Developing a Predictive Model for Internet Video Quality-of-Experience

Athula Balachandran, Vyas Sekar,
Aditya Akella, Srinivasan Seshan,
Ion Stoica, Hui Zhang

QoE → $$$

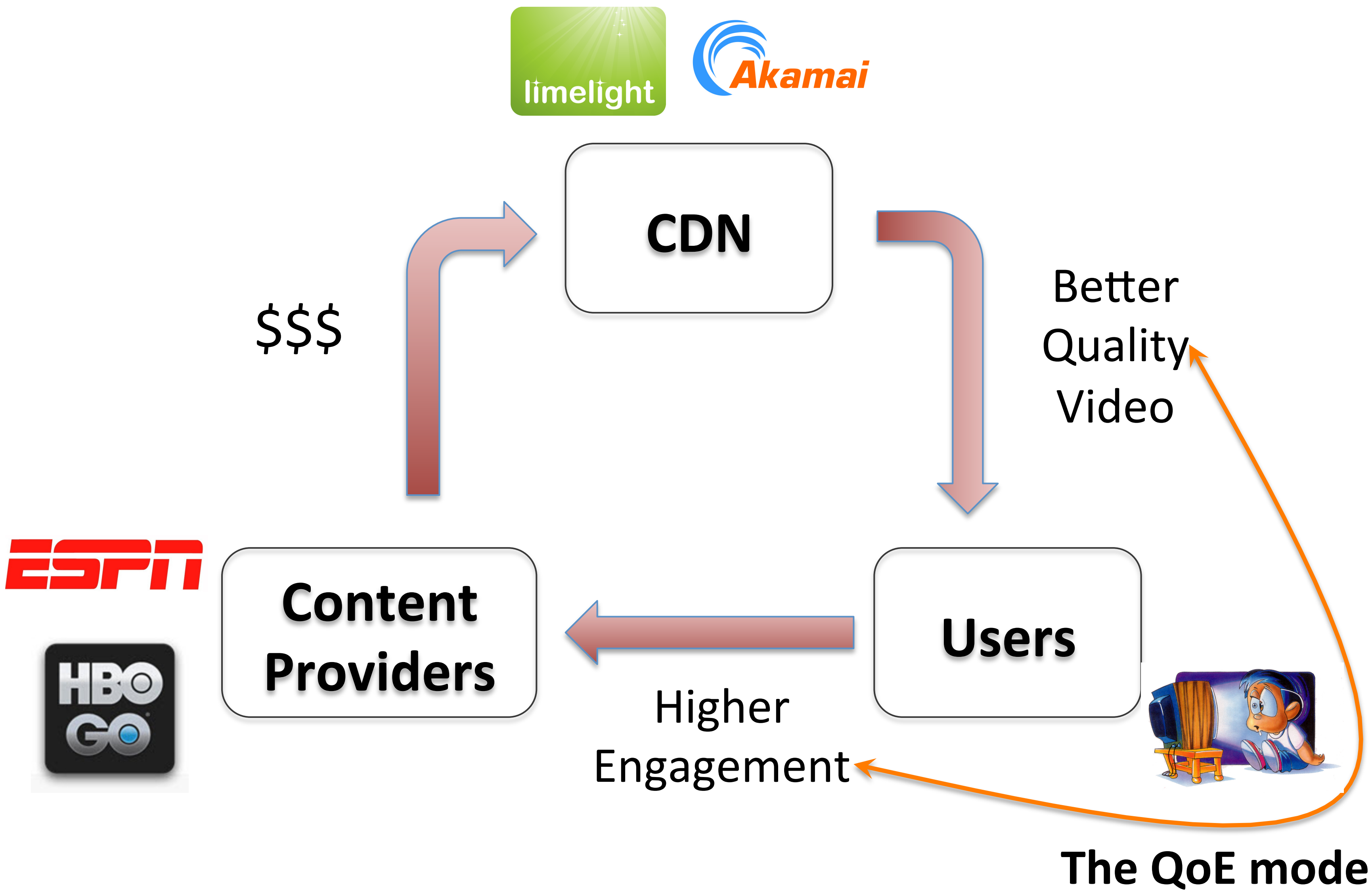
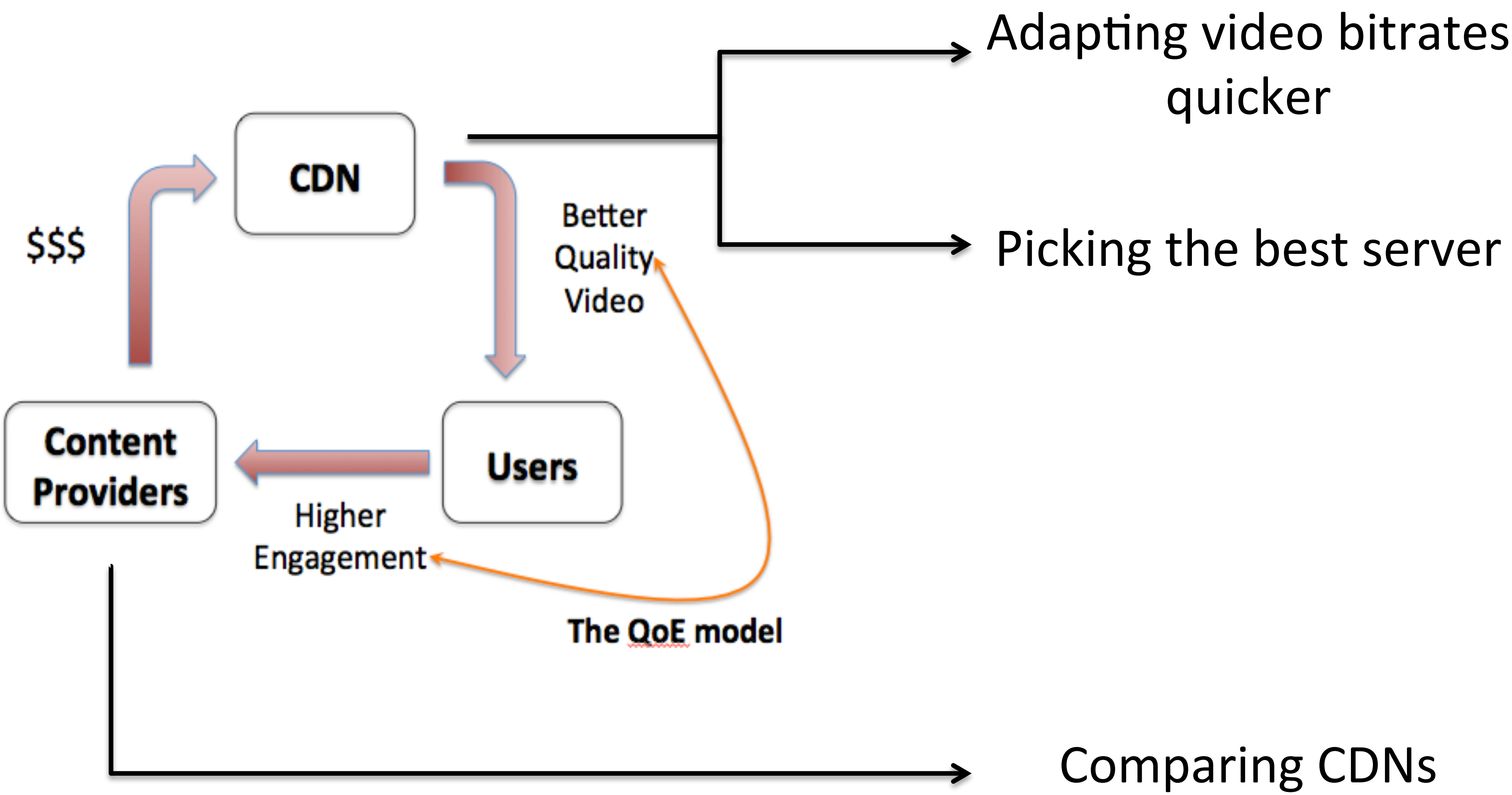


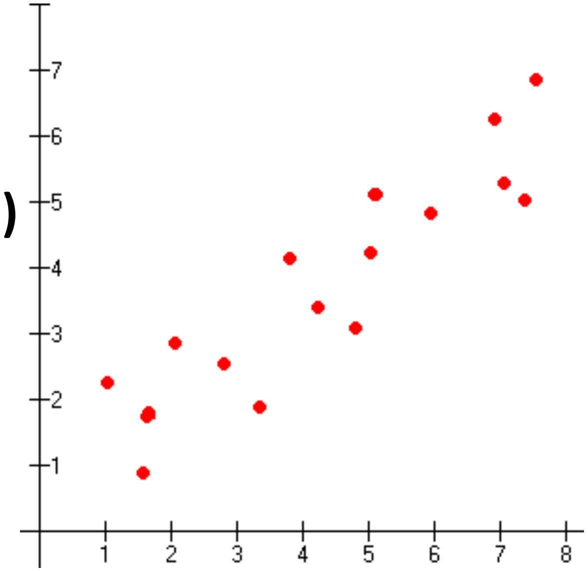
Diagram courtesy: Prof. Ramesh Sitaraman, IMC 2012

Why do we need a QoE model?



Traditional Video Quality Metrics

Subjective Scores
(e.g., Mean Opinion Score)



Objective Scores
(e.g., Peak Signal to Noise Ratio)

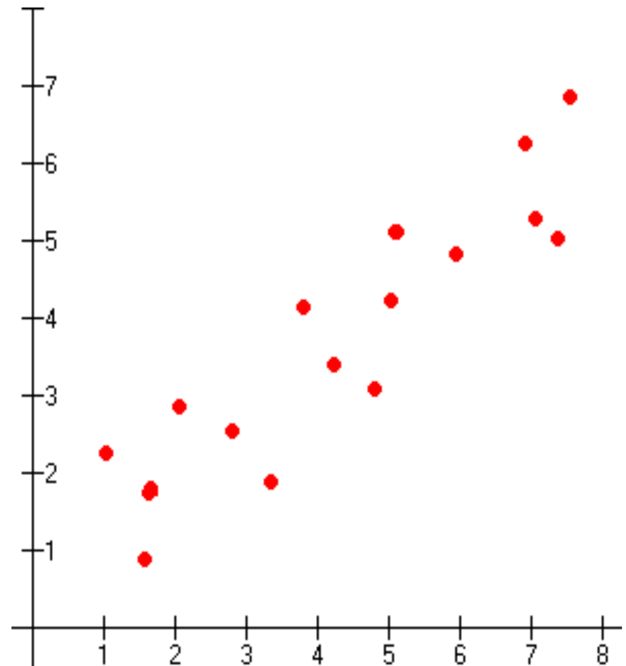
User studies not representative of "in-the-wild" experience

Does not capture new effects (e.g., buffering, switching bitrates)

Internet Video is a new ball game

~~**Subjective Scores (e.g., Mean Opinion Score)**~~

Engagement (e.g., fraction of video viewed)



~~**Objective Scores (e.g., Peak Signal to Noise Ratio)**~~

Quality metrics

Commonly used Quality Metrics

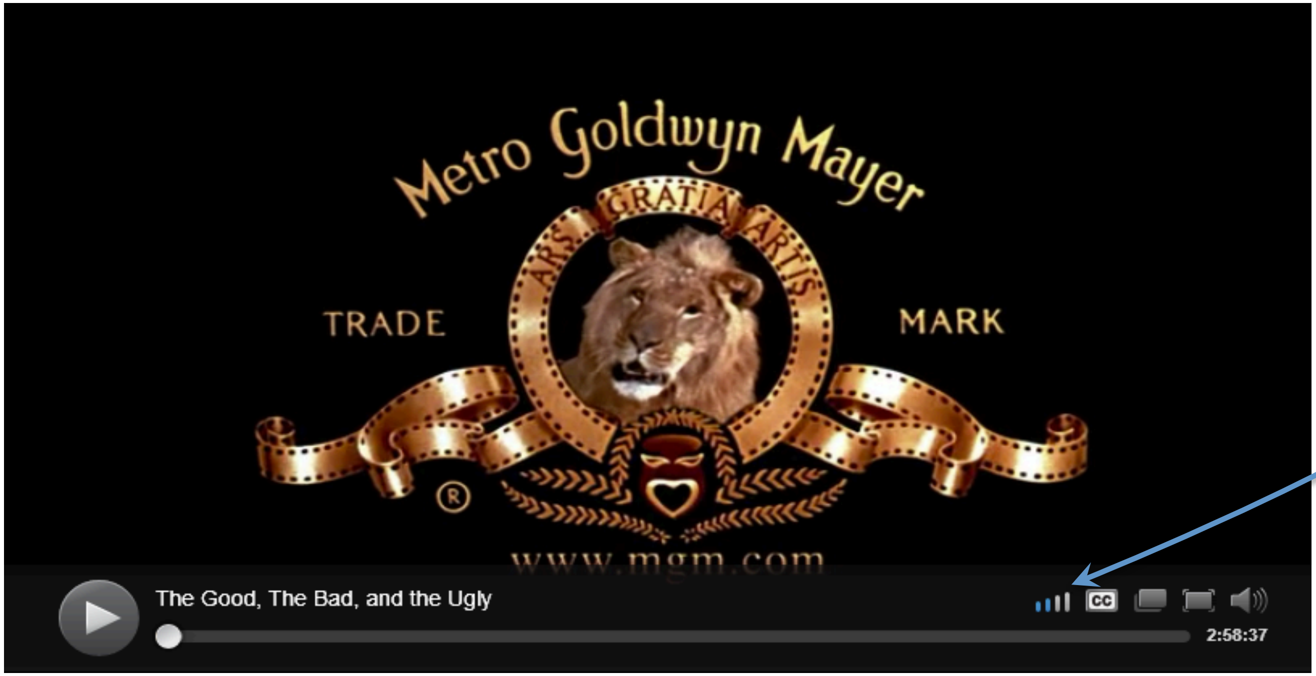
Join Time

Buffering ratio

Rate of switching

Rate of buffering

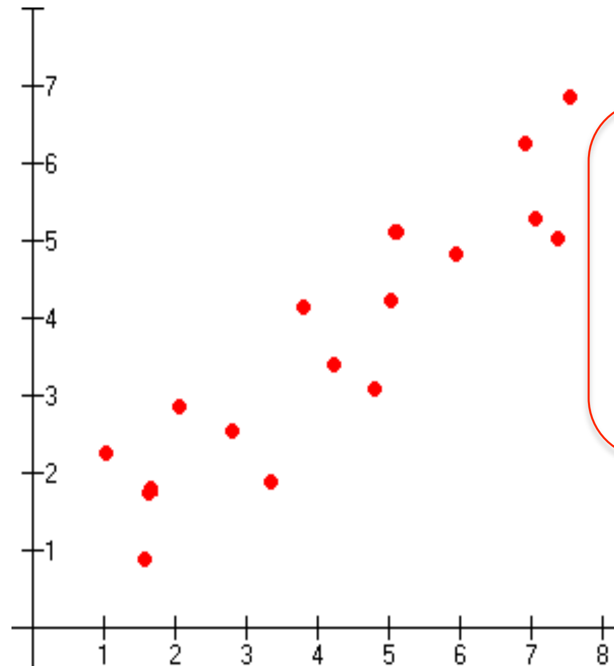
Average Bitrate



Which metric should we use?

~~**Subjective Scores**~~
~~**(e.g., Mean Opinion Score)**~~

Engagement
(e.g., fraction of video viewed)



Today:
Qualitative
Single-metric

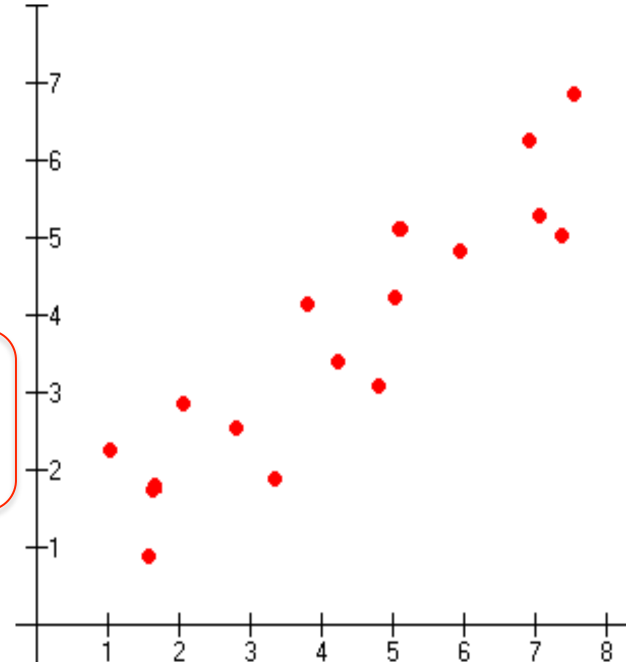
~~**Objective Scores**~~
~~**(e.g., Peak Signal to Noise Ratio)**~~

Quality metrics
Buffering Ratio, Average bitrate?

Unified and Quantitative QoE Model

~~Subjective Scores (e.g., Mean Opinion Score)~~

Engagement (e.g., fraction of video viewed)



~~Objective Scores (e.g., Peak Signal to Noise Ratio)~~

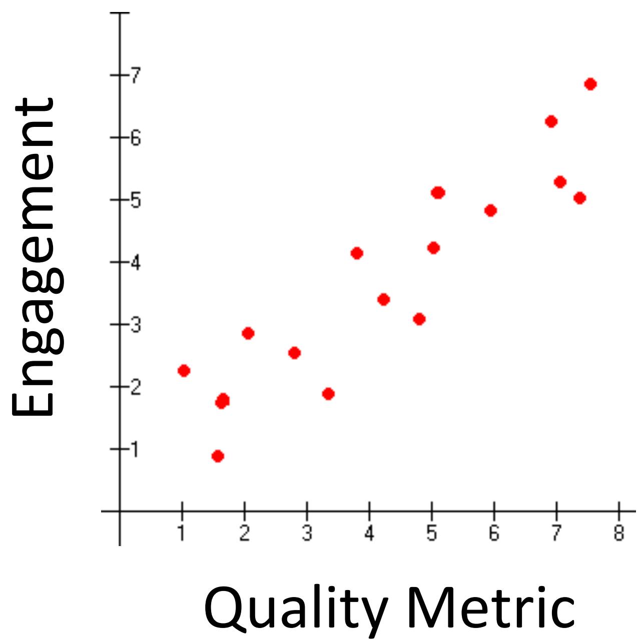
~~Quality metrics Buffering Ratio, Average bitrate?~~

f **(Buffering Ratio, Average bitrate,…)**

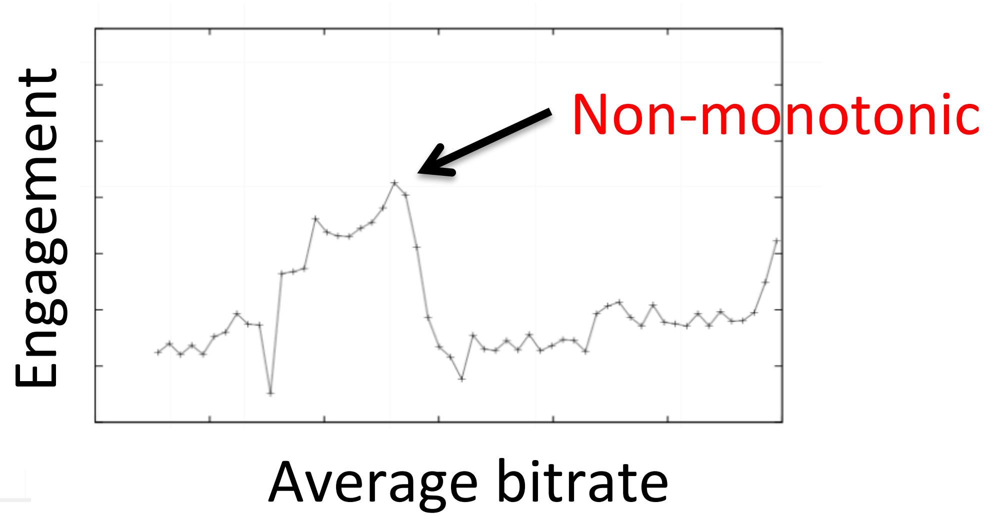
Outline

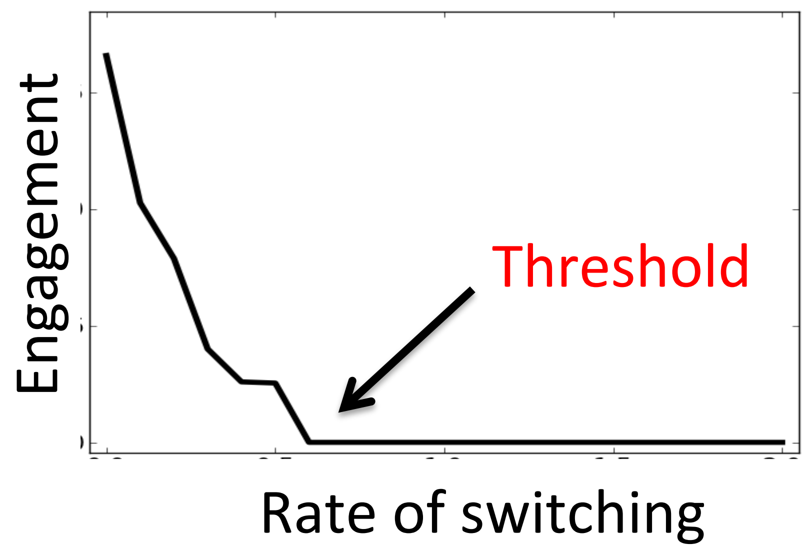
- **What makes this hard?**
- Our approach
- Conclusion

Complex Engagement-to-metric Relationships



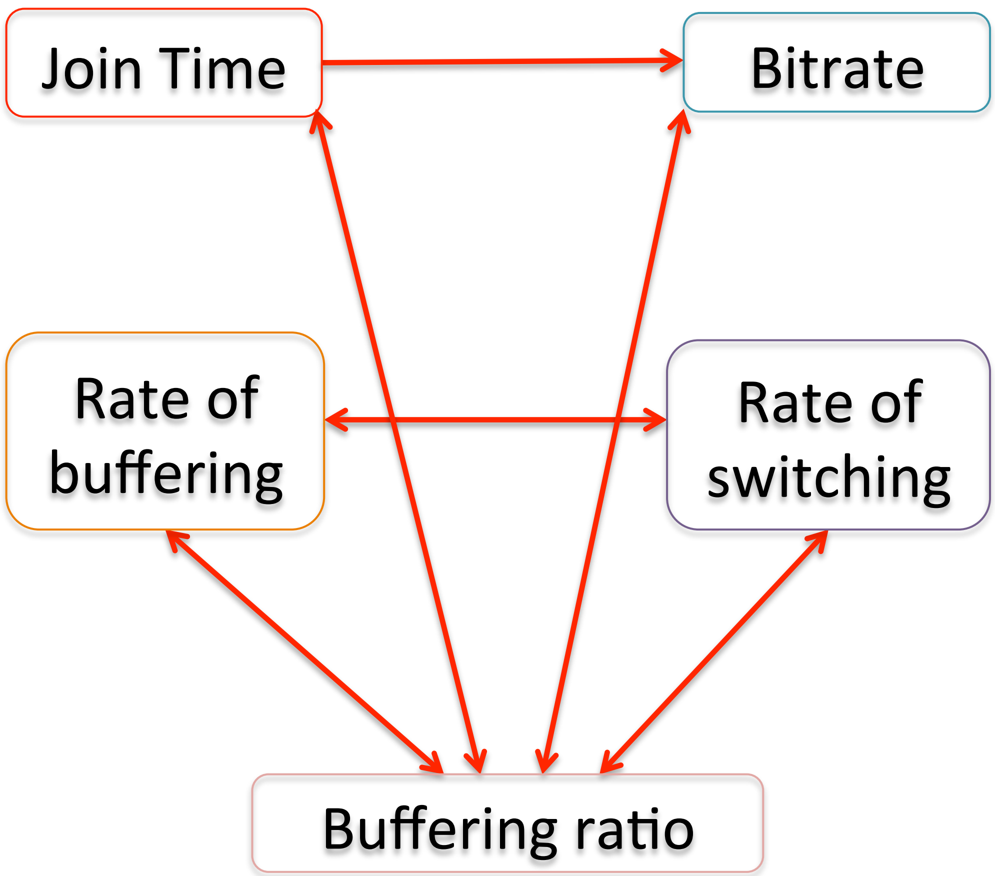
Ideal Scenario

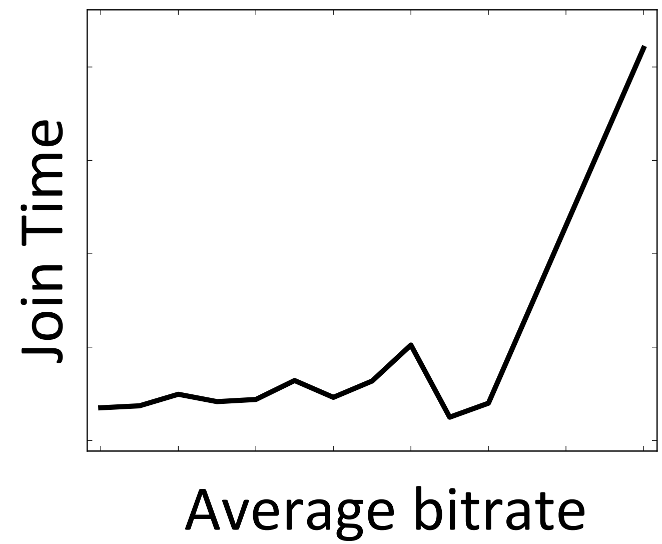


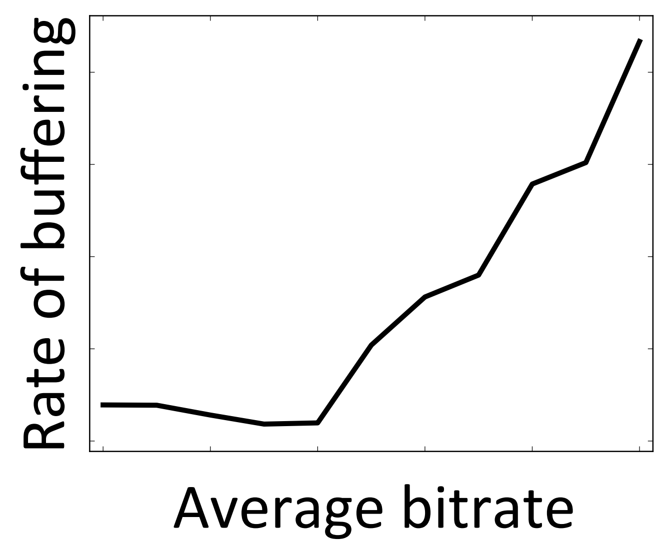


[Dobrian et al. Sigcomm 2011]

Complex Metric Interdependencies



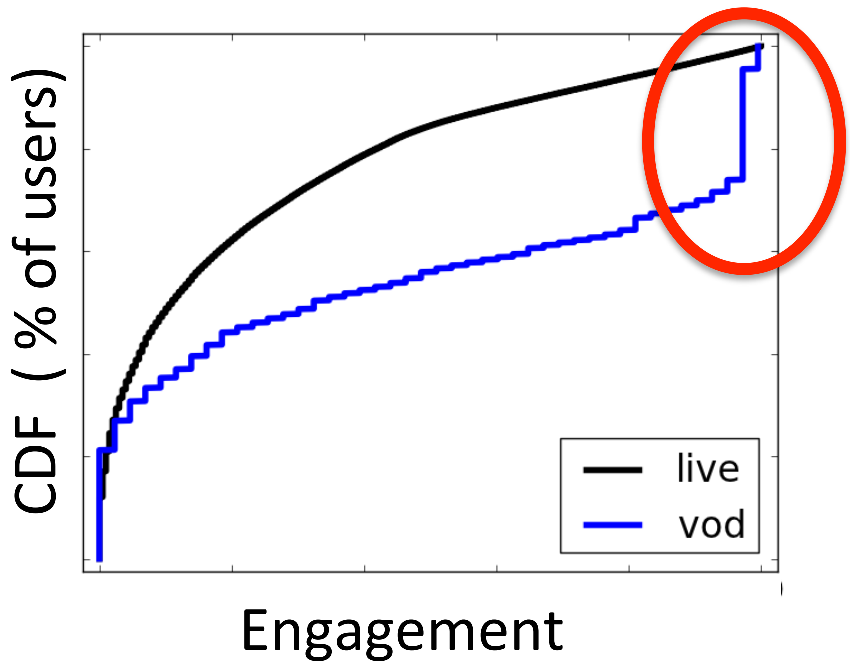




Confounding Factors

Confounding Factors can affect:

1) Engagement

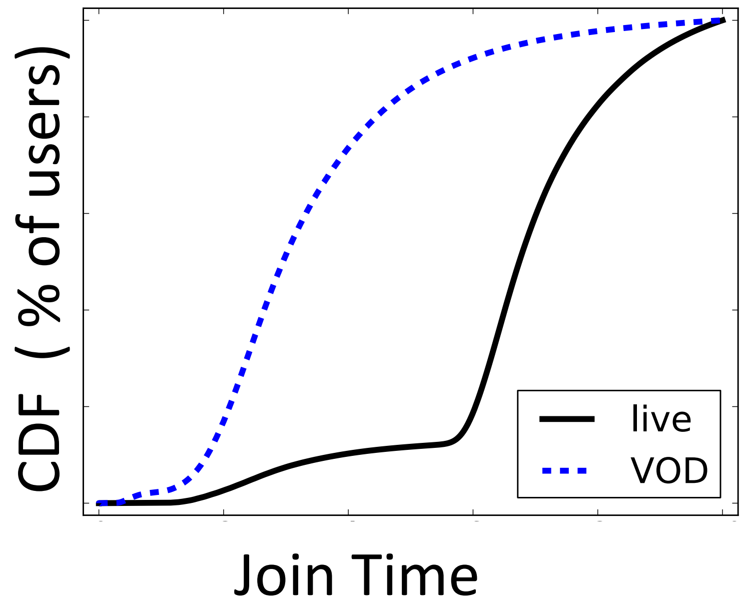


Live and Video on Demand (VOD) sessions have different viewing patterns.

Confounding Factors

Confounding Factors can affect:

1) Engagement
2) Quality Metrics



Live and Video on Demand (VOD) sessions had different join time distribution.

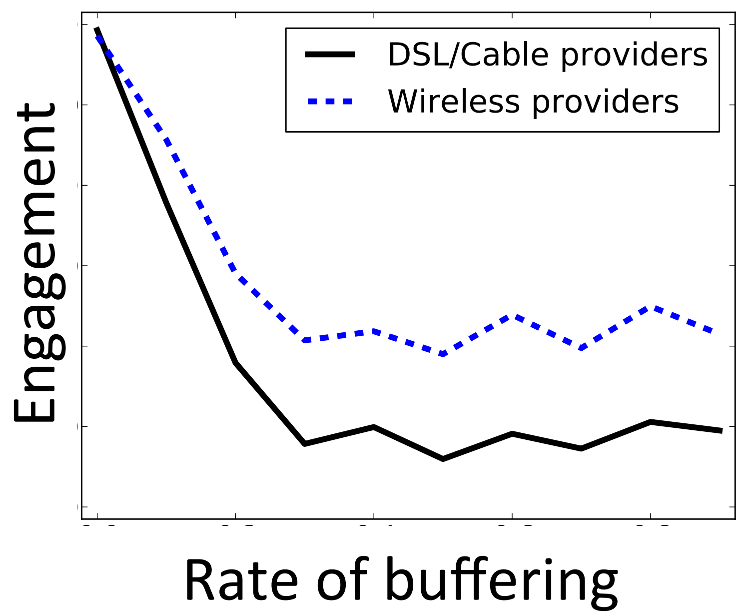
Confounding Factors

Type of Video

Connectivity → DSL/Cable, Wireless (3G/4G)

Confounding Factors can affect:

1) Engagement
2) Quality Metrics
3) Quality Metric → Engagement



Users on wireless connectivity were more tolerant to rate of buffering.

Confounding Factors

Device

Type of Video

Popularity

Location

Connectivity

Time of day

Day of week

Need systematic approach to
identify and incorporate confounding factors

Summary of Challenges

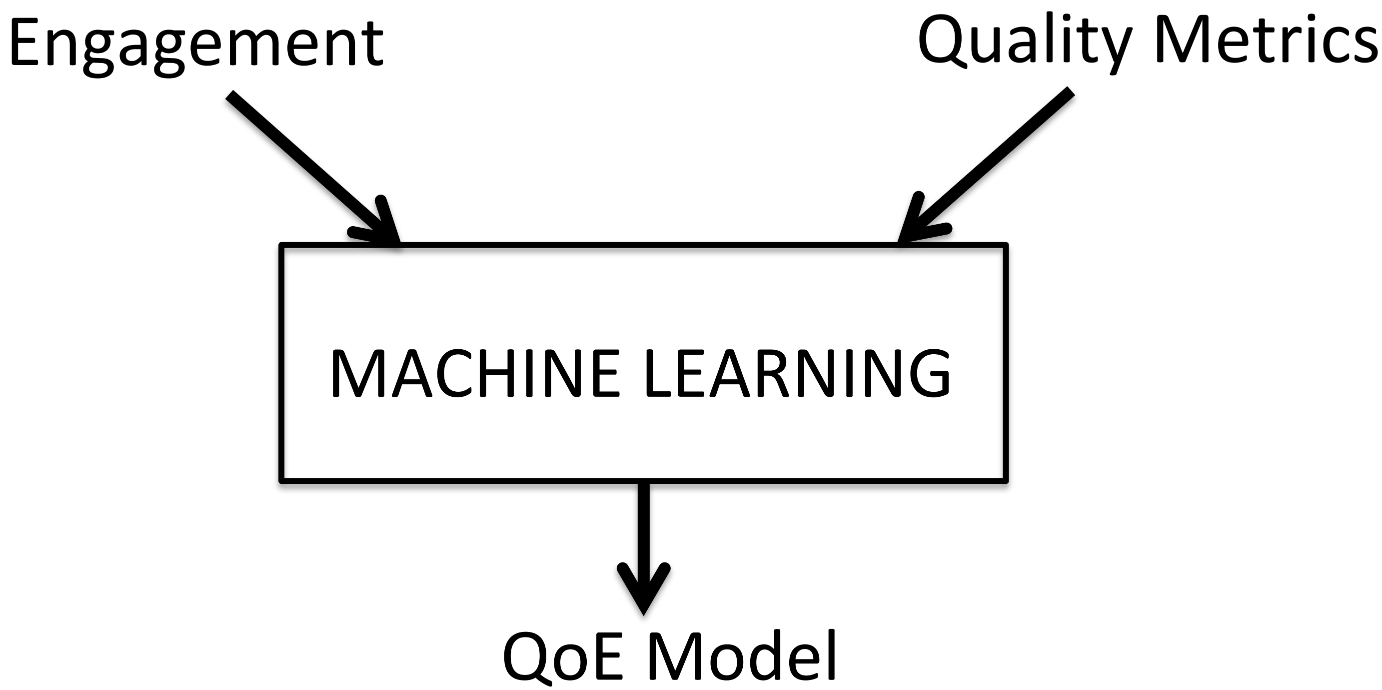
1. Capture complex engagement-to-metric relationships and metric-to-metric dependencies.
2. Identify confounding factors
3. Incorporate confounding factors

Outline

- What makes this hard?
- **Our approach**
- Conclusion

Challenge 1: Capture complex relationships

Cast as a Learning Problem

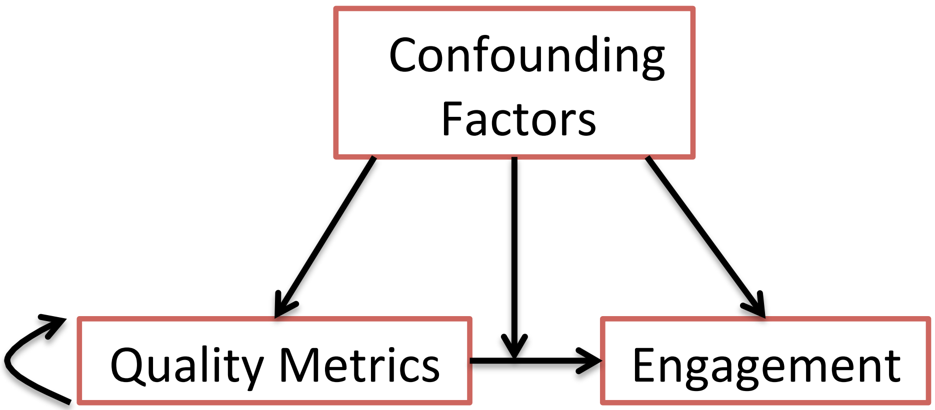


10 engagement classes: 0~10%, 10-20%, …., 90~100% of video length

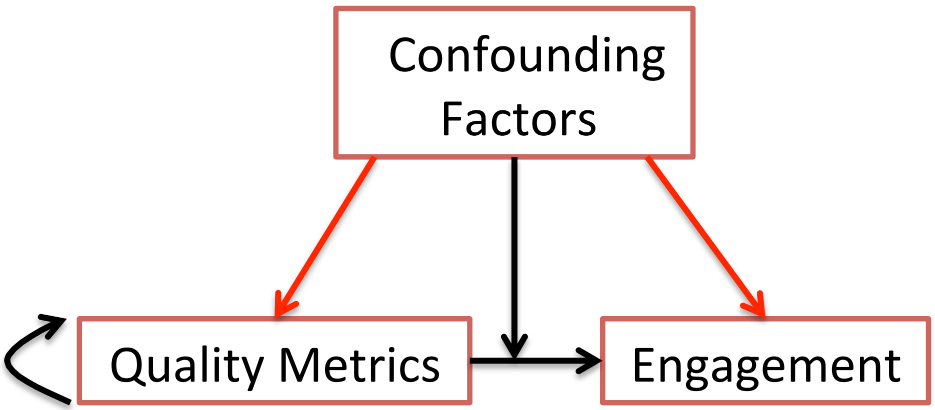
Decision Trees performed the best.
Accuracy of 40% for predicting within a 10% bucket.

Challenge 2: Identify the confounding factors

Test Potential Factors

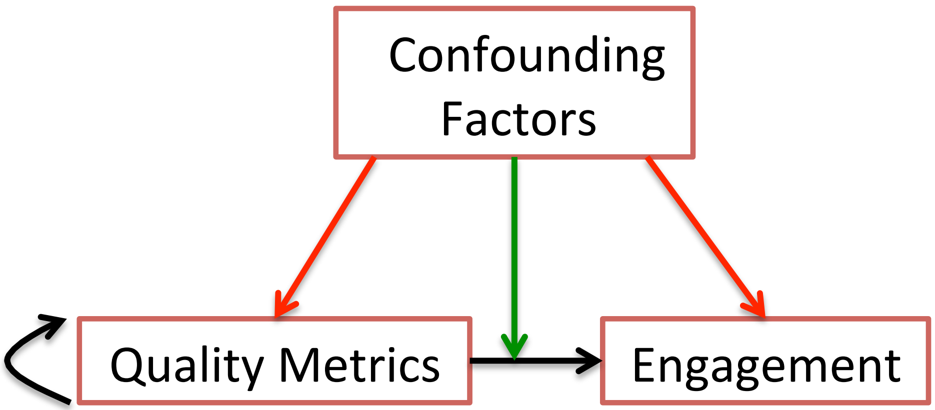


Test Potential Factors



Test 1: Relative Information Gain

Test Potential Factors



Test 1: Relative Information Gain

Test 2: Decision Tree Structure

Test 3: Tolerance Level

Identifying Key Confounding Factors

Factor	Relative Information Gain	Decision Tree Structure	Tolerance Level
Type of video	✓	✓	✓
Popularity	✗	✗	✗
Location	✗	✗	✗
Device	✗	✓	✓
Connectivity	✗	✗	✓
Time of day	✗	✗	✓
Day of week	✗	✗	✗

VOD users on different devices have different levels of tolerance for rate of buffering and average bitrate

Identifying Key Confounding Factors

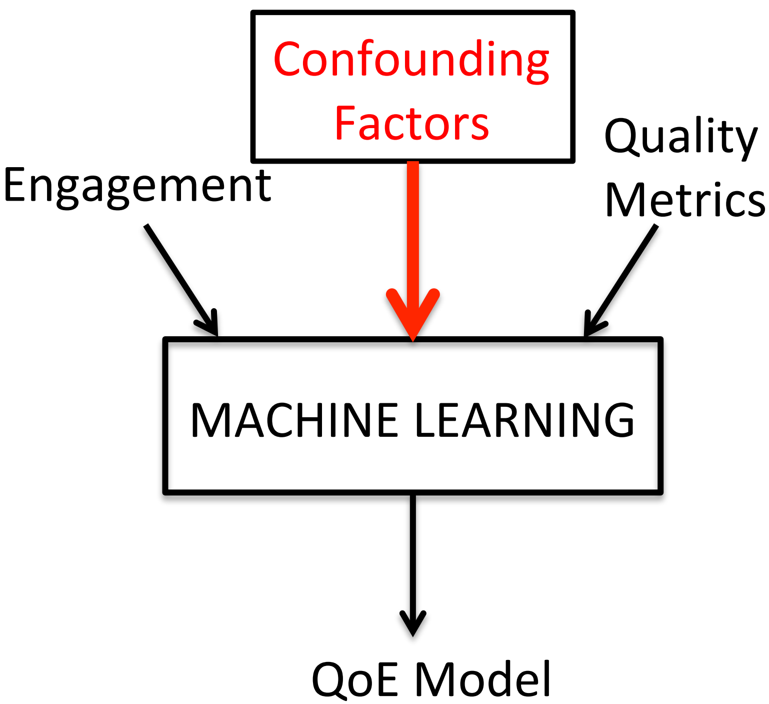
We are doing feature selection here:

Factor	Relative Information Gain	Decision Tree Structure	Tolerance Level
Type of video	✓	✓	✓
Popularity	✗	✗	✗
Location	✗	✗	✗
Device	✗	✓	✓
Connectivity	✗	✗	✓
Time of day	✗	✗	✓
Day of week	✗	✗	✗

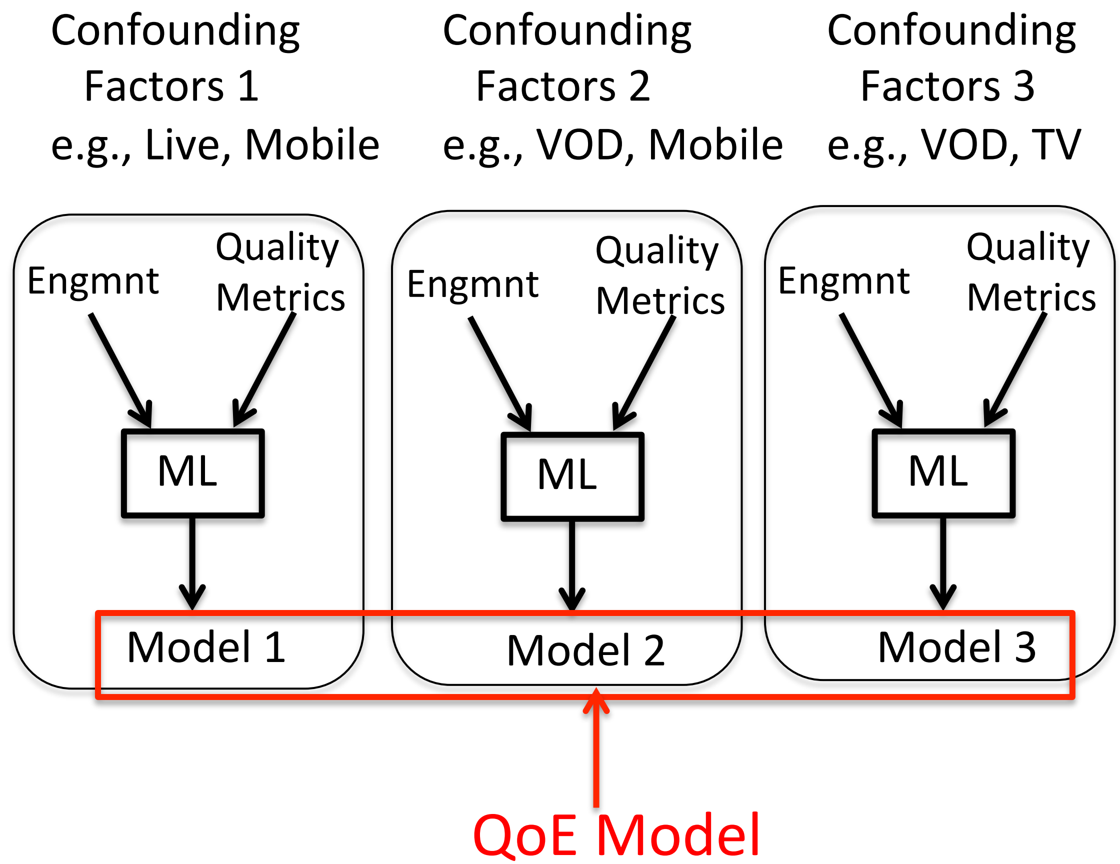
Challenge 3: Incorporate the confounding factors

Refine the Model

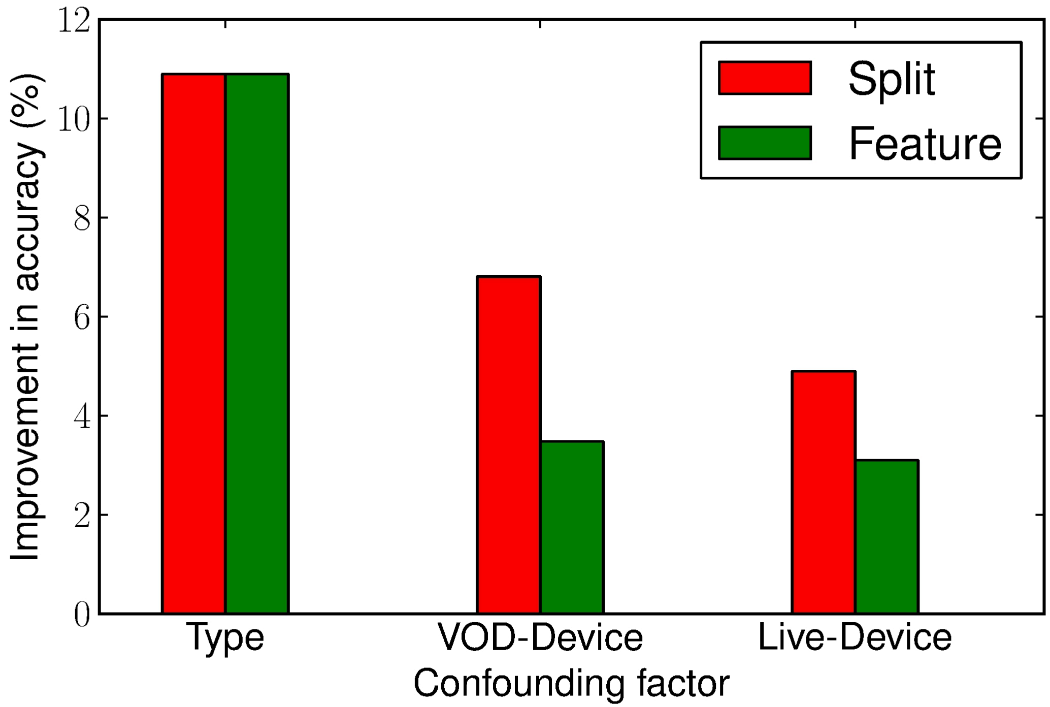
Adding as a feature



Splitting the data



Comparing Candidate Solutions



Final Model: Collection of decision trees
Final Accuracy- 70% (*c.f.* 40%) for 10% buckets

10 engagement classes: 0~10%, 10~20%, … , 90~100% of video length

Introduction to Mixture-of-Experts (MoE)

- **What is MoE?**
  - **Definition:** Mixture-of-Experts (MoE) is a machine learning paradigm that involves a set of models (the "experts") and a gating network. The gating network learns to decide which expert(s) to use for a given input.
- **How Does It Work?**
  - **Expert Models:** These are specialized models trained to perform well on a subset of the data or task.
  - **Gating Network:** A mechanism that dynamically allocates input data to the most relevant experts based on the input's characteristics.
- **Key Benefits:**
  - **Specialization:** By having experts specialize in different parts of the data or task, MoE can achieve higher performance on complex problems.
  - **Scalability:** It allows for efficient scaling by activating only relevant experts for a given input, reducing computational load.
  - **Flexibility:** Experts can be trained independently or jointly, offering flexibility in model design and training.

Applications:

- **Natural Language Processing (NLP):** MoE has been used to improve language models by handling diverse linguistic phenomena.
- **Computer Vision:** For tasks requiring detailed analysis of varied image data.
- **Recommendation Systems:** To cater to diverse user preferences and contexts.

Challenges:

- **Complexity:** Designing and training MoE models can be more complex than traditional models.
- **Resource Intensity:** Requires careful management of computational resources.

Conclusion:

- Mixture-of-Experts represents a powerful approach to handling complex and diverse datasets, enabling models to achieve superior performance by leveraging the strengths of multiple specialized experts

Summary of Our Approach

1. Capture complex engagement-to-metric relationships and metric-to-metric dependencies

   → Use Machine Learning

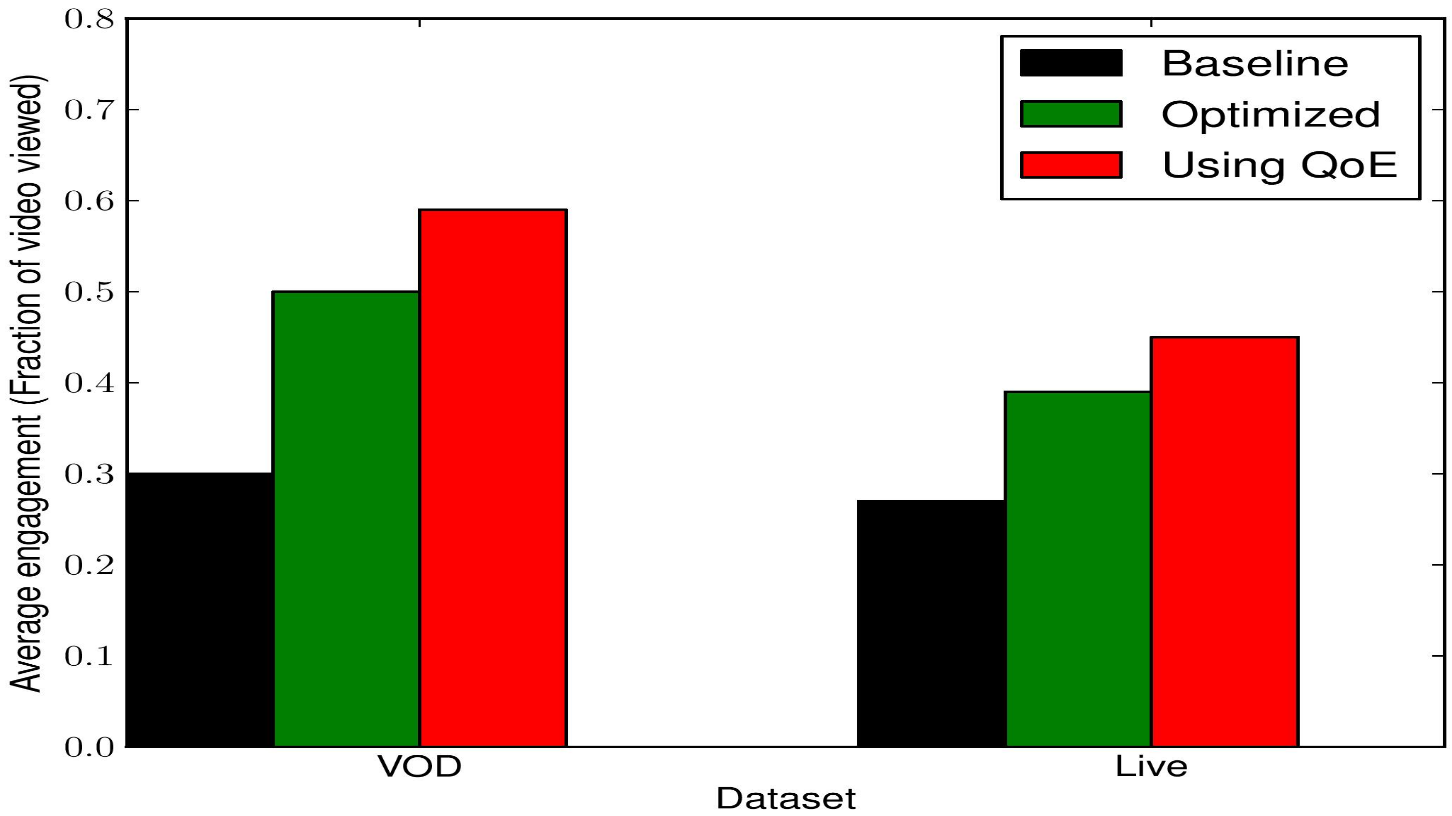
2. Identify confounding factors

   → Tests

3. Incorporate confounding factors

   → Split

Evaluation: Benefit of the QoE Model



Preliminary results show that using QoE model to select bitrate leads to 20% improvement in engagement

Conclusions

- Internet Video needs a unified and quantitative QoE model
- What makes this hard?
  - Complex relationships
  - Confounding factors (e.g., type of video, device)
- Developing a model
  - ML + refinements => Collection of decision trees
- Preliminary evaluation shows that using the QoE model can lead to 20% improvement in engagement
- What's missing?
  - Coverage over confounding factors
  - Evolution of the metric with time