CS2P: Improving Video Bitrate Selection and Adaptation with Data-Driven Throughput Prediction

Y. Sun, F. Lin, N. Wang @ ICT
X. Yin, J. Jiang, V. Sekar, B. Sinopoli @ Carnegie Mellon University
T. Liu @ IQIYI
Bitrate adaptation is key for QoE

- DASH = Dynamic Adaptive Streaming over HTTP
- Entail new QoE metrics, e.g., low buffering, high video quality
- Need intelligent bitrate control and adaptation

Prior work: Accurate throughput prediction can help!
Accurate throughput prediction → Better initial bitrate selection

- Fixed bitrate
- Adaptive bitrate
Accurate throughput prediction → Better midstream adaptation

- Replicate the analysis by Yin et al. at SIGCOMM2015\[1\]

\[
\text{Normalized QoE} = \frac{\text{Actual QoE}}{\text{Theoretical optimal}}
\]


Open questions on predictability!

- Our understanding of throughput variability and predictability is quite limited.

- What types of prediction algorithms to use?
  - In the context of video bitrate adaptation
  - Prior approaches: 30%+ of predictions with error $\geq 0.2$
Our work and contributions

A large-scale analysis, providing data-driven insights for predicting the throughput accurately.


A practical implementation of CS2P and the demonstration of improvements in video QoE.
Outline

- Motivation

- Data-driven Observations

- CS2P Approach

- Evaluation
From operational platform of *iQIYI*.
iQIYI is a leading online video content provider in China.

**20M+** sessions, 8 days in Sep. 2015,
• Each session records avg. throughput per 6-second *epoch*.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client IP</td>
<td>3.2M</td>
</tr>
<tr>
<td>Client ISP</td>
<td>87</td>
</tr>
<tr>
<td>Client AS</td>
<td>161</td>
</tr>
<tr>
<td>Province</td>
<td>33</td>
</tr>
<tr>
<td>City</td>
<td>736</td>
</tr>
<tr>
<td>Server</td>
<td>18</td>
</tr>
</tbody>
</table>
Observation 1:
Significant variability within a session.

- 50% of sessions with $n$-stddev $\geq 30$
- 20% of sessions with $n$-stddev $\geq 50$

CDF

Normalized stddev

20% of sessions with $n$-stddev $\geq 50$
50% of sessions with $n$-stddev $\geq 30$
Observation 2: Stateful/persistent characteristics.

An example session:
Throughput variation across two consecutive epochs with a particular IP/16 prefix.
Observation 3: Similar session $\rightarrow$ Similar throughput

Throughput at different session clusters with particular IP/8 prefixes.
Observation 4: Complex relationship between session feature ↔ throughput

The impact of the same feature on different sessions could be variable.
Outline

- Motivation
- Data-driven Observations
- CS2P Approach
- Evaluation
<table>
<thead>
<tr>
<th>Observation</th>
<th>Idea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many sessions exhibit stateful characteristics in the evolution of the throughput.</td>
<td>CS2P learns Hidden-Markov Models (HMM) to capture the states and state transitions.</td>
</tr>
<tr>
<td>Sessions sharing similar critical characteristics tend to exhibit similar throughput patterns.</td>
<td>CS2P groups similar sessions sharing the same critical feature values and uses Cross-Session prediction methodology.</td>
</tr>
<tr>
<td>The relationship between session features and throughput are quite complex.</td>
<td>CS2P learns a separate model for each similar session clusters instead of using a global model.</td>
</tr>
</tbody>
</table>
Workflow of CS2P

1. Throughput Measurements

Step 1: Session Clustering

Step 2: Model Training

Prediction Engine

1. Initial Throughput
2. Prediction Model

Step 3: Throughput Prediction and Bitrate Selection

Video Server

Video Player Clients
Session clustering-finding critical features

All the sessions for training

- Session under prediction

A given subset of session features
Try another session feature subset

Repeat these procedures to find the critical feature set, which yields the most accurate throughput prediction of

Sessions matching selected features with

Predict the throughput of with these filtered sessions
Throughput prediction with HMM


Throughput prediction and bitrate selection (online).
Outline

• Motivation

• Data-driven Observations

• CS2P Approach

→ Evaluation
Trace-driven simulation setup

**Algorithms to compare:**

1. **History-based predictor:**
   - Last Sample, Harmonic-Mean, Auto Regression

2. **ML-based predictor:**
   - SVR, Gradient Boosting Regression trees

3. **CFA**[1]

**Bitrate selection method:**

- State-of-art: MPC[2]

**iQIYI throughput trace:**

- Non-overlapping traces of training and testing

**Video source:**

- “Envivio” from dash.js test website
- Encoded in H.264/MPEG-4 in 5 bitrate levels

---

Throughput Prediction Accuracy

Takeaway

- **Midstream Epoch**
  - Reduce median error by 50%

- **Multi-epoch Ahead**
  - 9% prediction error for 10 epoch ahead
  - 50% improvement

Midstream Throughput
Video QoE

- Normalized QoE = \frac{Actual QoE}{Theoretical\ optimal}

- QoE\[^1\] is a linear combination of avg. video quality, quality variation, total rebuffer time and startup delay.

Pilot deployment: multi-city test

<table>
<thead>
<tr>
<th>Metrics</th>
<th>vs. HM+MPC</th>
<th>vs. BB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Bitrate</td>
<td>10.9%</td>
<td>9.3%</td>
</tr>
<tr>
<td>Good Ratio</td>
<td>2.5%</td>
<td>17.6%</td>
</tr>
<tr>
<td>Bitrate Variability</td>
<td>-2.3%</td>
<td>5.6%</td>
</tr>
<tr>
<td>Startup Delay</td>
<td>0.4%</td>
<td>-3.0%</td>
</tr>
<tr>
<td>Overall QoE</td>
<td>3.2%</td>
<td>14.0%</td>
</tr>
</tbody>
</table>

**Takeaway:**
1. CS2P improves most of the QoE metrics, except longer startup delay than BB and higher bitrate variability than HM.
2. The overall QoE improvement of CS2P is 3.2% to HM and 14% to BB.
Conclusions

● Good prediction ➔ Better bitrate selection & adaptation ➔ Improved video QoE

● Key insights on throughput variability
  □ Evolution of intra-session throughput exhibits stateful characteristics.
  □ Similar sessions have similar throughput structures.

● CS2P: Cross-session HMM-based approach
  ● Outperform prior predictors by 50% in midstream prediction error.
  ● Achieve 3.2% improvement to HM and 14% to BB in video QoE.