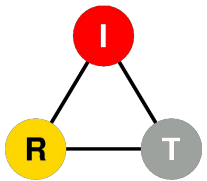


# Social Distancing and the Internet: What Can Network Performance Measurements Tell Us?

Jan Janak

with Jessica De Oliveira Moreira, Amey Praveen Pasarkar,  
Wenjun Chen, Wenkai Hu, and Henning Schulzrinne



Dept. of Computer Science, Columbia University  
February 17, 2021

Internet Real Time Laboratory: <https://www.cs.columbia.edu/irt/>



# Research Questions

1. Impact of COVID-19 on US fixed broadband internet performance?
  - a. look at throughput, packet loss, upstream performance (video conferencing)
  - b. disaggregate by population density, geography, time period
2. Correlate internet usage with changing user behavior?
3. Response of users to COVID-19 measures from network metrics?
  - a. daytime and evening data usage (work from home, distance learning)
  - b. stay-at-home on weekends

# FCC Measuring Broadband America (MBA)

- US broadband internet performance measurements since 2011
- Covers major ISPs, geographic regions, connection types
- 4,000-5,000 active volunteers (~7,500 installed test units in total)
- SamKnows Whitebox: Active & passive performance measurements
- Annual reports covering ~ 80% of population
- Raw data:
  - ~ 7,500 test units, 20 GB per month
  - Upload / download speed, latency, packet loss
  - Web, VoIP, and video streaming performance
  - Anonymized user data usage



Source: <https://samknows.one/>

# Downloads by Population Demographics

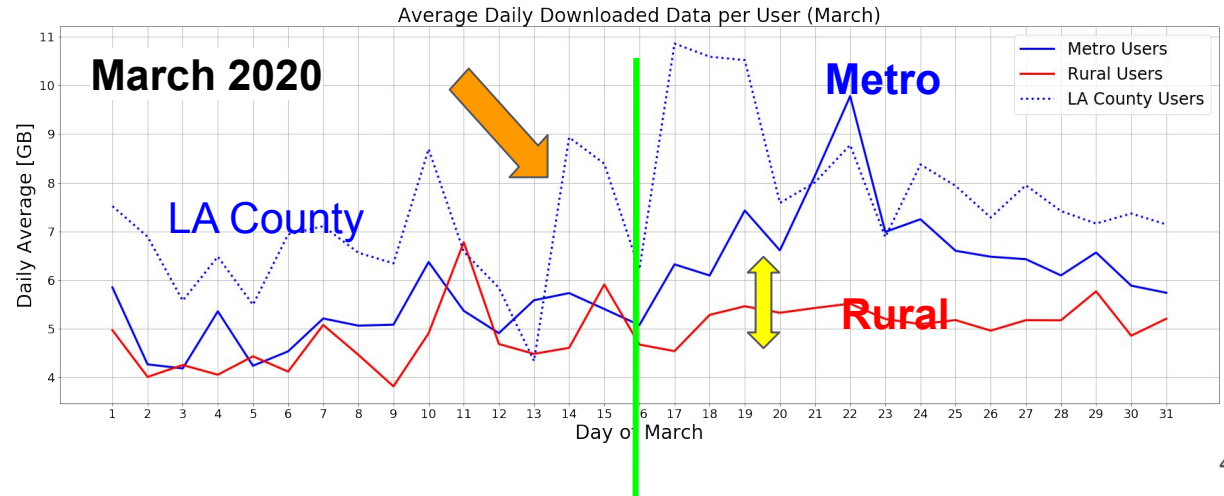
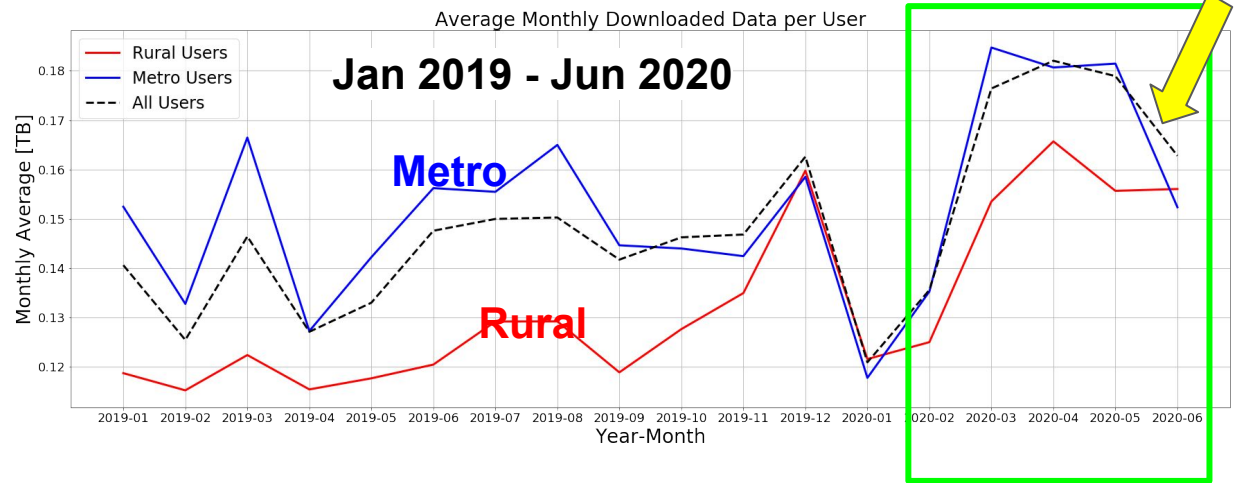
2019 US Census Data

Rural: < 100,000

Metro: > 1,000,000

20% of dataset is rural

44 units in LA county



# Hourly Downloads

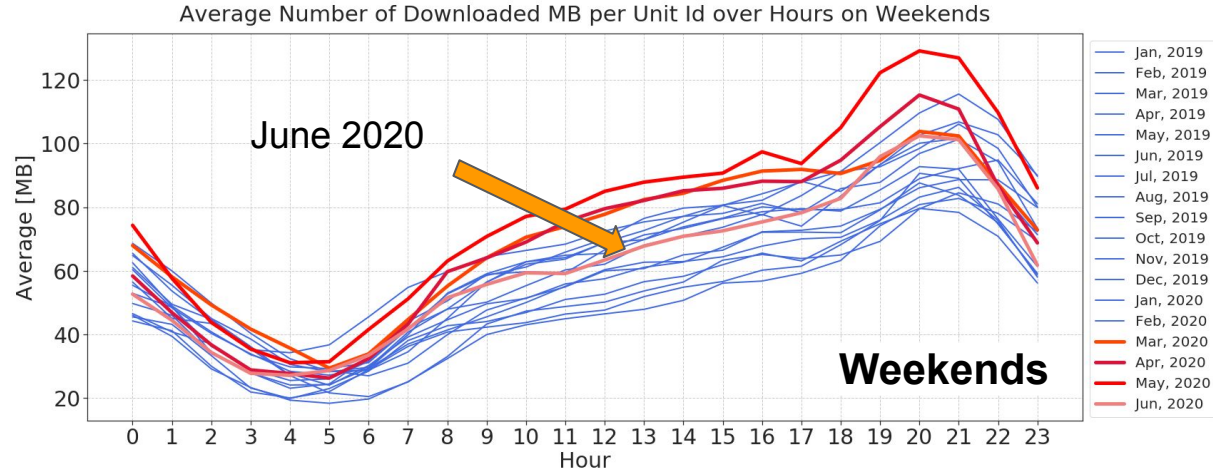
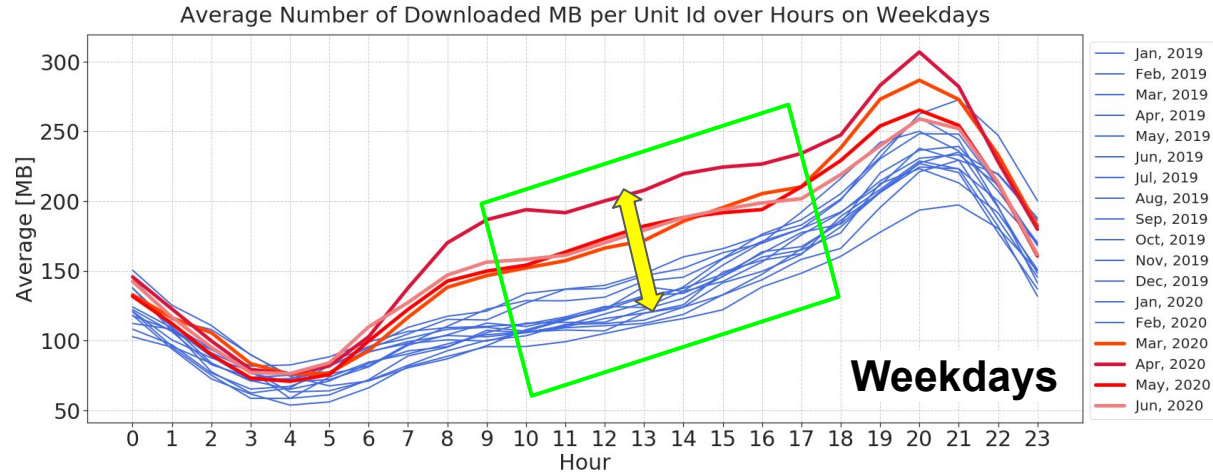
2019 vs 2020, ~ 3,000 test units included

Higher values on weekdays

Highest levels in April 2020

Weekday pattern resembles weekend

Return to normal on weekends in June 2020

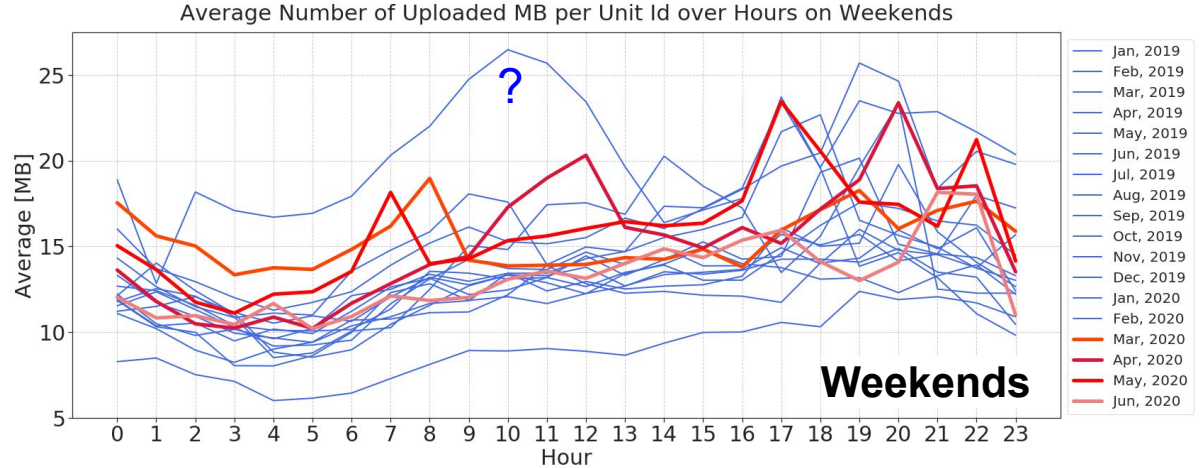
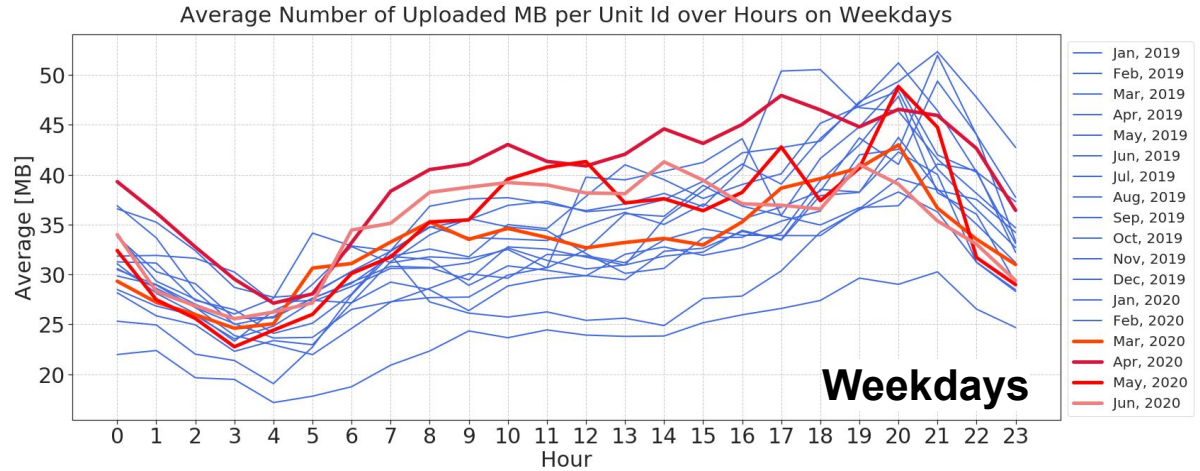


# Hourly Uploads

Very noisy data

Little difference between weekdays and weekends

Inconclusive: does not show increase in upload volumes due to video conferencing



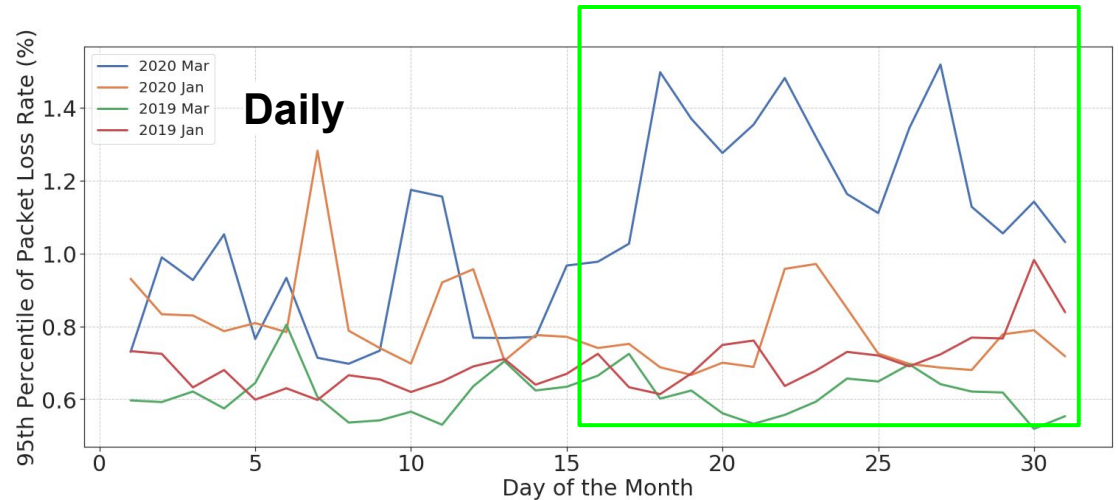
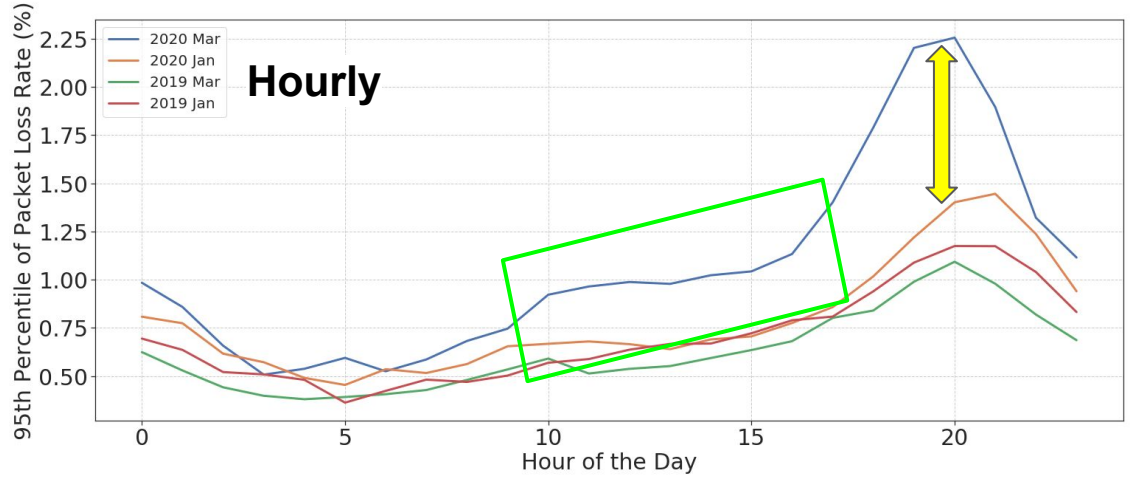
# Packet Loss Analysis

Tested in periods of no user activity (no user congestion)

Focuses test unit subset with worst packet loss (worst 5%)

1 percentage point increase starting from mid-March

Indicates ISP network congestion in 2nd half of March



# Summary (more in the paper)

- Data usage changes consistent with lockdown timeline
- Weekday daytime traffic pattern -> weekend pattern (normal in June 2020)
- Later (and smaller) spikes in rural areas compared to metro areas
- ISP network congestion in the second half of March
- 5% decrease in average speed for 60-70% of test units (in the paper)

## **Some Limitations**

- Older (2018) test unit profile data (not all test units included)
- Possible sampling bias (MBA is a volunteer-based program)
- Lack of application-level traffic information (no activity-based disaggregation)



# Suggestions for Future Improvements

## Datasets

- More frequent publication of MBA raw data (monthly?)
- Publish up-to-date unit profile file (not ISP cross-validated?)
- Correlate # of test units in ISP service class with the class' popularity?
  - could be fixed by ISPs

## Methodology

- Disaggregate by network type (DSL?)
- Correlate with other data (mobile networks, service providers)
- Eliminate sampling bias

*Acknowledgment: Craig Newmark Philanthropies, DIMACS REU 2020 summer research program*