. *But can the dataset for A be partially reused for B, if not fully reused?*

Second, cellular network measurement is more sensitive to dynamic environment, which limits the reuse value of data, especially data collected from small-scale experiments. Findings might conceptually hold but numeric results are often biased, largely affected by experiment settings such as areas (with distinct cellular network deployment and coverage), time (busy or idle hours with different network traffic loads), mobility (static, walking or driving), user traffic (e.g., single or multiple mobile apps, heavy or light, streaming or bursty). Similar to the wired broadband measurement, it is also ad-hoc and opportunistic [5]. *But can we take these factors into account and extract benchmark for data reuse in different settings?*

What we can do? We believe that the answers are promising. Based on lessons and insights gained in our past experience, we recommend the following several actions plausible for the community.

◦ *Design modular experiments.* An experiment consists of multiple standard or common modules, which makes easy to learn common experiment parts between two research projects and infer whether (and how) data can be reused. For instance, we have conducted two distinct studies to measure network latency [6,7] and handover quality [8,9]. Both run similar traffic (at least ping every second); It is hard to keep running latency experiments for a five-year longitudinal study, but the dataset for the handover study is largely reusable and make this longitudinal study possible [10].

◦ *Breakdown analysis.* Measurement is hard because it is impacted by a number of factors. Decoupling complexity can be done via breakdown analaysis. Take the latency measurement as an example. End-to-end latency is contributed by the transmission time on the last-hop radio link and non-radio links, sequential functions at different layers (e.g., from the application and DNS to physical radio link establishment), changing overhead in the control-plane and data-plane [7]. End-to-end latency is actually a sum of multiple parts $\sum \tau_i$ through vertical latency breakdown. It makes easier to reuse datasets from other research groups or plan experiments in a cost-effective manner; Researchers can reuse third-party datasets to quantify one latency component τ_i, if not all latency components. Researchers can also run more experiments to measure certain latency components of their focus (e.g., latency on radio links) given their limited resources. This complements the above modular experiment design. To make it, we need to collect not only primary data (here, latency performance samples) but secondary data to understand why (here, signaling messages needed for breakdown analysis).

◦ *Recommend baseline experiments.* Conducting a variety of measurement experiments repeatedly to build up datasets is time, manpower, and fund consuming. A cost-effective way is to conduct baseline experiments to measure common bottlenecks in a largely controlled manner. Other experiments can be built up over these baselines with their own experiment variables. Specifically, the last-hop radio access quality is of common interests in cellular network research. Conducting several benchmark experiments to measure radio link quality under typical traffic settings is much easier and less resource-consuming than running end-to-end measurements that traverse both radio links and non-radio links with more experiment variables to consider. We recommend baseline experiments in a controlled manner to minimize the impacts beyond our control (e.g., accompanying traffic along with speedtest impacts the performance results). Inexpensive experiments are preferred; Most measurements want to reduce data usage given that cellular data plan is not really unlimited. For research projects that do not depend on heavy traffic (e.g., studies on radio coverage, blockage in 5G, control-plane signaling, security), light traffic (e.g, ping) is much more cost-effective than speedtest. In a nutshell, we want to build up high-quality, small data as the core which are extended and expanded with lower-quality (more noisy) and bigger data.

◦ *Develop advanced ML algorithms for data reuse.* Advanced ML algorithms, e.g., transfer learning, multi-task learning, can maximize the power of data reuse. Clearly, we should develop customized algorithms with domain knowledge in cellular network functions and procedures. For example, cellular network functions and protocols can be modeled as finite state machines (FSM); Exploiting FSMs obtained from 3GPP standards or inferred from practice [11] is of great value to design ML algorithms to reuse data relevant to a sequence of states in FSMs to study network behaviors.

No doubt, there is more to do, technically and non-technically, as discussed at WOMBIR [12], including not limited to fund incentives for measurement researchers, building open experimentation platforms (e.g, PAWR), enhancing academia-industry collaboration etc. Lowering the technical barrier of data reuse in cellular networks is quite achievable by the community. The most recommended action is to define standards and recommendations on how to design experiments, annotate data, share artifacts, build benchmarks and streamline common data processing, and/or support and reward efforts that empower them.

References

[1] FCC. Measuring mobile broadband, 2014 - 2021. https://www.fcc.gov/general/measuring-mobile-broadband-performance.

[2] Considerations for the structure of the bureau of cyber statistics, 2020. https://www.lawfareblog.com/considerations-structure-bureau-cyber-statistics.

[3] David Clark and Kc Claffy. The role of academic research in a bureau of cyber statistics, 2021. NSF Workshop on Overcoming Measurement Barriers to Internet Research (WOMBIR 2021).

[4] Sigcomm artifacts evaluation committee, 2017 - 2021. https://www.sigcomm.org/content/sigcomm-artifacts-evaluation-committee.

[5] Mark Allman. Moving from opportunistic to systematic measurement, 2021. NSF Workshop on Overcoming Measurement Barriers to Internet Research (WOMBIR 2021).

[6] Yuanjie Li, Zengwen Yuan, and Chunyi Peng. A Control-Plane Perspective on Reducing Data Access Latency in LTE Networks. In *ACM International Conference on Mobile Computing and Networking*, MobiCom'17, pages 56–69, 2017.

[7] Zengwen Yuan, Yuanjie Li, Chunyi Peng, Songwu Lu, Haotian Deng, Zhaowei Tan, and Taqi Raza. A Machine Learning Based Approach to Mobile Network Analysis. In *IEEE International Conference on Computer Communication and Networks*, ICCCN, 2018. (Invited Paper).

[8] Haotian Deng, Chunyi Peng, Ans Fida, Jiayi Meng, and Charlie Hu. Mobility Support in Cellular Networks: A Measurement Study on Its Configurations and Implications. In *ACM Internet Measurement Conference*, IMC'18, 2018.

[9] Haotian Deng, Qianru Li, Jingqi Huang, and Chunyi Peng. iCellSpeed: Increasing Cellular Data Speed with Device-Assisted Cell Selection. In *ACM International Conference on Mobile Computing and Networking*, MobiCom'20, 2020.

[10] Zhehui Zhang Zhaowei Tan Haotian Deng Jinghao Zhao Qianru Li Yunqi Guo Kai Ling Boyan Ding Hewu Li Songwu Lu Yuanjie Li, Chunyi Peng. Experience: A Five-Year Retrospective of MobileInsight. In *ACM International Conference on Mobile Computing and Networking*, MobiCom'21, 2021.

[11] Yuanjie Li, Chunyi Peng, Zengwen Yuan, Jiayao Li, Haotian Deng, and Tao Wang. MobileInsight: Extracting and Analyzing Cellular Network Information on Smartphones. In *ACM International Conference on Mobile Computing and Networking*, MobiCom'16, 2016.

[12] Nsf workshop on overcoming measurement barriers to internet research (wombir 2021), 2021. https://www.caida.org/workshops/wombir/2101/.