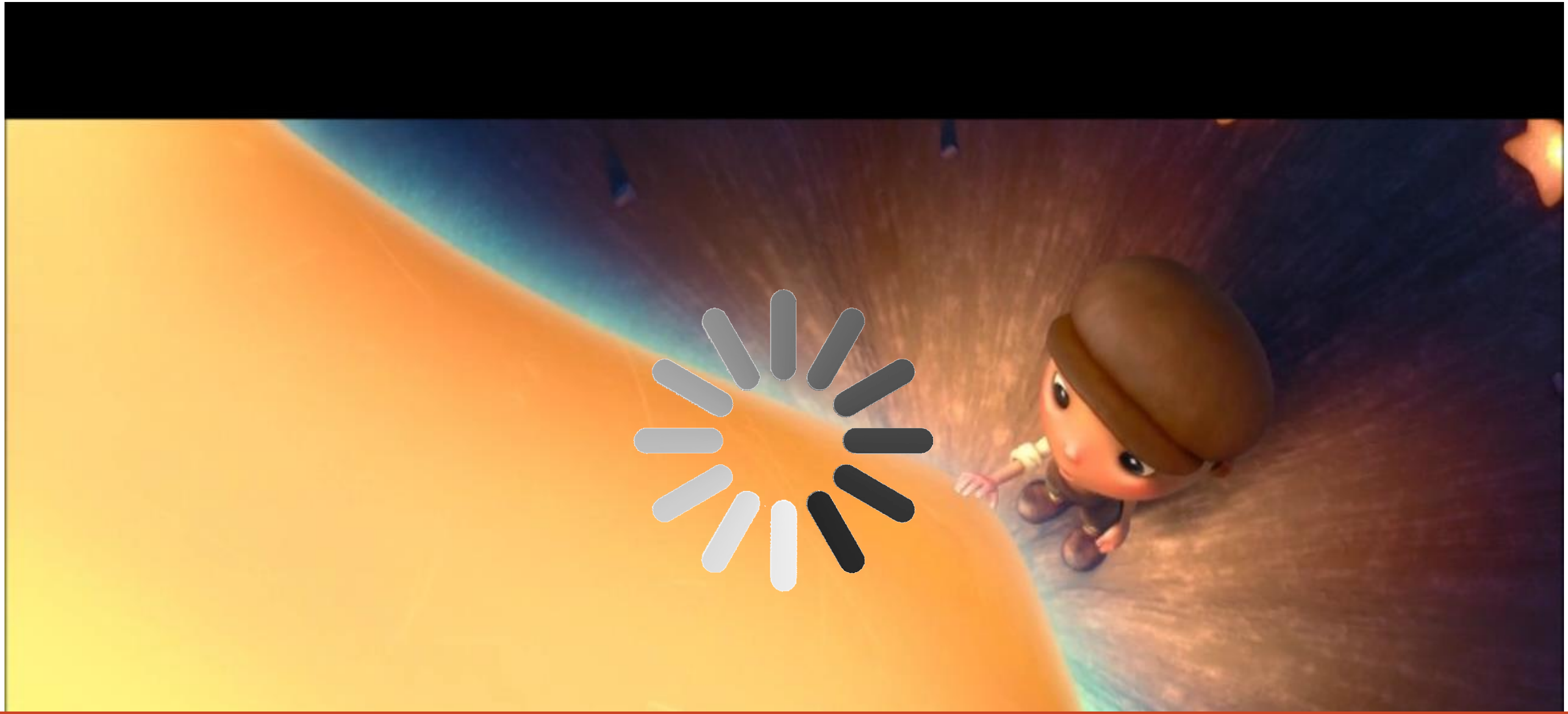
Neural Adaptive Video Streaming with Pensieve

Hongzi Mao

Ravi Netravali   Mohammad Alizadeh

Massachusetts Institute of Technology

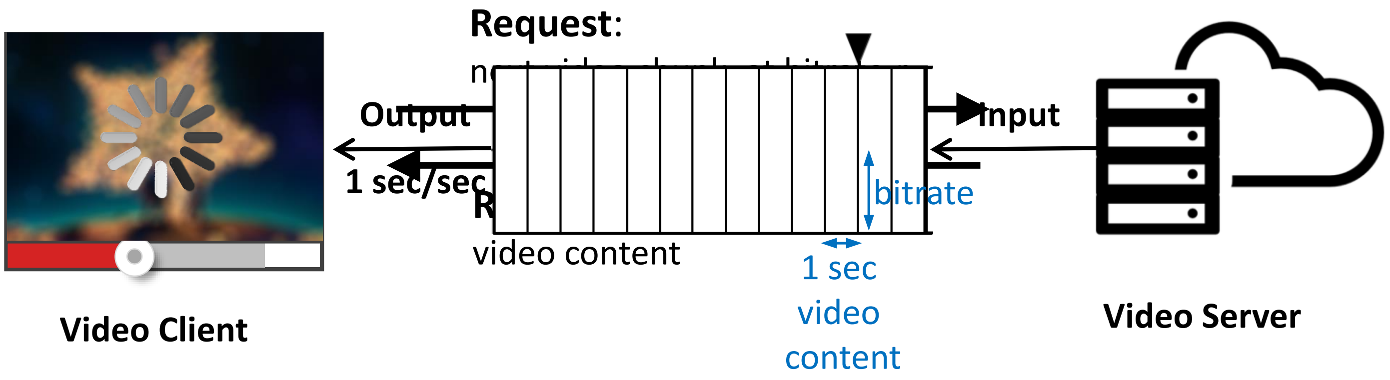
CSAIL



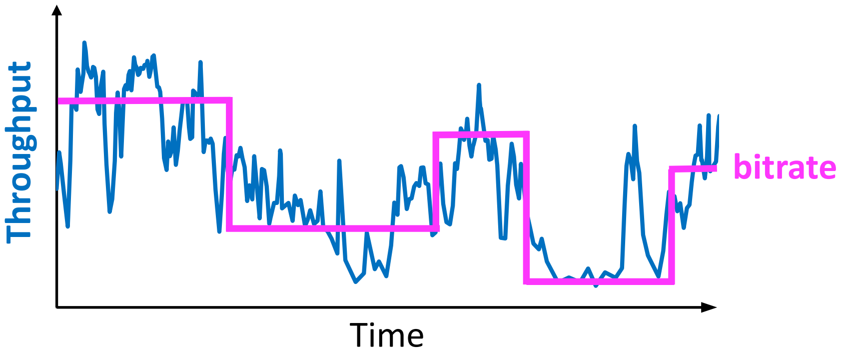
Users start leaving if video doesn’t play in 2 seconds

https://gigaom.com/2012/11/09/online-viewers-start-leaving-if-video-doesnt-play-in-2-seconds-says-study/
Video: La Luna (Pixar 2011)

Dynamic Streaming over HTTP (DASH)

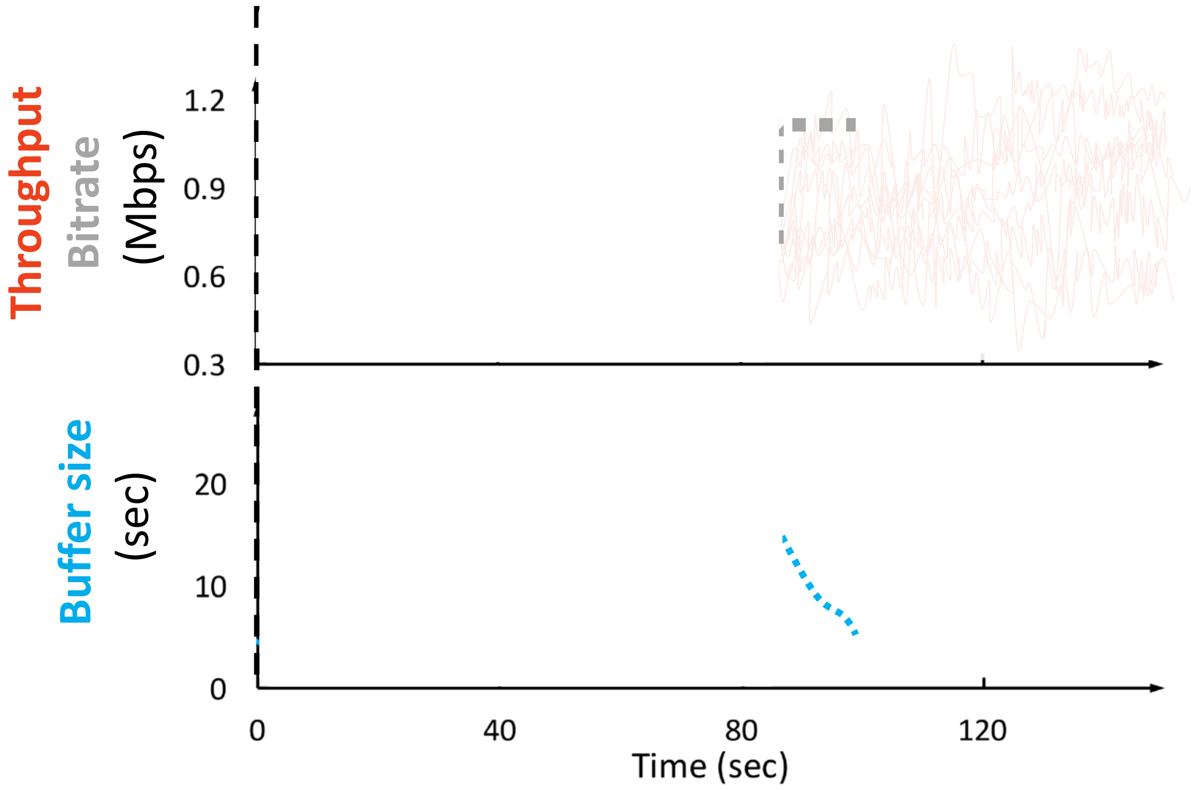


Adaptive Bitrate (ABR) Algorithms



Animation borrowed from Te-Yuan Huang (SIGCOMM '14) http://conferences.sigcomm.org/sigcomm/2014/doc/slides/38.pdf

Why is ABR Challenging?



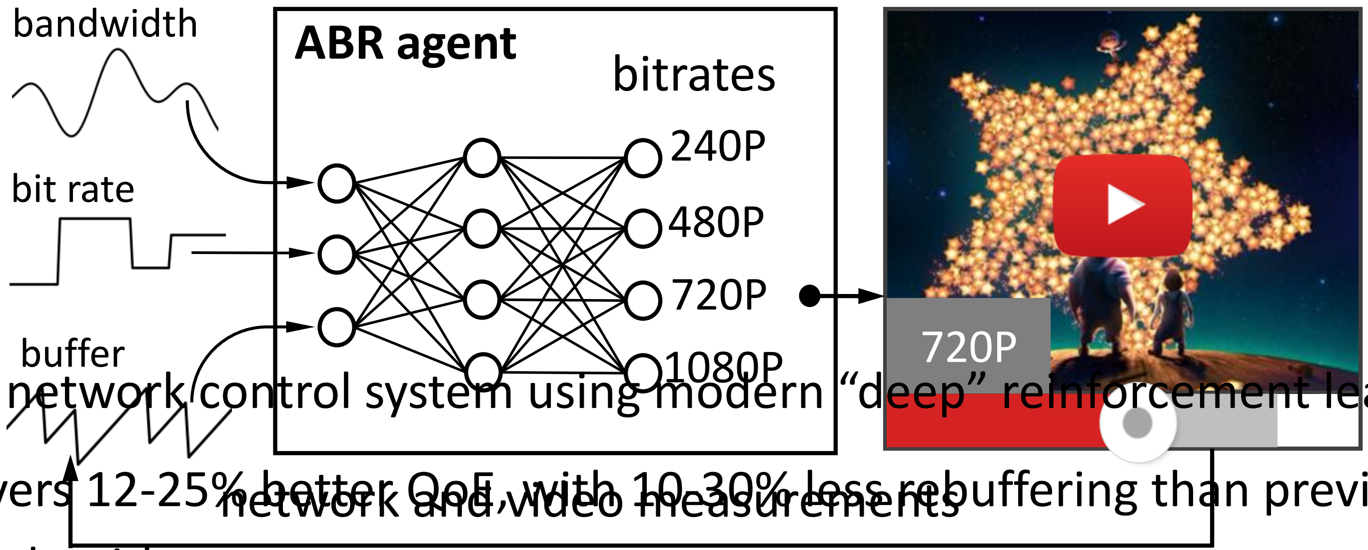
Network throughput is variable & uncertain

Conflicting QoE goals

- Bitrate
- Rebuffering time
- Smoothness

Cascading effects of decisions

Our Contribution: Pensieve



1. First network control system using modern "deep" reinforcement learning
2. Delivers 12-25% better QoE, with 10-30% less rebuffering than previous ABR algorithms
3. Tailors ABR decisions for different network conditions in a data-driven way

Pensieve learns ABR algorithm automatically through experience

Previous Fixed ABR Algorithms

- Rate-based: pick bitrate based on predicted throughput
  - FESTIVE [CoNEXT’12], PANDA [JSAC’14], CS2P [SIGCOMM’16]
- Buffer-based: pick bitrate based on buffer occupancy
  - BBA [SIGCOMM’14], BOLA [INFOCOM’16]
- Hybrid: use both throughput prediction & buffer occupancy
  - PBA [HotMobile’15], MPC [SIGCOMM’15]

Simplified inaccurate model leads to suboptimal performance

Example: Model Predictive Control

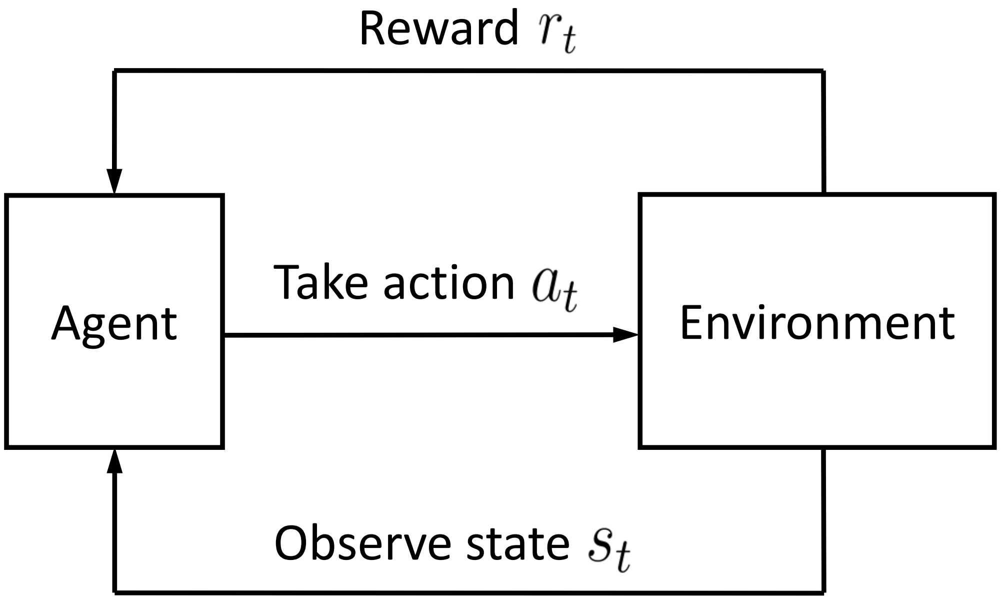
maximize QoE (t, t + T)
subject to system dynamics

Prediction

Solution: learn from video streaming sessions in actual network conditions

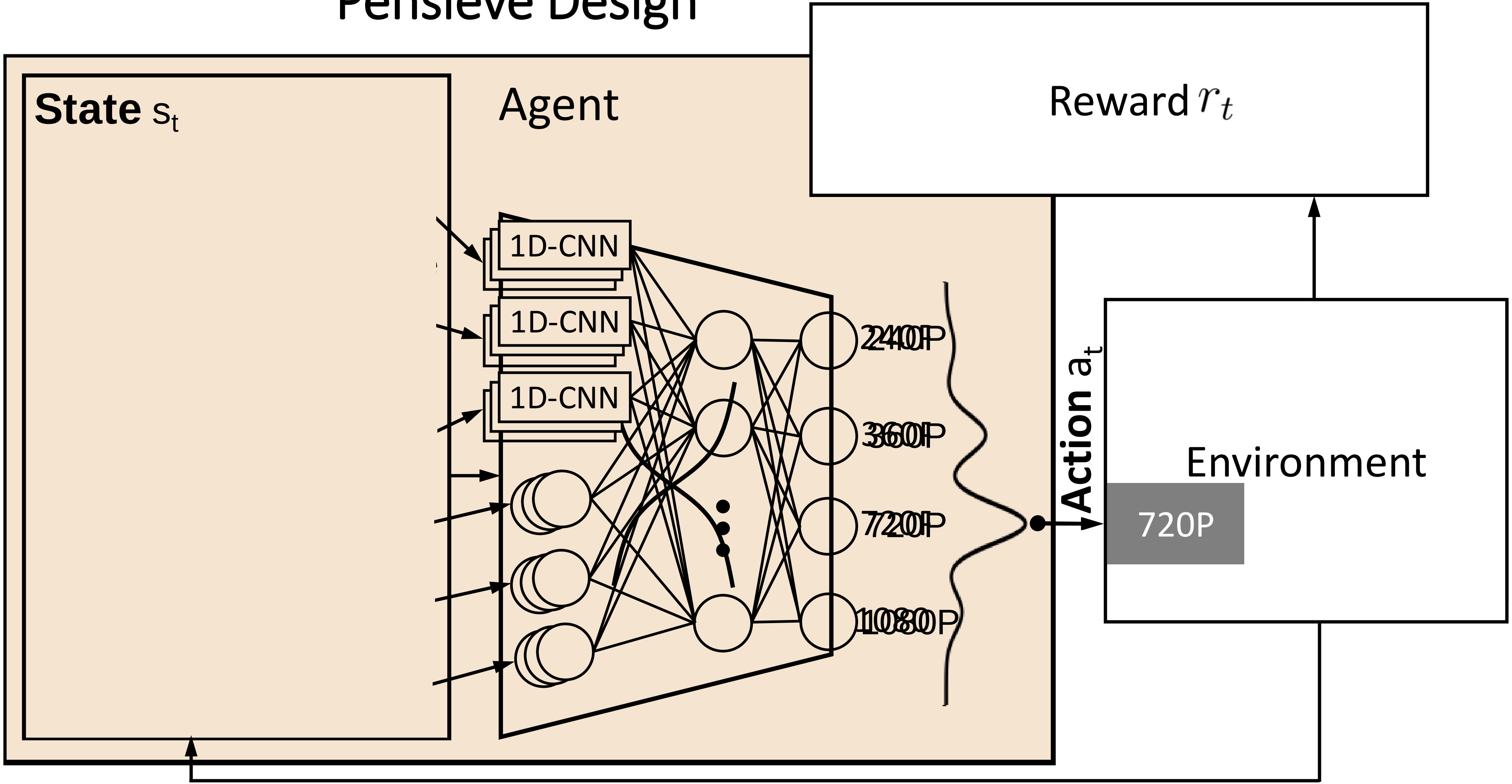
Problem: Needs accurate throughput model

Reinforcement Learning

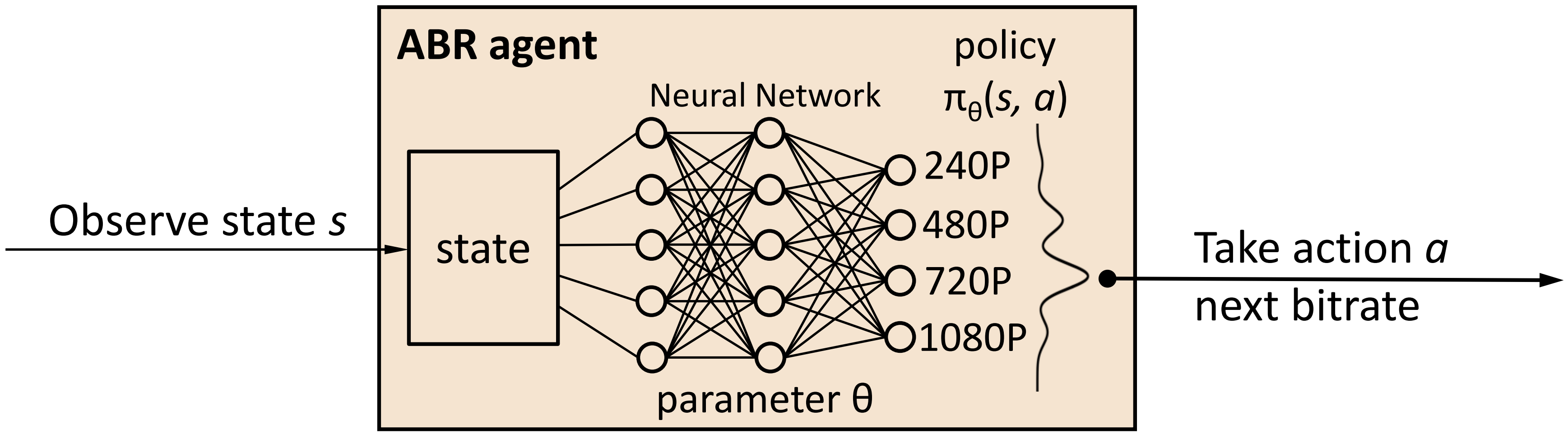


Goal: maximize the cumulative reward $\sum_t r_t$

Pensieve Design



How to Train the ABR Agent



Collect experience data: trajectory of [state, action, reward]

Training: $$\theta \leftarrow \theta + \alpha \nabla_\theta \mathbb{E}_{\pi_\theta}\left[\sum_t r_t\right]$$

estimate from empirical data

What Pensieve is good at

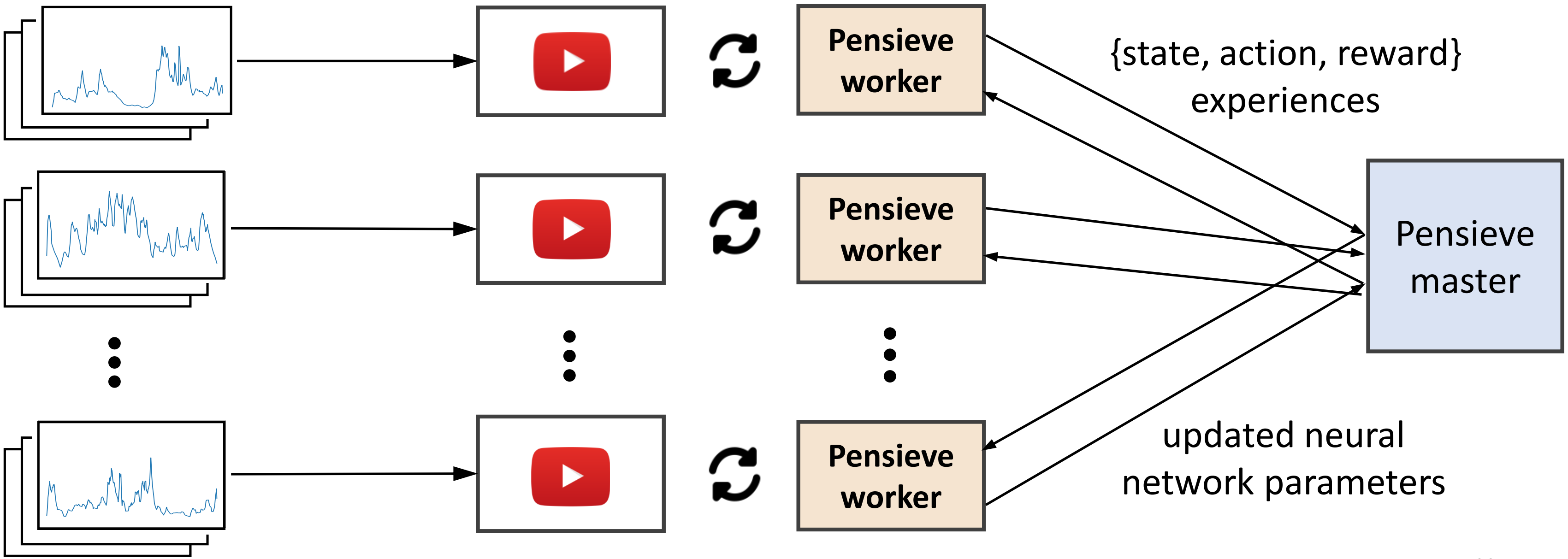
- Learn the dynamics **directly from experience**
- Optimize the high level QoE objective **end-to-end**
- Extract control rules from **raw high-dimensional** signals

Pensieve Training System

Large corpus of network traces
cellular, broadband, synthetic

Video playback
Fast chunk-level simulator

Model update
TensorFlow



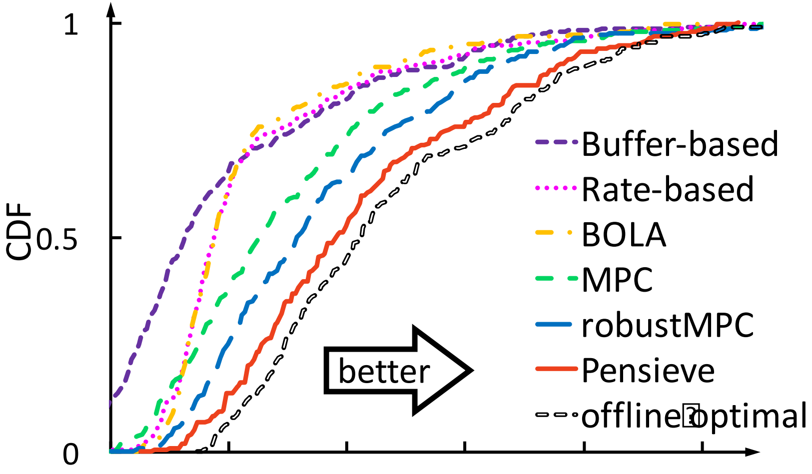
Demo

Pensieve
240P
MPC
240P
Pensieve
buffer (sec)
20
10
0
Throughput
(mbps)
2.0
1.5
1.0
0.5
chances of outage
MPC
buffer (sec)
20
10
0
Rebuffering
Time

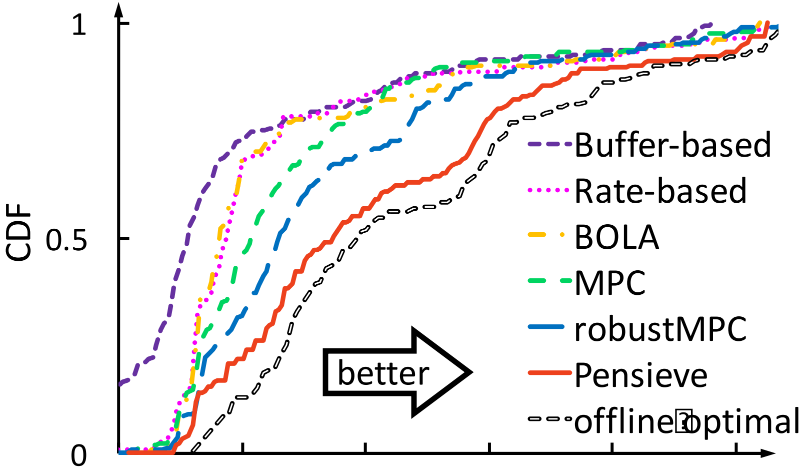
Trace-driven Evaluation

- **Dataset:** Two datasets, each dataset consists of 1000 traces, each trace 320 seconds.
- **Video:** 193 seconds. encoded at bitrates: {300, 750, 1200, 1850, 2850, 4300} kbps.
- **Video player:** Google Chrome browser **Video server:** Apache server

Norway 3G cellular dataset

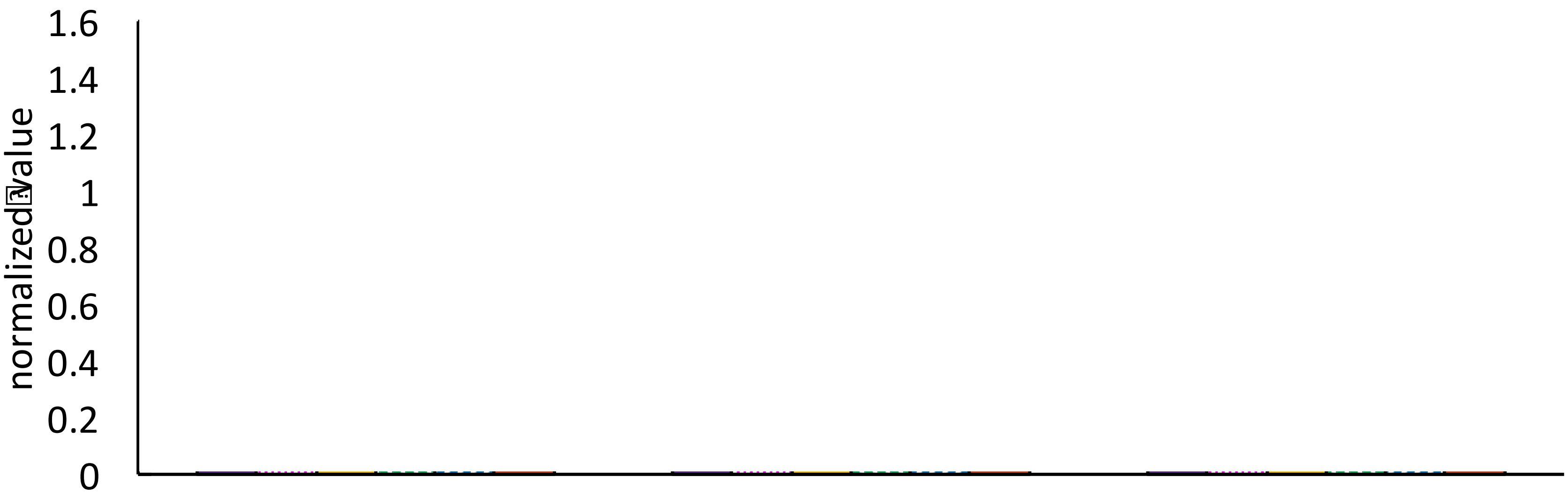


FCC broadband dataset



Pensieve improves the best previous scheme by 12-25% and is within 9-14% of the offline optimal

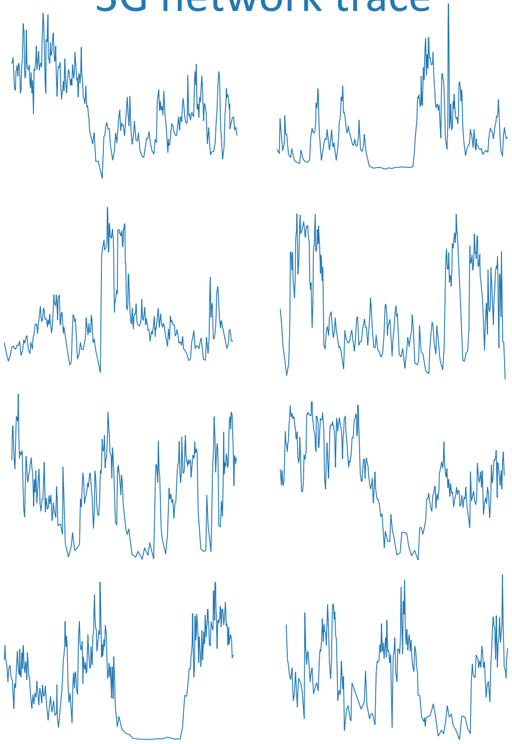
QoE Breakdown



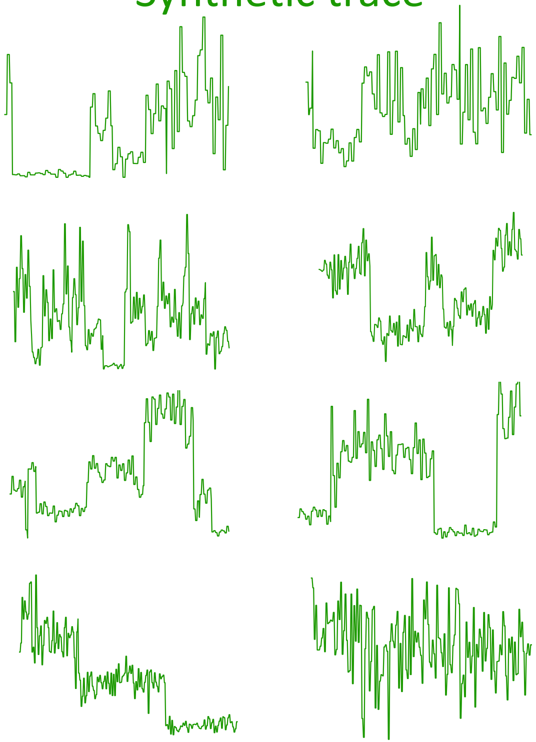
Pensieve reduces rebuffering by 10-32% over second best algorithm

Does Pensieve Generalize?

3G network trace

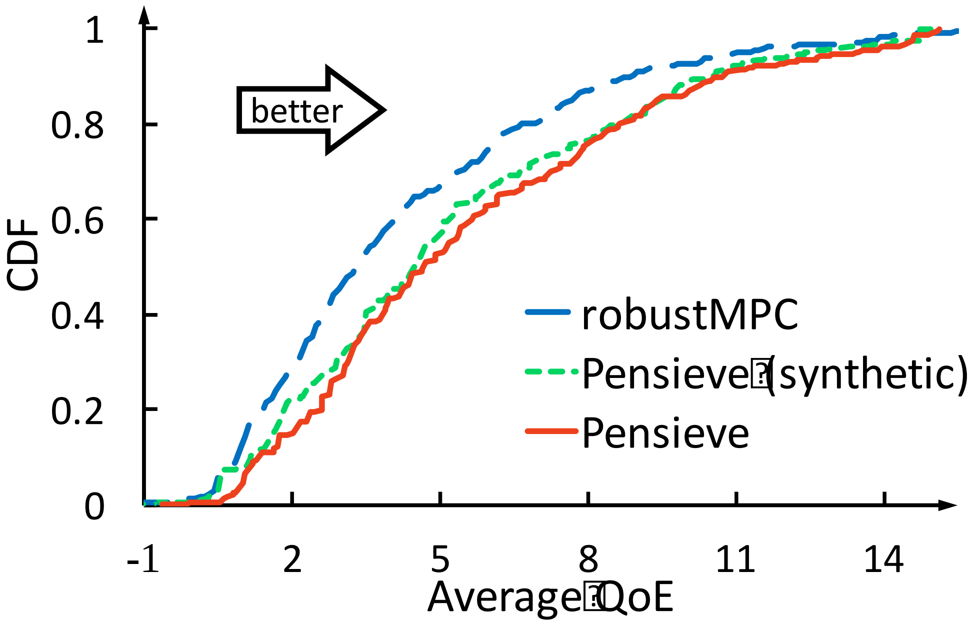


Synthetic trace



- Trace generated from a Hidden Markov model
- Covers a wide range of average throughput and network variation

Does Pensieve Generalize?



Train on synthetic traces then test on real 3G network trace

Only 5% degradation compared with Pensieve trained on real network trace

Other Evaluations

- Experiments in the wild (LTE, public WiFi, international link)
- Controlled experiment for testing optimality
- Multi-video extension
- Sensitivity analysis

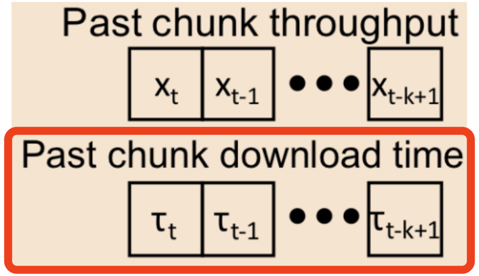
Lessons We Learned

1. Build a fast experimentation/simulation platform

Coarse-grain chunk simulator

2. Data diversity is more important than "accuracy"

3. Think carefully about controller state space (observation signals)
    - Too large a state space ⟶ slow & difficult learning
    - Too small a state space ⟶ loss of information
    - ⟶ When in doubt, include rather than cut the signal



Summary

- Pensieve uses Reinforcement Learning to generate ABR algorithms
- Pensieve optimizes different network conditions through experience
- Pensieve outperforms existing approaches across a wide range of network environments and QoE preferences
- Policies generated by Pensieve have strong ability to generalize

http://web.mit.edu/pensieve/