



# EdgeDroid

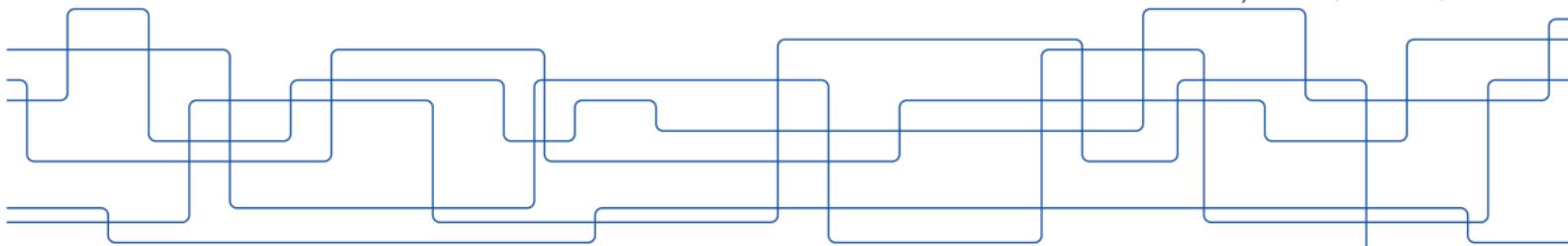
*An Experimental Approach to Benchmarking Human-in-the-Loop Applications*

M. Olguín Muñoz<sup>†</sup>, J. Wang<sup>‡</sup>, M. Satyanarayanan<sup>‡</sup> and J. Gross<sup>†</sup>

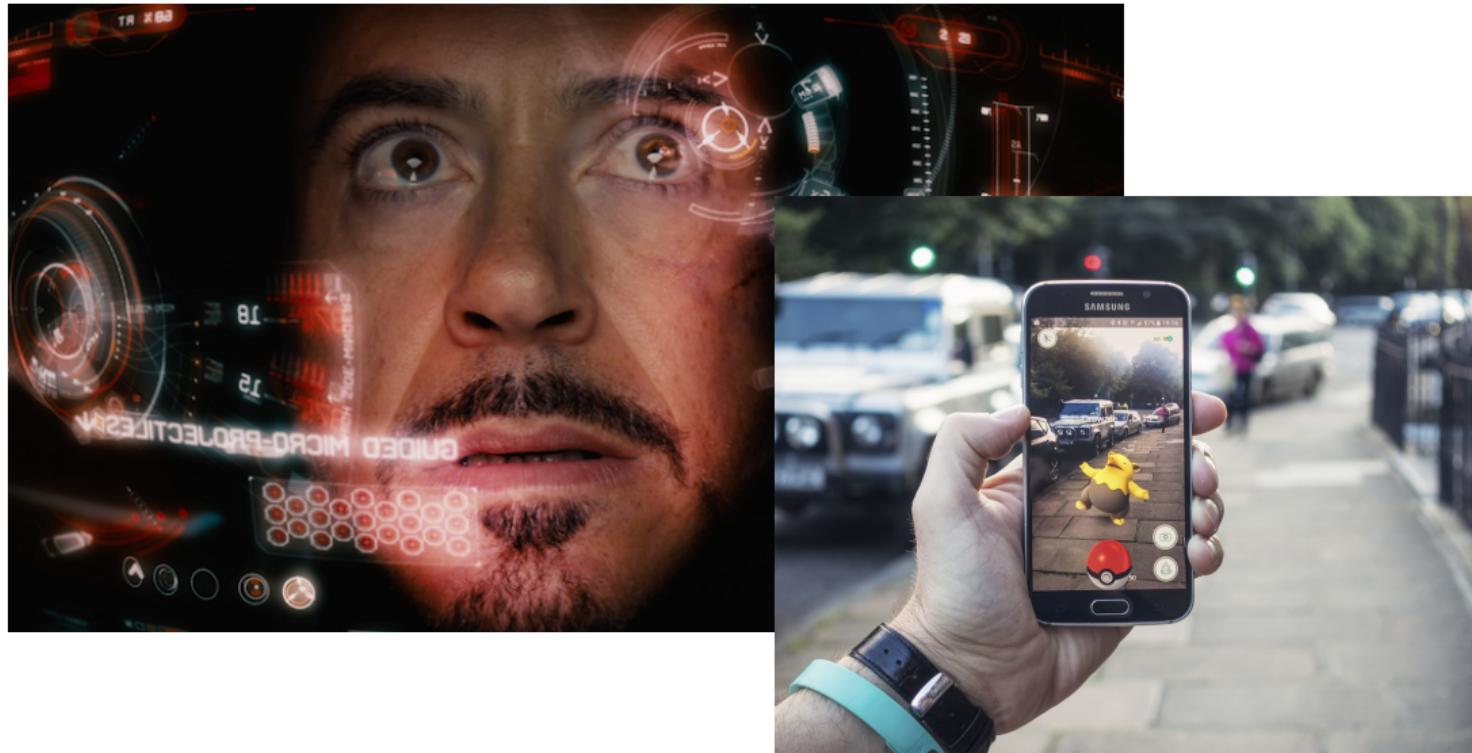
<sup>†</sup> KTH Royal Institute of Technology

<sup>‡</sup> Carnegie Mellon University

*HotMobile'19 Session 5: February 28<sup>th</sup> 2019, Santa Cruz, CA*



# Human-in-the-Loop Applications

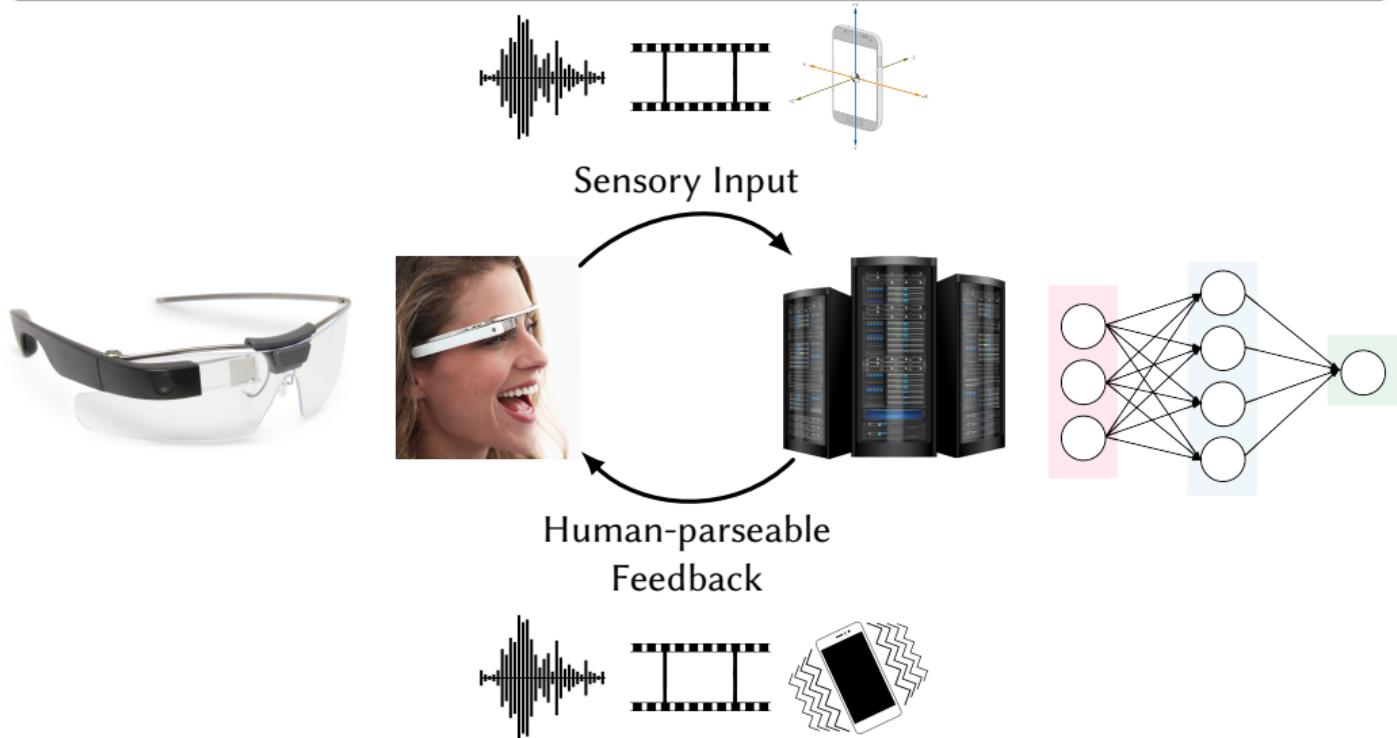


# Cognitive Assistance

Characterize Cognitive Assistance and Task Guidance?

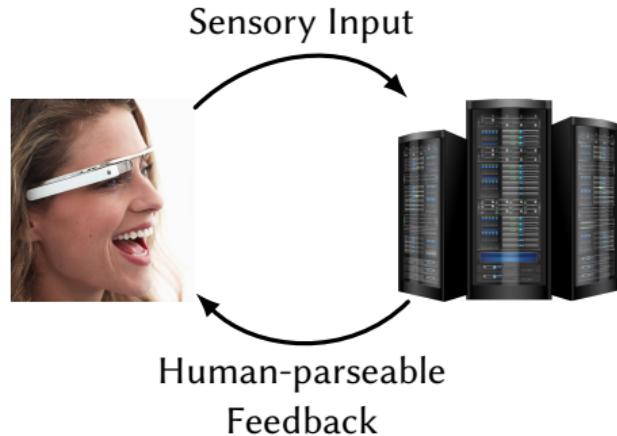
# Cognitive Assistance

Move this?



# Scaling

Add images of more users



Human users make it hard to study scaling behavior.

# Our Contributions

## Our Contributions

- ▶ A methodology for benchmarking human-in-the-loop applications.

# Our Contributions

- ▶ A methodology for benchmarking human-in-the-loop applications.
- ▶ EdgeDroid: A benchmarking tool-suite.

# Our Contributions

- ▶ A methodology for benchmarking human-in-the-loop applications.
- ▶ EdgeDroid: A benchmarking tool-suite.
- ▶ A set of experiments and measurements which show the effectiveness of the approach.

# Outline

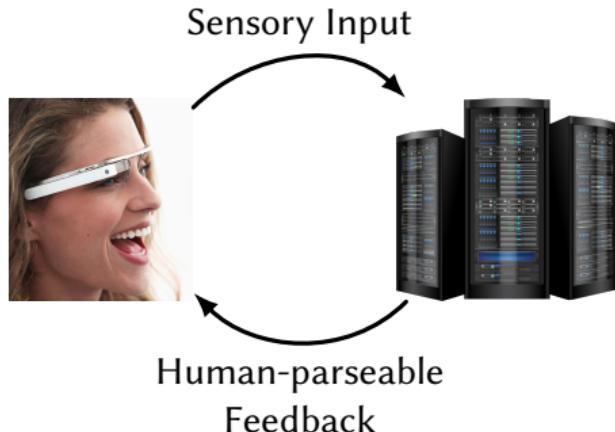
- Introduction
- **Background**
  - Previous & Related Work
  - Motivation
- EdgeDroid: Experimentally Benchmarking Human-in-the-Loop
- Conclusions

## Previous & Related Work

References to previous work. Mention wearable cognitive assistance is prime candidate for Edge Computing

# Motivation

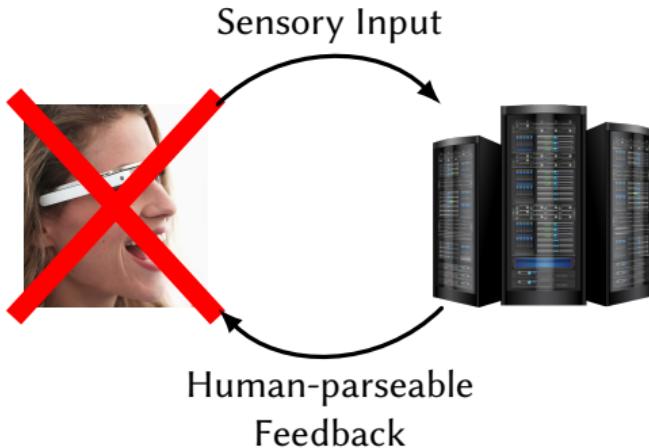
Add more users!



Benchmarking human-in-the-loop applications is HARD!

# Motivation

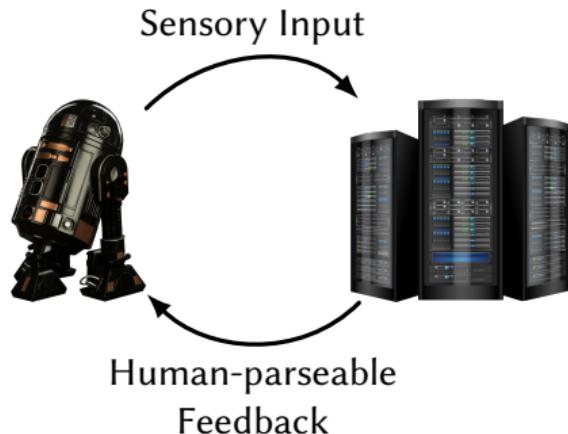
Add more users!



What if we could do away with the human users?

# Motivation

Add more users!



What if we could do away with the human users?

# Outline

- Introduction
- Background
- **EdgeDroid: Experimentally Benchmarking Human-in-the-Loop**
  - Requirements
  - Implementation
  - Evaluation
- Conclusions

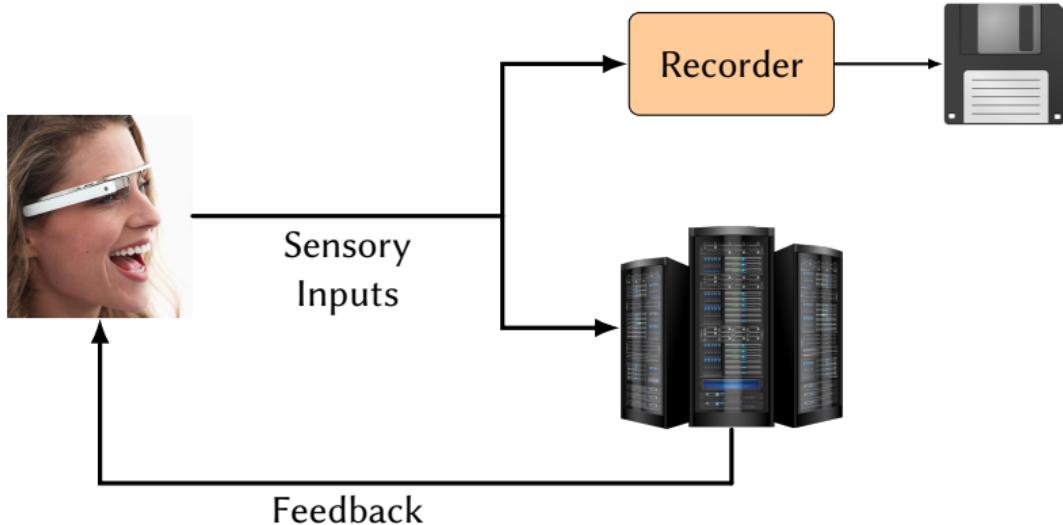
# Requirements

- ▶ Generate realistic, high-dimensional, real-time inputs.
- ▶ Correctly and realistically react to feedback.



# Input Generation

- ▶ Trace of human-generated inputs.



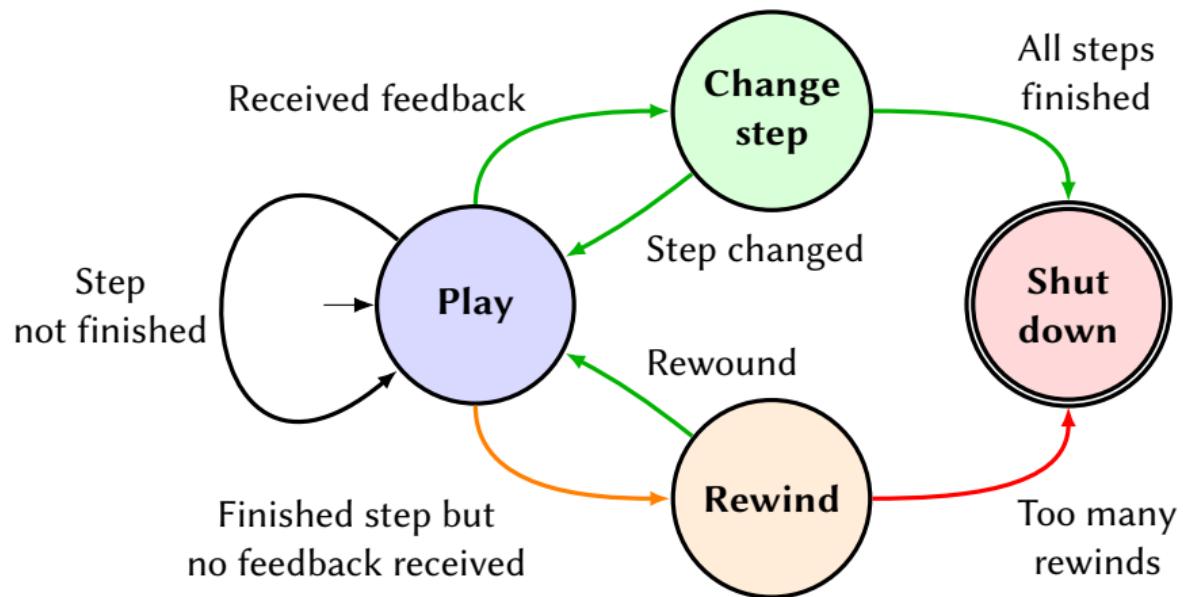
Add another figure showing usage of trace?

## Trace-based approach: Issues

Mention problems with pure trace-based approach: timing issues, unreliability of CV

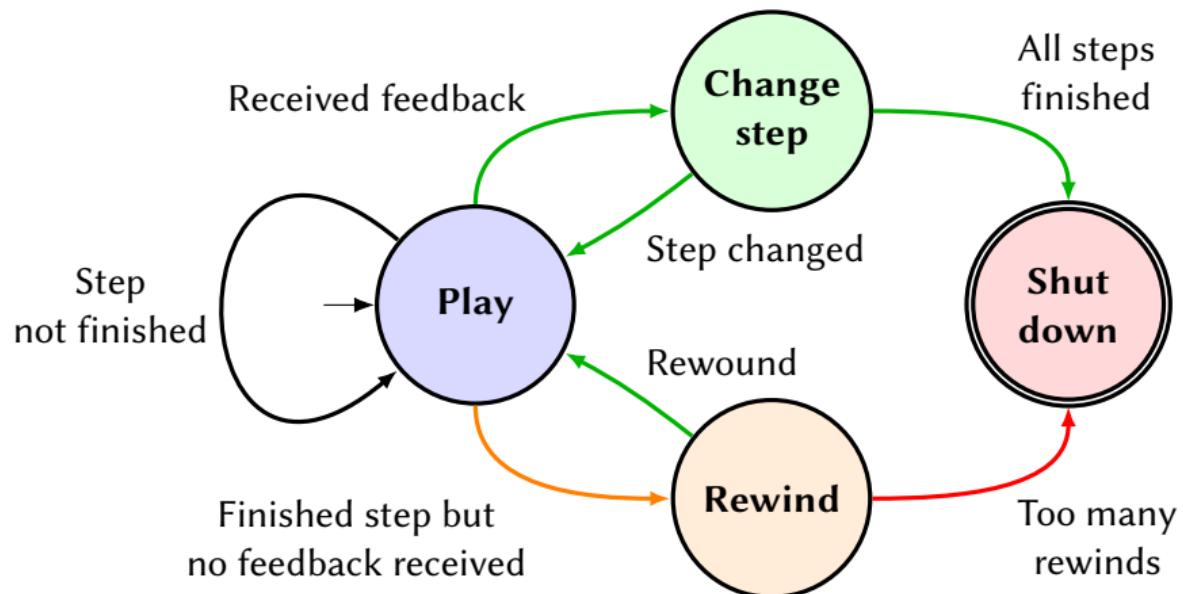
# Reaction to Feedback: User Model

- ▶ Model of human interaction.



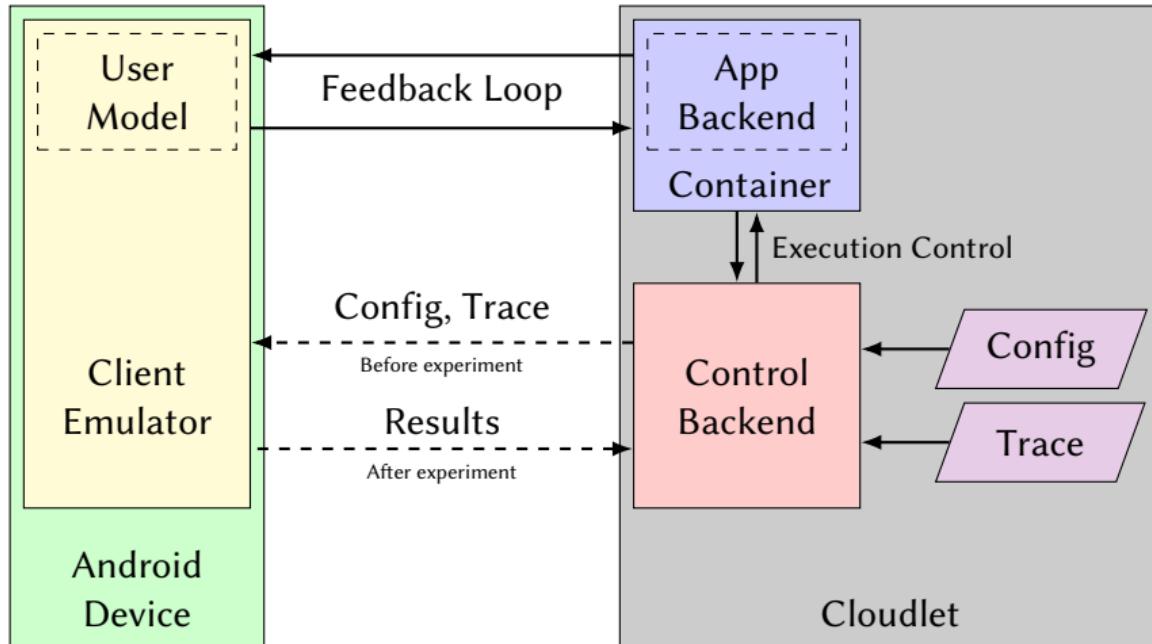
# Reaction to Feedback: User Model

- ▶ Model of human interaction.

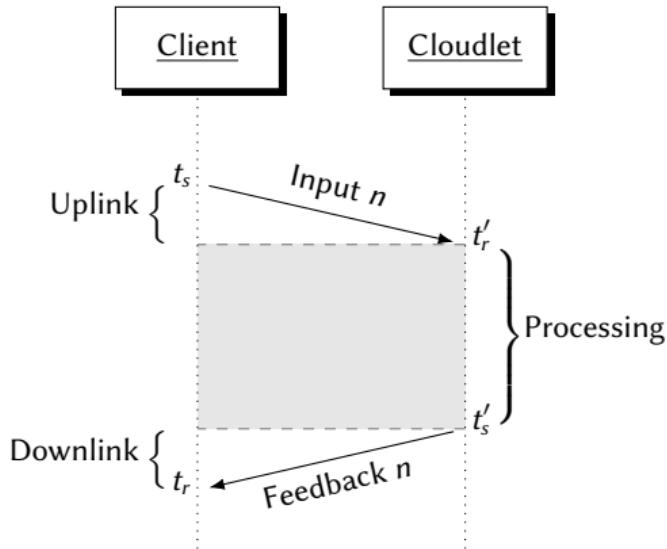


Currently working on a more thorough characterization of human behavior.

# Implementation



# Timestamping



Clocks are synchronized previous to the experiment.

Timestamps at key points to obtain:

$$\Delta T_{\text{up}} = t'_r - t_s \quad (1)$$

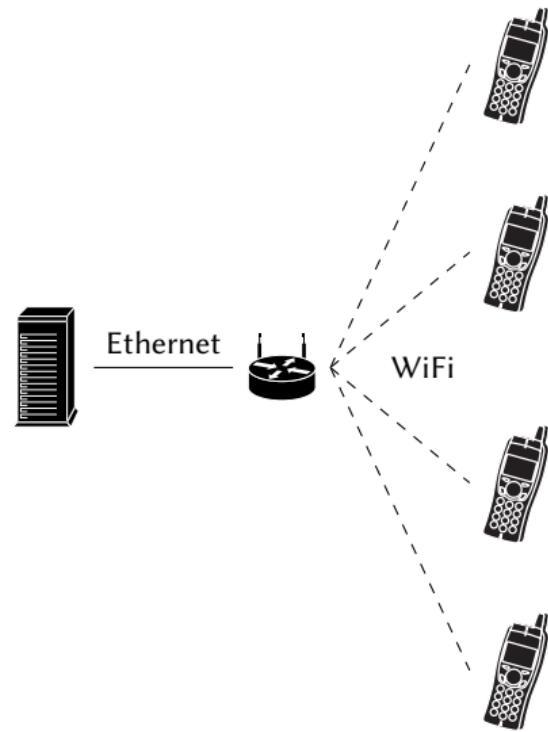
$$\Delta T_{\text{proc}} = t'_s - t'_r \quad (2)$$

$$\Delta T_{\text{down}} = t_r - t'_s \quad (3)$$

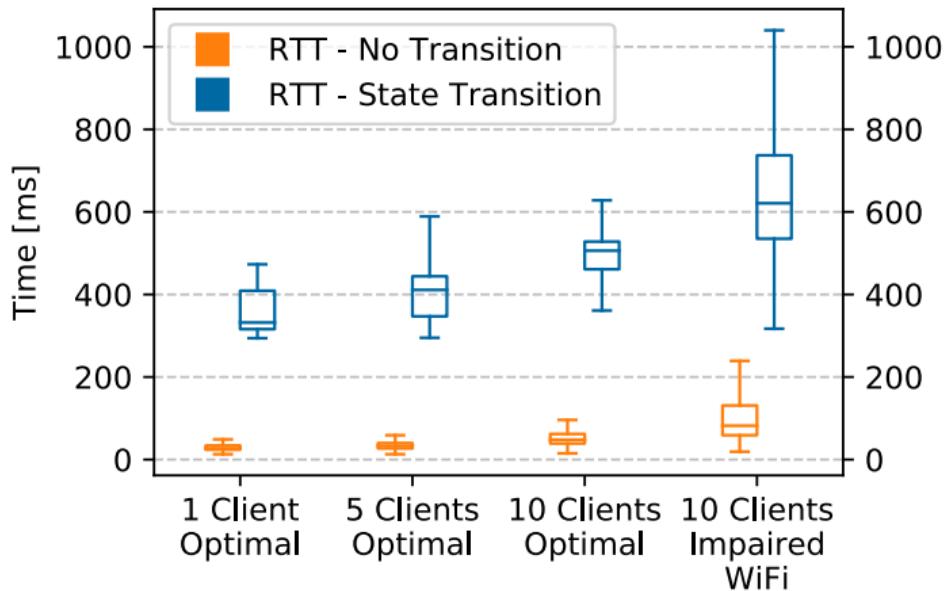
$$\Delta T_{\text{total}} = \Delta T_{\text{up}} + \Delta T_{\text{proc}} + \Delta T_{\text{down}} = t_r - t_s \quad (4)$$

# Evaluation

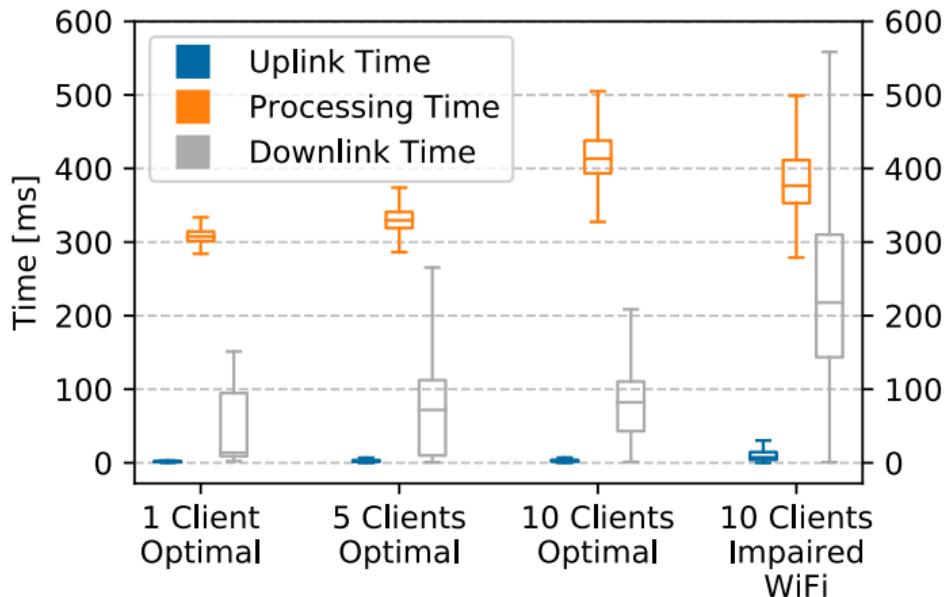
Insert pictures of LEGO Assistant



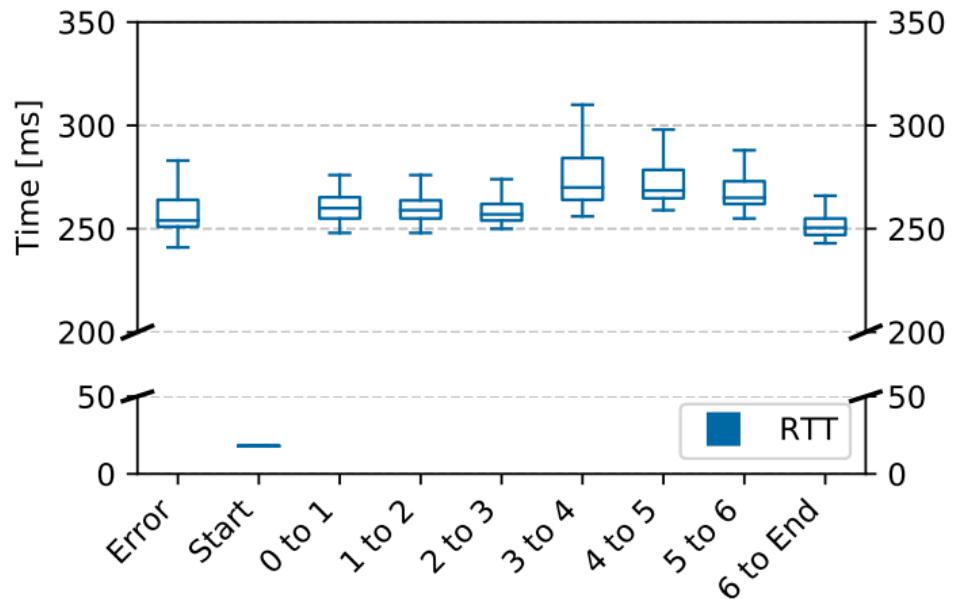
# Results



# Results



# Results



# Outline

- Introduction
- Background
- EdgeDroid: Experimentally Benchmarking Human-in-the-Loop
- **Conclusions**
  - Future Work
  - Summary



## Future Work

- ▶ More accurate user model.
- ▶ Expand to other types of Applications.

# Summary

- ▶ There's a need to study the scaling of Human-in-the-Loop applications.
  - ▶ This is difficult due to human users.
- ▶ We present a methodology + tool suite for benchmarking:
  - ▶ **EdgeDroid**
  - ▶ Trace based.
  - ▶ Model of human behavior.
- ▶ We present results which show the utility of EdgeDroid.