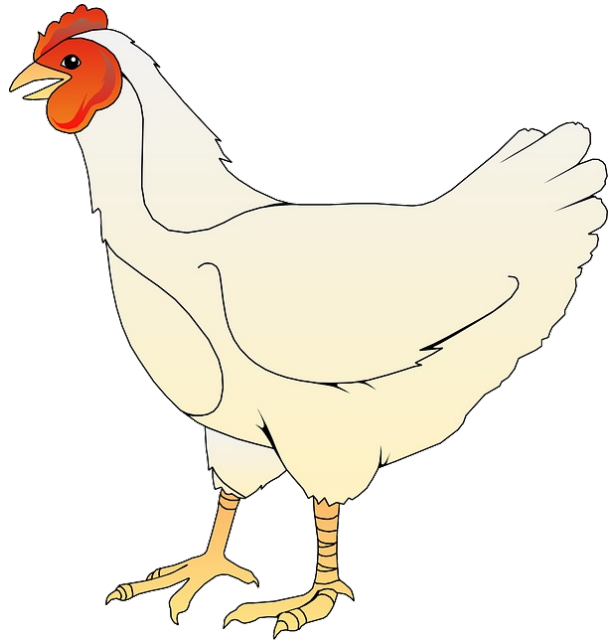


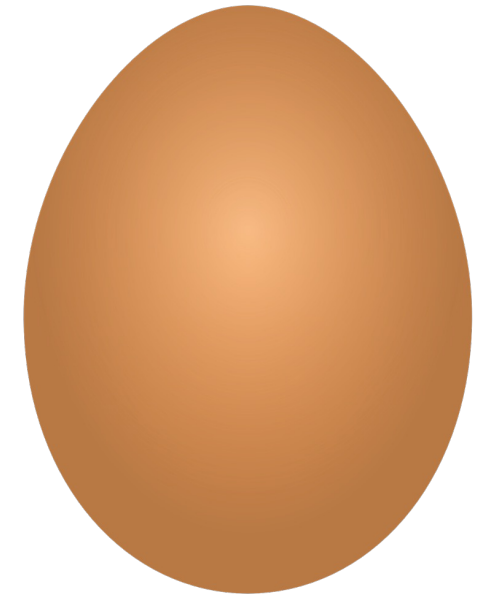
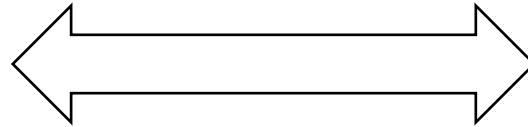
A First Look at the Spatial and Temporal Variability of Internet Performance Data in Hyperlocal Geographies

Taveesh Sharma, Jonatas Marques, Nick Feamster, Nicole Marwell
University of Chicago

The Internet Measurement Sampling Problem



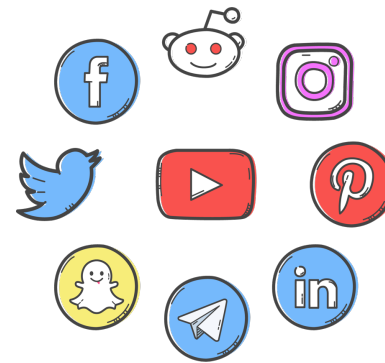
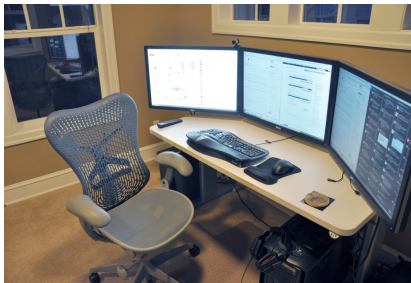
Better data



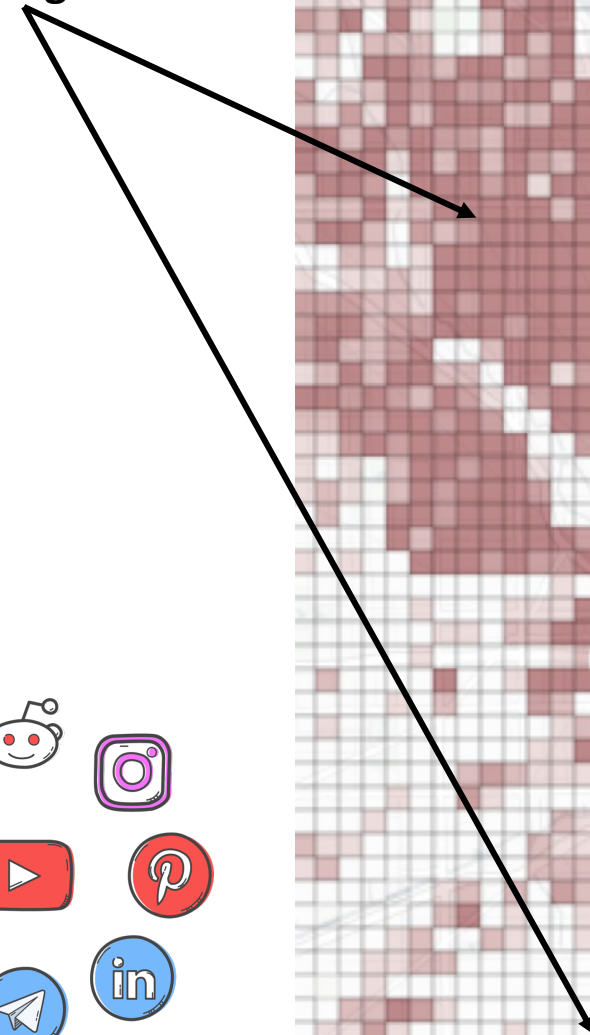
High Quality Samples

Motivation

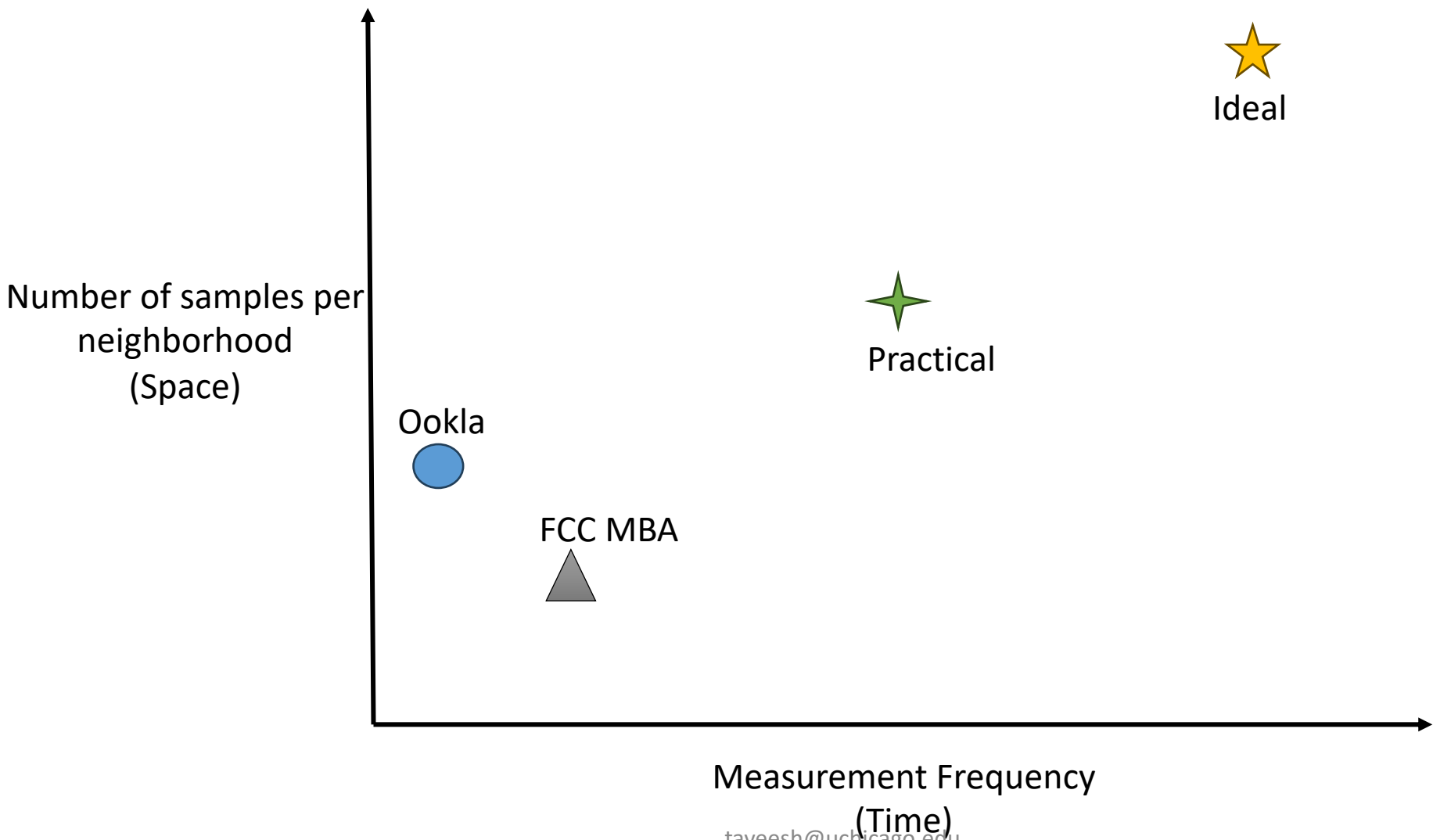
- Internet access is critical to work, education, healthcare and day-to-day communication
- Not all users receive the same level of internet services across **space** and **time** (Sharma, 2022)
- Existing broadband speed test datasets lack coverage across both dimensions (Saxon, 2022)



Sampling bias

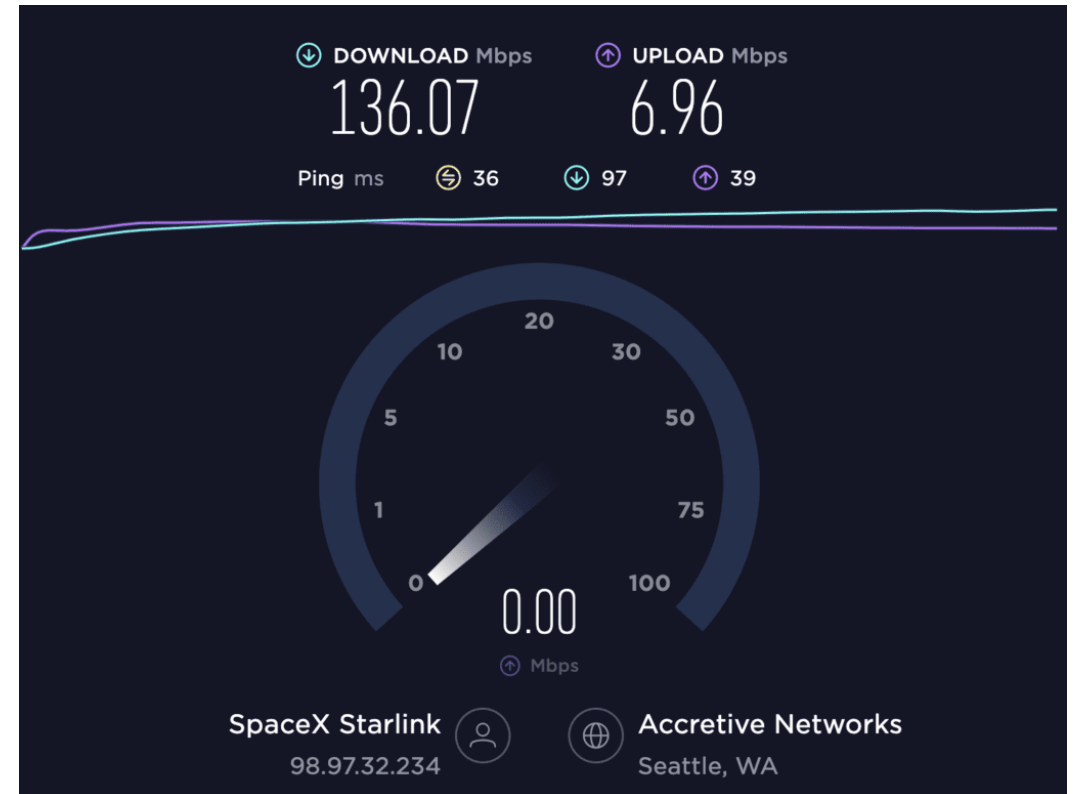


The Sampling Tradeoff



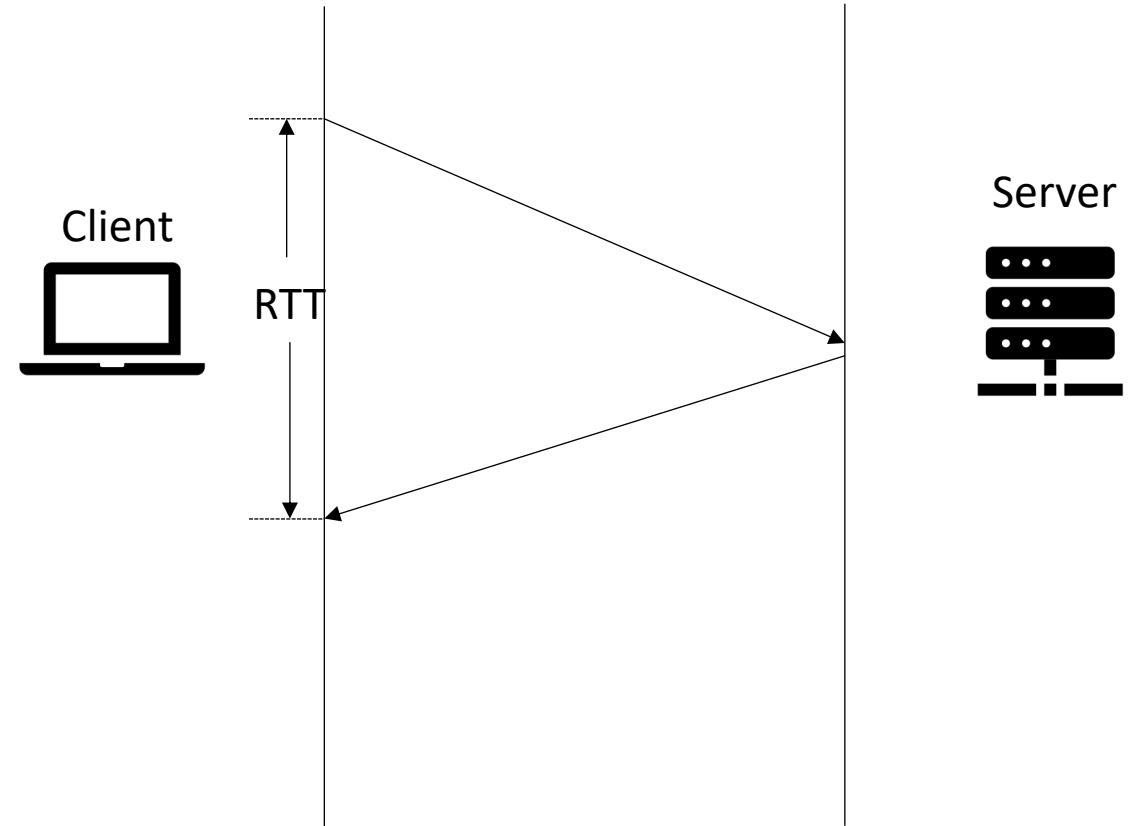
Challenges

- High data consumption overheads
- Different methodologies offered by a range of speed test tools (Macmillan, 2023)
- Limited visibility into the network



Using Lightweight Measures of Performance

- Round Trip Time (RTT) – popular measure of latency
- Latency Measurements can be collected at high granularity and lower costs
- Latency correlates highly with geographical distance between the client and the testing server (Landa, 2013)

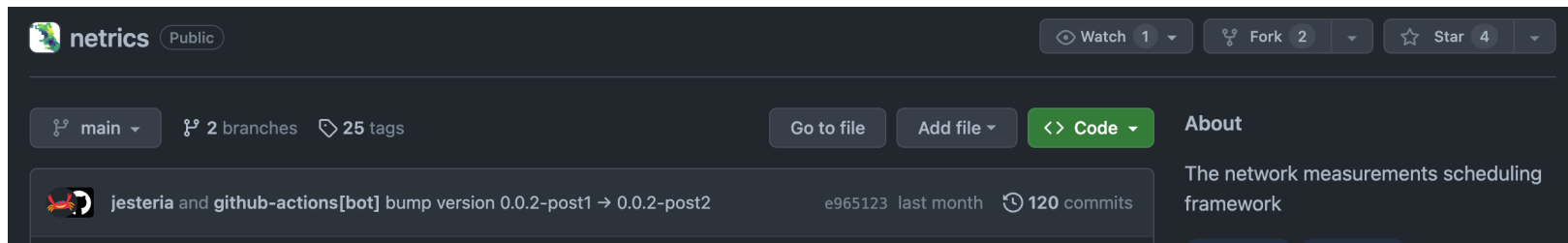
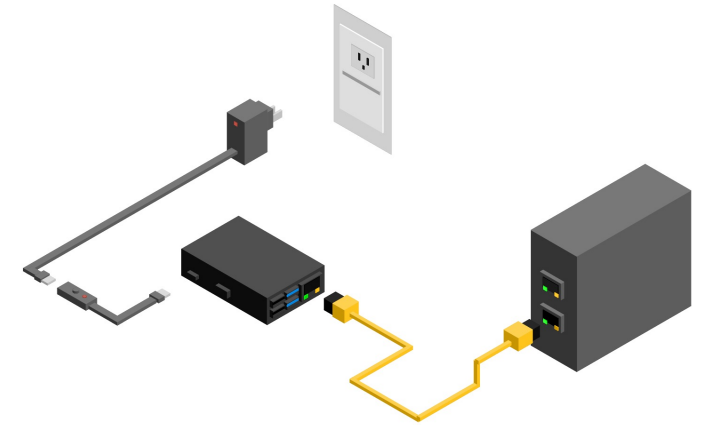


Research Question

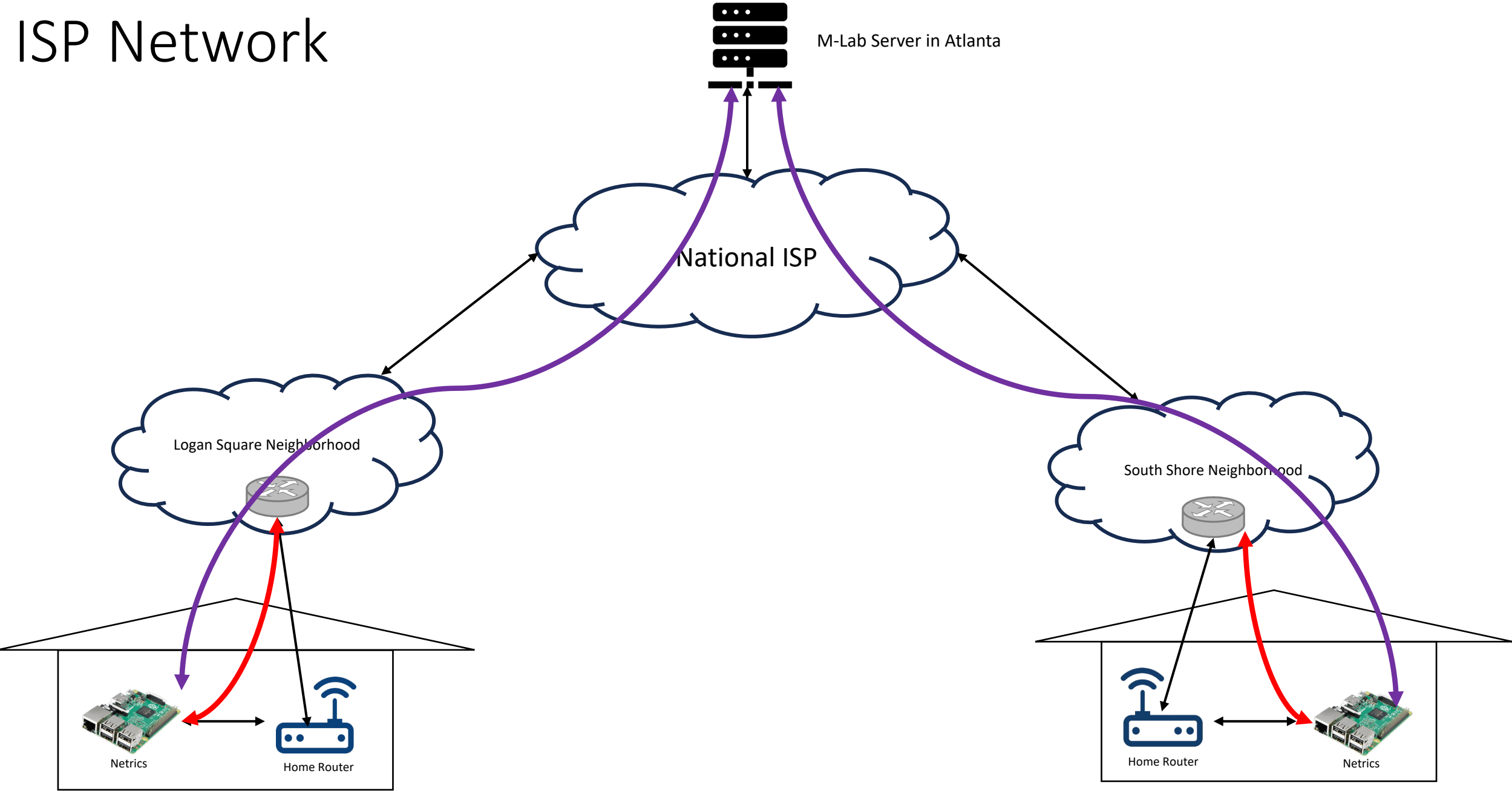
How can we use latency data to devise measurement sampling strategies across space and time?

Dataset

- 88 Netrics devices
- 22 Chicago neighborhoods
- 2021 – Present day
- One latency measurement every 5 minutes
- 1-5 speed test measurements everyday



ISP Network

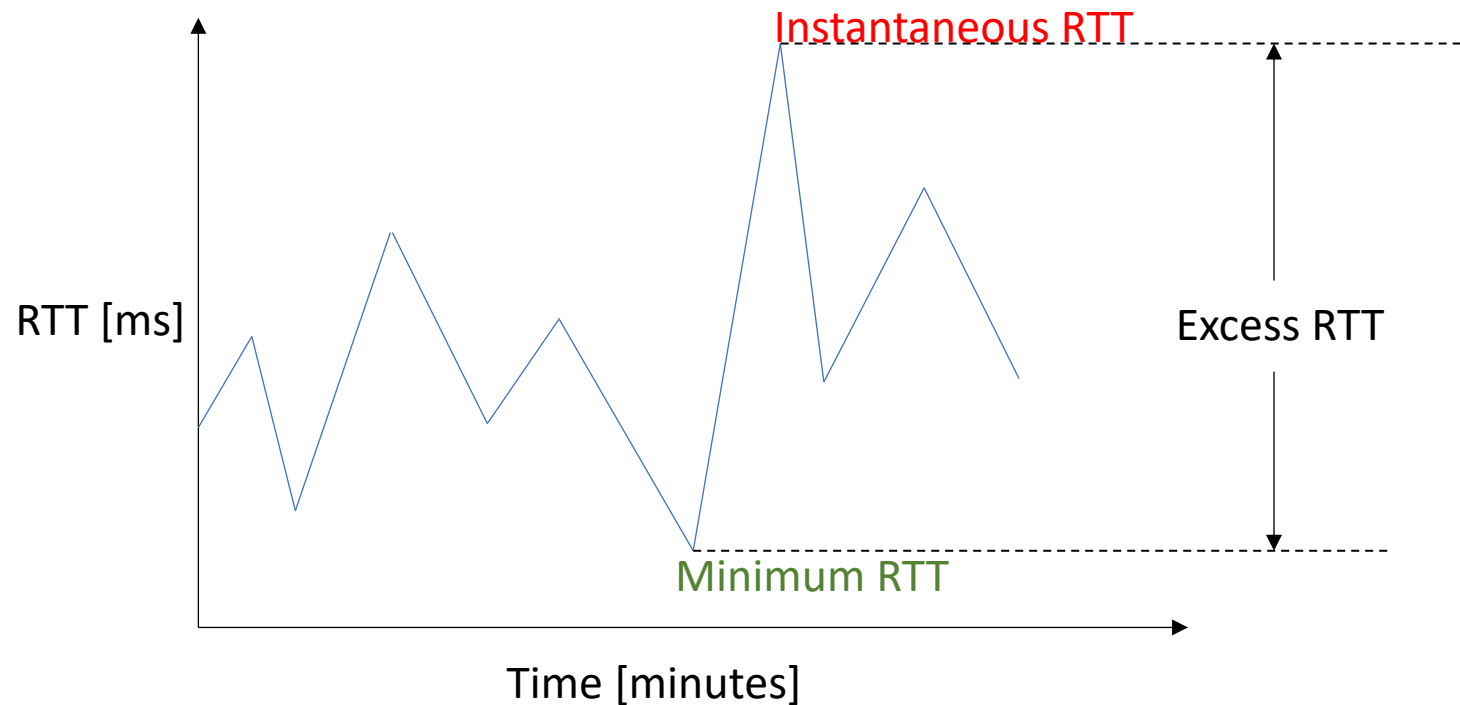


Logan Square Home

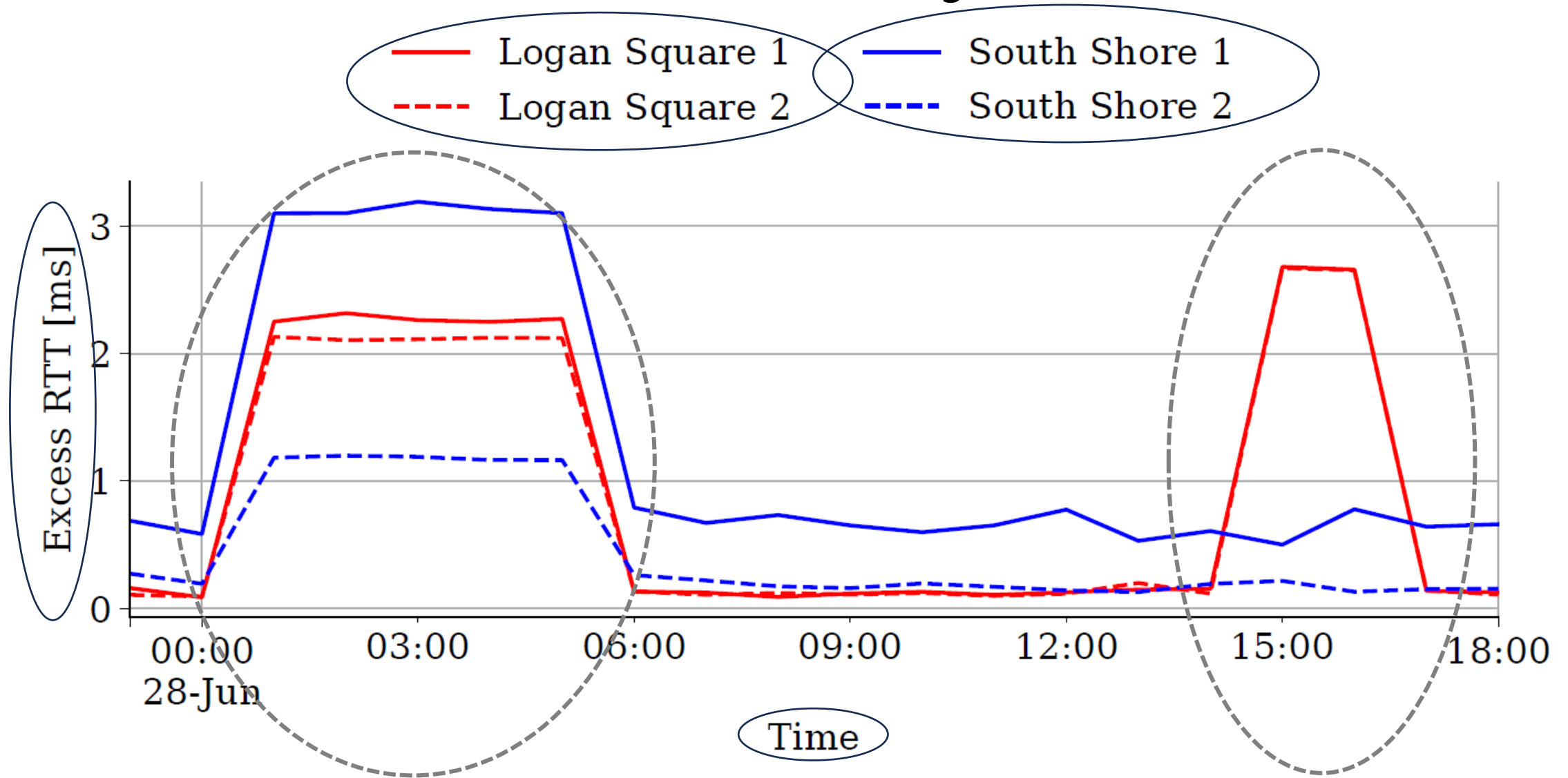
South Shore Home

Excess Round Trip Time (eRTT)

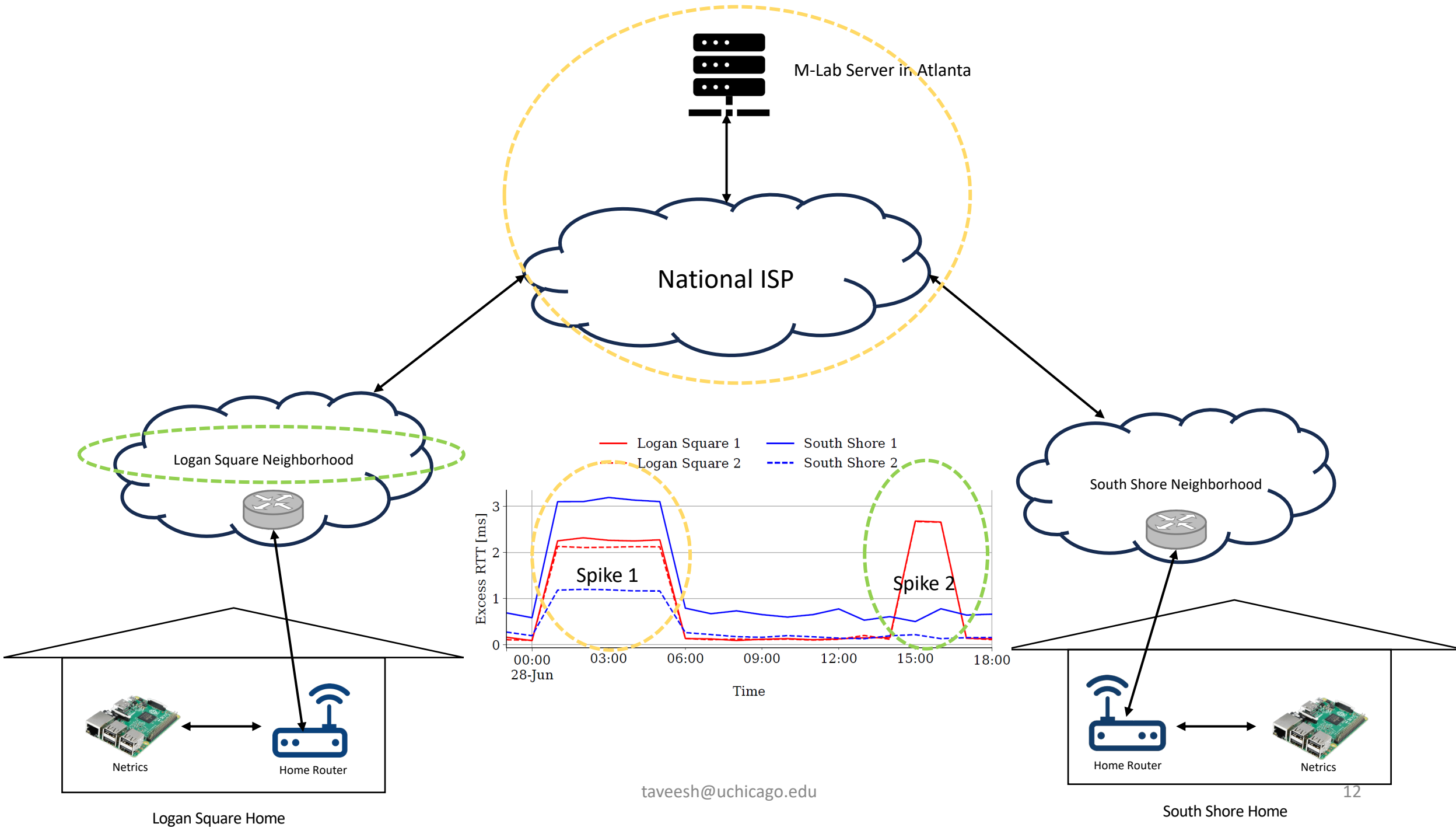
Difference between an **instantaneous RTT sample** and the **minimum RTT** across a time window



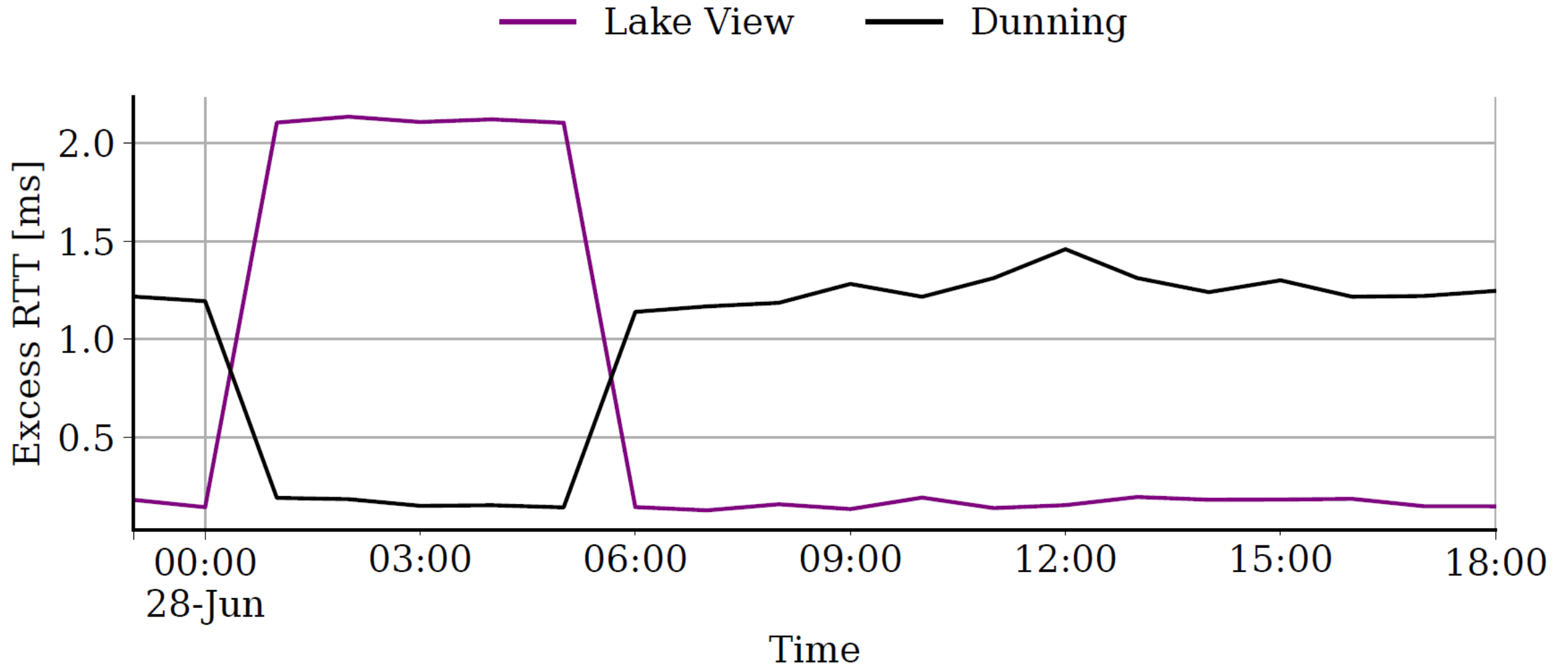
“Devices Sharing a persistent end-to-end latency spike may not always be located in the same neighborhood”



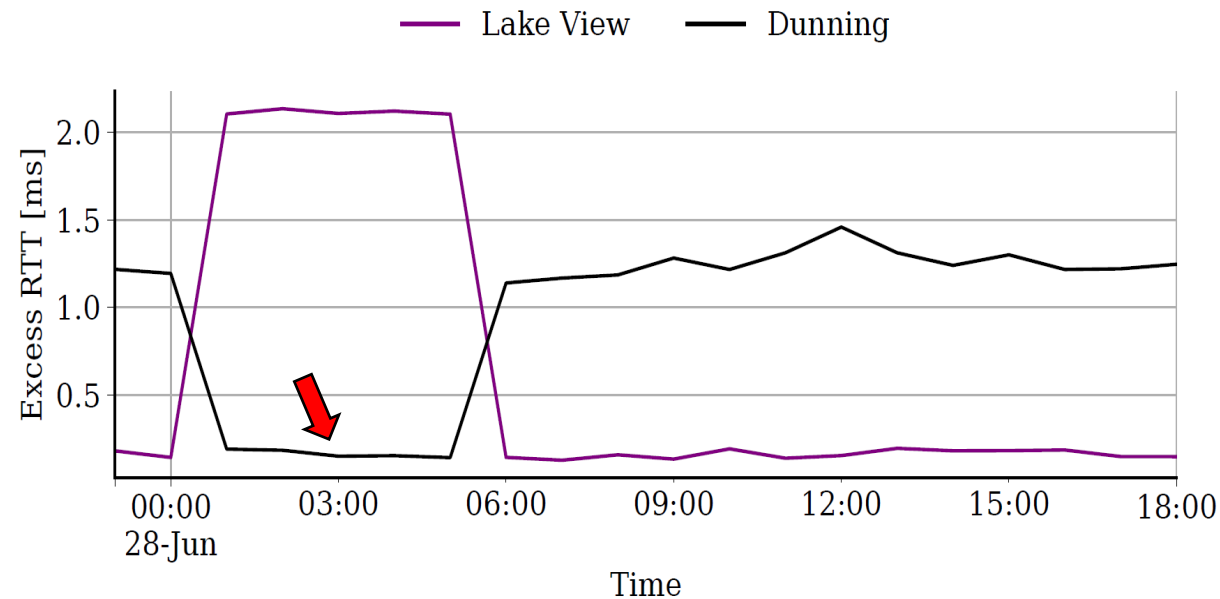
No persistent spikes observed across the same time at the last mile



Abnormal behavior around Spike 1 but flat lines around Spike 2



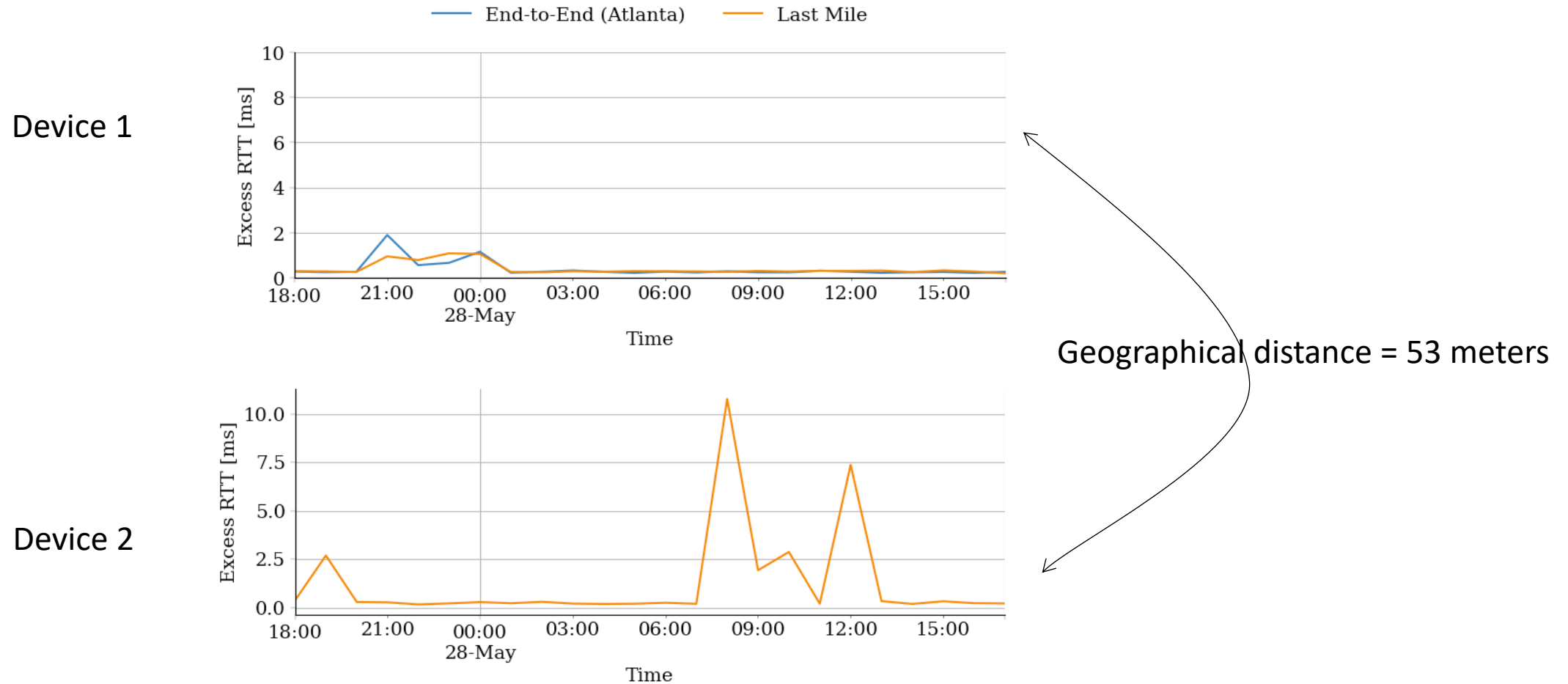
What about the dip for Dunning?



Effect of load balancing applied by the ISP:

- The device in Dunning routed through a less congested national router
- Other devices assigned different routers with higher loads around the same time
- Change of paths occurring around midnight of 28th June

“Devices located very close to each other may not always exhibit similar last-mile latency spikes”



Takeaways

- The uncertainty in locating network bottlenecks makes sampling hard
- Spatial proximity does not imply similarity in latency trends
- Data ↔ Samples
 - Deploy multiple devices in the same building?
 - Understand how load balancing is implemented?
 - Measure additional links along the path?

Come work with us!

For more information

Visit our website to check our data and code!

<https://internetequity.uchicago.edu/>

Internet Equity Initiative Contacts:

Nzinga-Ain Barberousse, Community Project Manager, nzingaain@uchicago.edu

Marc Richardson, Technical Project Manager, mtrichardson@uchicago.edu

Alexis Schrubbe, Executive Director, schrubbe@uchicago.edu

Questions?

References

- Sharma, Ranya, Tarun Mangla, James Saxon, Marc Richardson, Nick Feamster, and Nicole P. Marwell. "Benchmarks or Equity? A New Approach to Measuring Internet Performance." *A New Approach to Measuring Internet Performance (August 3, 2022)* (2022).
- Saxon, James, and Dan A. Black. "What we can learn from selected, unmatched data: measuring Internet inequality in Chicago." *Computers, Environment and Urban Systems* 98 (2022): 101874.
- Landa, Raul, Richard G. Clegg, Joao Taveira Araújo, Eleni Mykoniati, David Griffin, and Miguel Rio. "Measuring the relationships between Internet geography and RTT." In *2013 22nd International Conference on Computer Communication and Networks (ICCCN)*, pp. 1-7. IEEE, 2013.