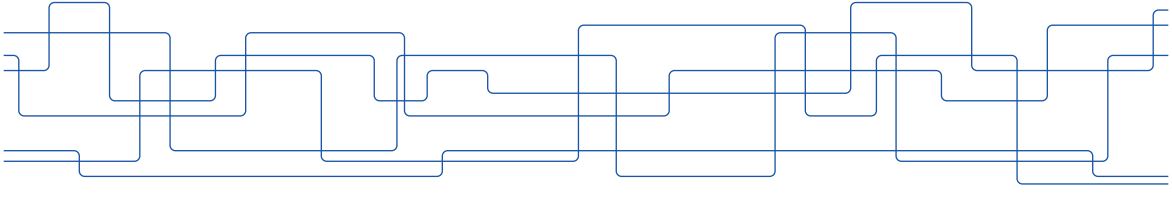
# EdgeDroid

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

M. Olguín Muñoz[†], J. Wang[‡], M. Satyanarayanan[‡] and J. Gross[†]

[†] KTH Royal Institute of Technology
[‡] Carnegie Mellon University

*HotMobile'19 Session 5: February 28th 2019, Santa Cruz, CA*




65
126ft
NE
354


SAMSUNG


Sensory Input


Human-parseable
Feedback

# Human-in-the-Loop Applications

- Application Developers
- Infrastructure Providers
- Researchers

Too much text?

# Human-in-the-Loop Applications

- Application Developers
  - Debugging
  - Resource Consumption
  - Performance and Optimization
- Infrastructure Providers
- Researchers

Too much text?

# Human-in-the-Loop Applications

- Application Developers
  - Debugging
  - Resource Consumption
  - Performance and Optimization
- Infrastructure Providers
  - Performance and Optimization
  - Scaling
- Researchers

Too much text?

# Human-in-the-Loop Applications

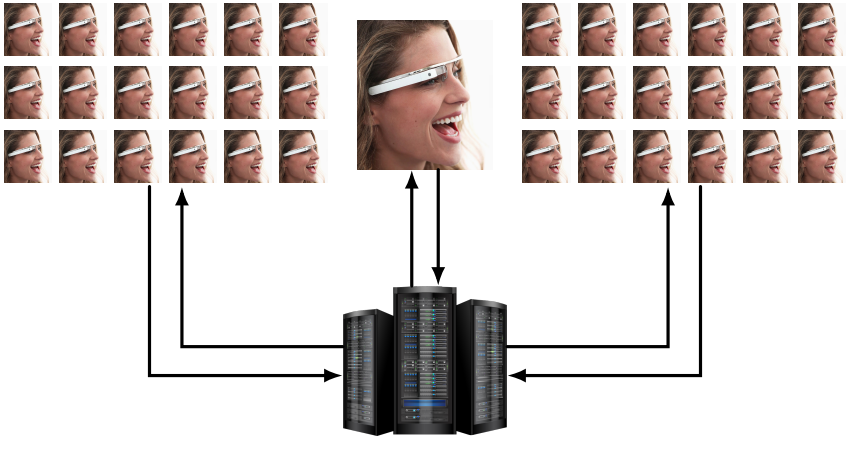
- Application Developers
  - Debugging
  - Resource Consumption
  - Performance and Optimization
- Infrastructure Providers
  - Performance and Optimization
  - Scaling
- Researchers

**How to obtain these measurements?**

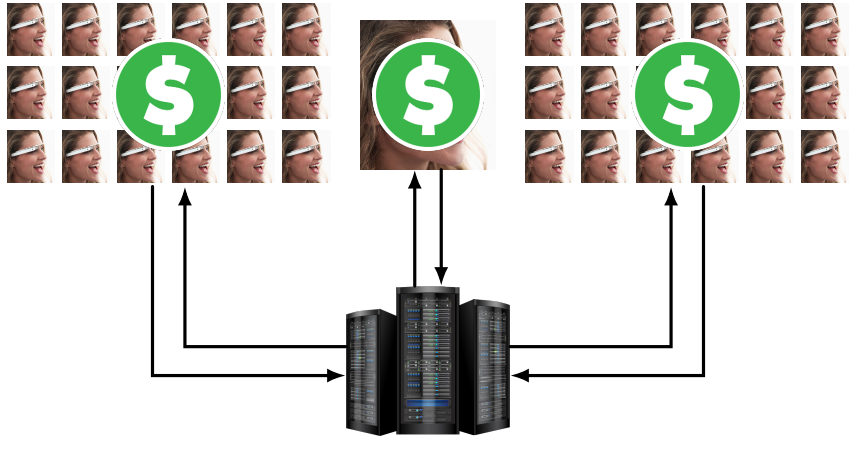
Too much text?

# Studying Human-in-the-Loop Applications

# Studying Human-in-the-Loop Applications

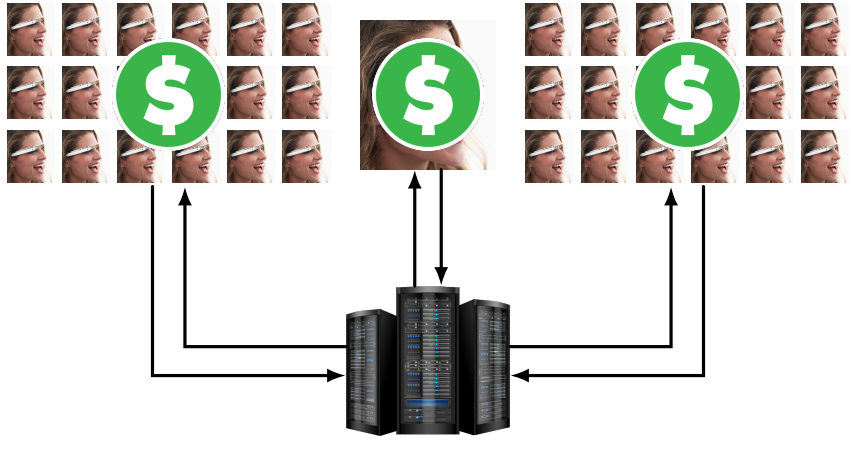


# Studying Human-in-the-Loop Applications



**Costly**

# Studying Human-in-the-Loop Applications



**Costly, poor repeatability**

# Previous & Related Work

split this up?

- "Towards Wearable Cognitive Assistance" (Ha *et al.* 2014)
- "Early Implementation Experience with Wearable Cognitive Assistance Applications" (Chen *et al.* 2015)
- "An Empirical Study of Latency in an Emerging Class of Edge Computing Applications for Wearable Cognitive Assistance" (Chen *et al.* 2017)

# Our Contributions

# Our Contributions

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

# Our Contributions

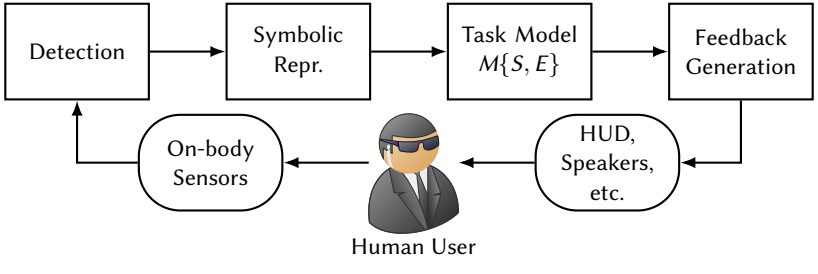
- A methodology for benchmarking human-in-the-loop applications.
- EdgeDroid: A benchmarking tool-suite.

# Our Contributions

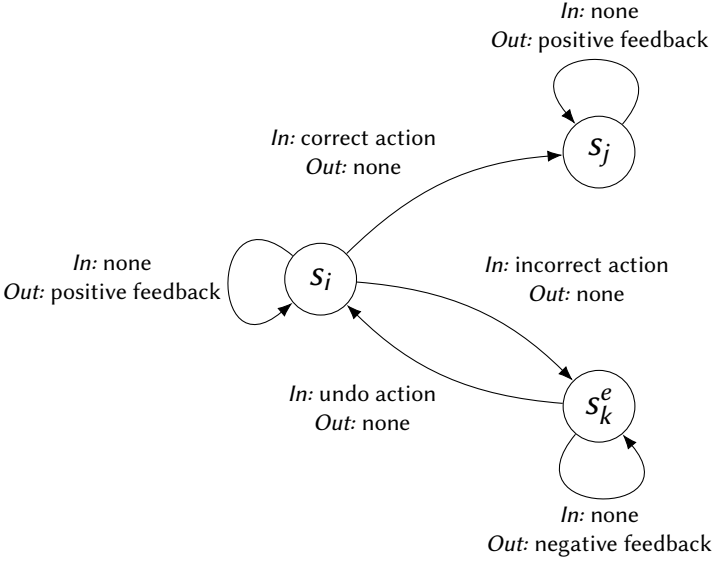
- A methodology for benchmarking human-in-the-loop applications.
- EdgeDroid: A benchmarking tool-suite.
- A set of experiments and measurements which show the effectiveness of the approach.

# Task-guidance Cognitive Assistance

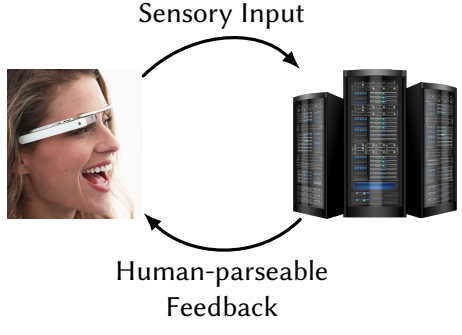
Some image of task guidance



# The Task Model

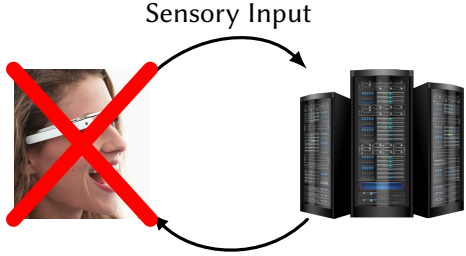


# Approach: Motivation



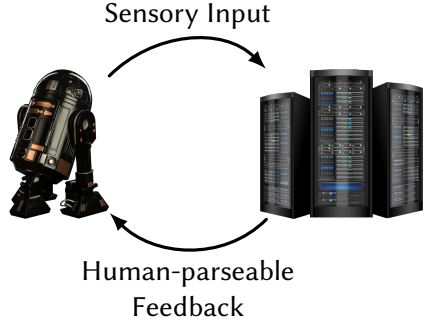

Benchmarking human-in-the-loop applications is HARD

# Approach: Motivation




What if we could do away with the human users?

# Approach: Motivation




What if we could do away with the human users?

# Requirements

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

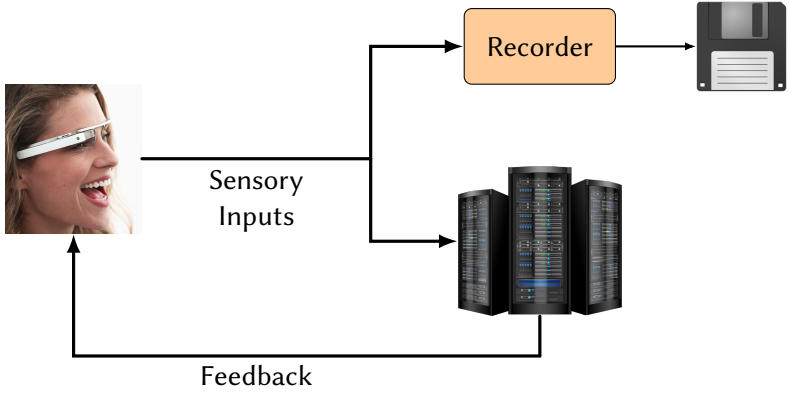
# Requirements

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

**Trace of pre-recorded inputs**

# Input Generation

- Trace of human-generated inputs.




Add another figure showing usage of trace?

# Trace-based approach: Issues

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

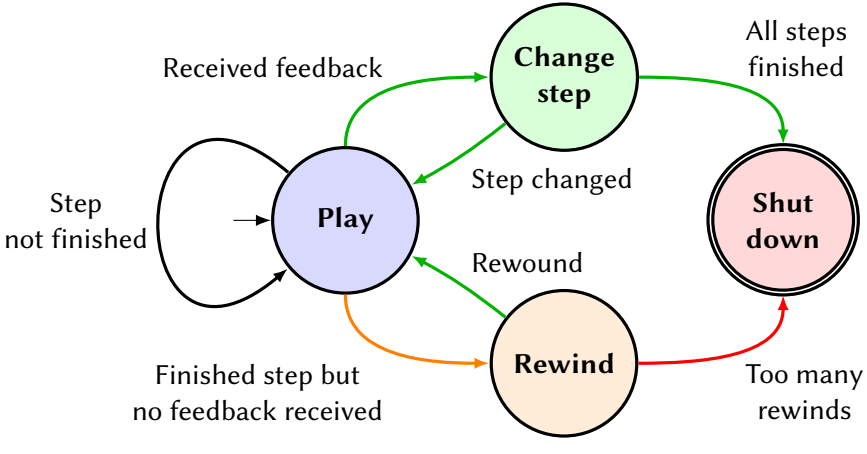
# Requirements

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

**Trace of pre-recorded inputs**
**& a model of user behavior**

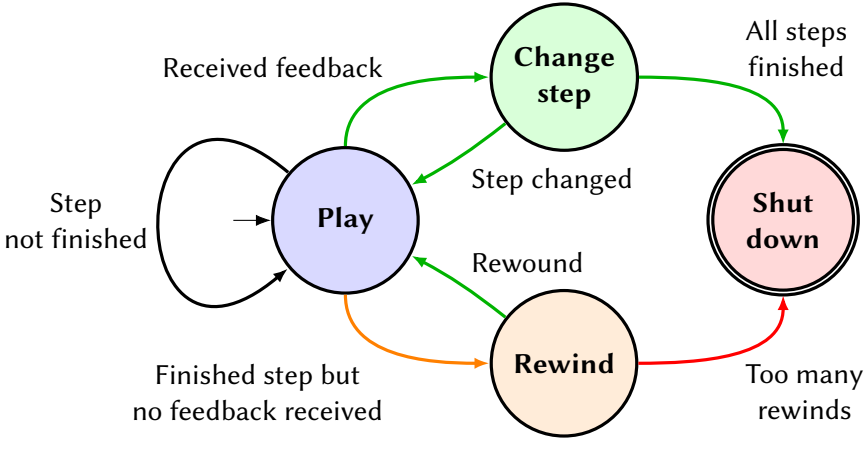
# Reaction to Feedback: User Model

- Model of human interaction.



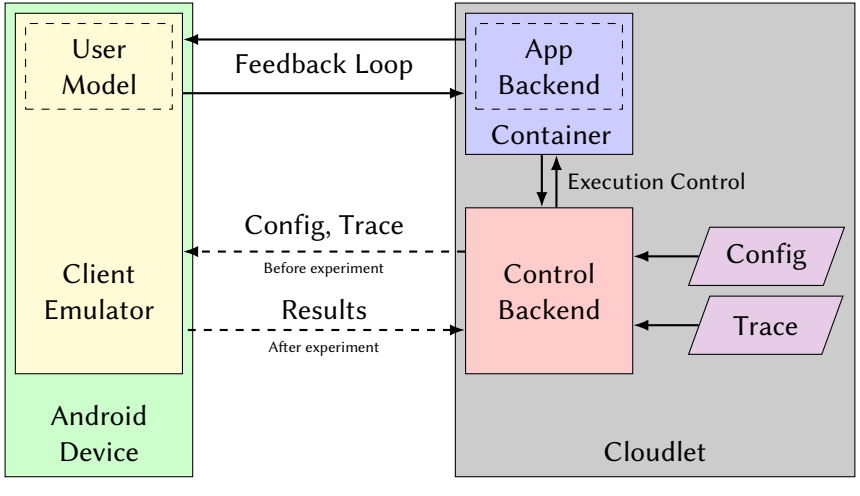
# Reaction to Feedback: User Model

- Model of human interaction.

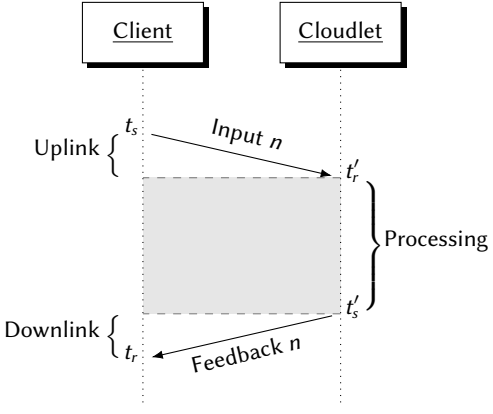



Currently working on a more thorough characterization of human behavior.

# Implementation



# Timestamping




Clocks are synchronized previous to the experiment.

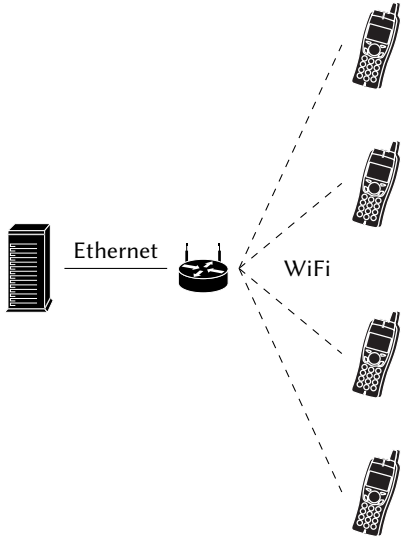
Timestamps at key points to obtain:

$$\Delta T_{\text{up}} = t'_r - t_s \tag{1}$$
$$\Delta T_{\text{proc}} = t'_s - t'_r \tag{2}$$
$$\Delta T_{\text{down}} = t_r - t'_s \tag{3}$$

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

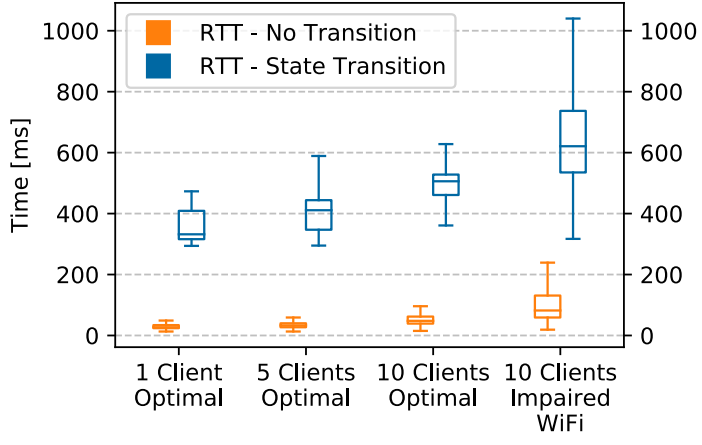
# Evaluation

Insert pictures of LEGO Assistant



# Results

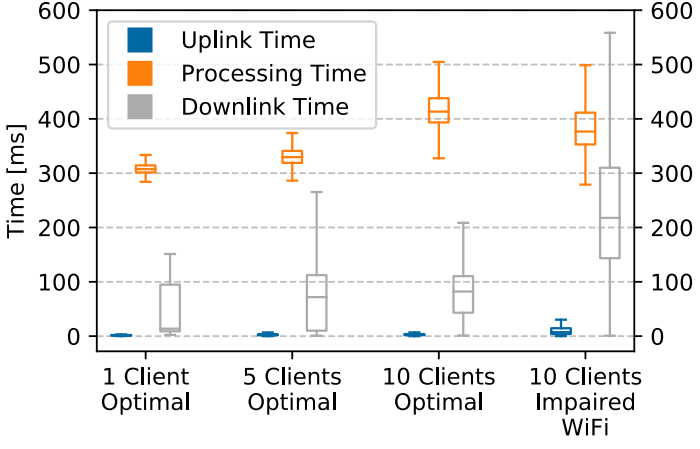
Figure labels




I haven't explained the task model, maybe skip this graph?

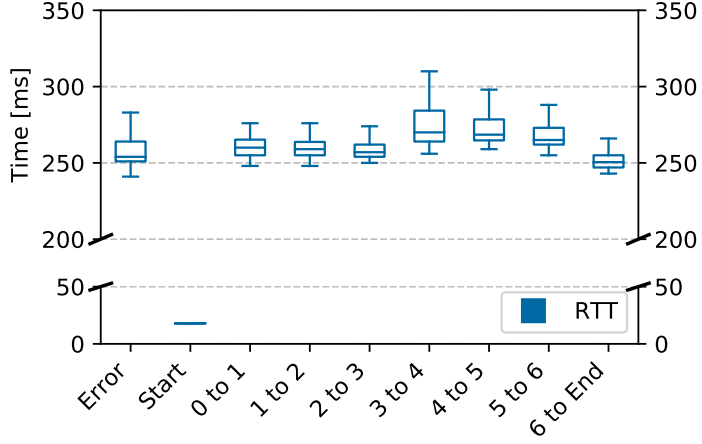
# Results

## Figure labels



# Results

Figure labels



# 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.