Adaptive Video Streaming and Quality of Experience

Anirban Lekharu
Research Scholar

Department of Computer Science and Engineering
Indian Institute of Technology, Guwahati

Monday 18th November, 2019

Outline

Introduction

- Mobile data traffic will grow at a compound annual growth rate (CAGR) of 47 percent from 2016 to 2021[1]
- Smartphones will surpass four-fifths of mobile data traffic (86 percent) by 2021.
- More than three-fourths of the worlds mobile data traffic will be video by 2021

[1] **Cisco-VNI-1**.

Introduction

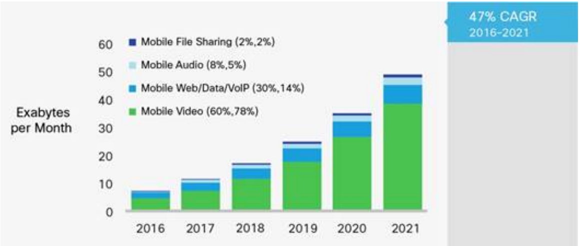


Figure: Global Mobile Traffic Growth by Device Type[1]

[1] **Cisco-VNI-1**.

Challenges

- Providing stutter-free video viewing experience
- Minimizing flickers in the video outputs
- Maximize the aggregate video quality

Proposed Framework

- Video Quality Adaptation Framweork *(VQAF)*[1]
- Video Bit-rate Prediction Model *(VBPM)*[2]
- Deep Learning Based Prediction Model for Adaptive Video Streaming *(LASH)* [3]

---

[1]Lekharu Anirban, et al A QoE Aware SVC Based Client-side Video Adaptation Algorithm for Cellular Networks, ACM ICDCN 18

[2]Lekharu Anirban, et al "A QoE aware LSTM based bit-rate prediction model for DASH video", IEEE, COMSNETS 2018

[3]Anirban Lekharu, et al, "Deep Learning based Prediction Model for Adaptive Video Streaming", IEEE, COMSNETS 2020

Dynamic Adaptive Streaming over HTTP

- Dynamically adjusts the transmission bit rate of a flow over time, based on Mobile Device capability and estimated link conditions
- Multiple versions of the same content offered
- DASH defines: XML Document MPD and Segment formats

Dynamic Adaptive Streaming over HTTP

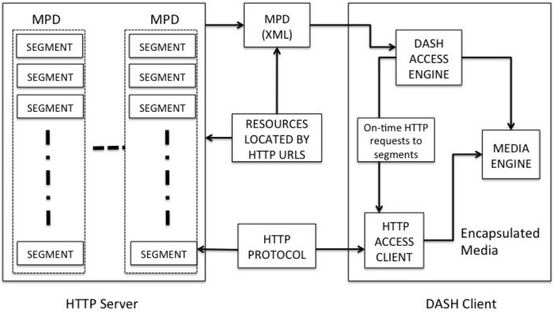


Figure: A Typical DASH system

Media Presentation Description (MPD)

- Description of the available media: Media Presentation Description (MPD)

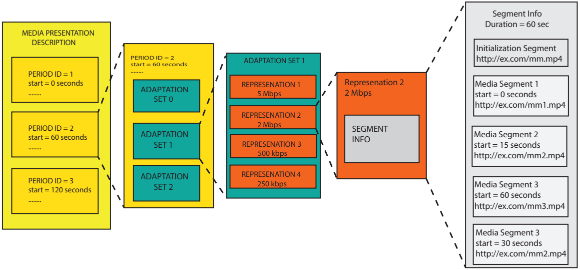


Figure: Hierarchical Structure of the MPD Data Model

Demonstration

- Let's have a Look At How DASH works!!!!

QoE Modelling

The QoE corresponding to a video playback is considered to be composed of three important components:

- Improving the overall Video Quality: The Encoding qualities / enhancement levels at which the video segments are being played.
- Reducing the Flickering Effect: Degree of variation in enhancement levels of consecutive segments, over a certain video playout duration.
- Minimizing the buffering: Maintaining a safe buffer status for client-side video playout.

QoE Modelling

The goal of QoE modeling is to judiciously select enhancement levels of the video segment such that overall encoding quality of the video output is maximized while maintaining playout buffer sizes above a safe threshold and restricting the degree of switching as far as possible.

Related Work

Existing Schemes based on Bandwidth Measurement

- Liu et al.[1] proposed Rate Adaptation for Adaptive HTTP Streaming (RAHS), to improve the QoE
- Uses Equation (1) to estimate the available bandwidth μ

$$\mu = \frac{MSD}{SFT} \tag{1}$$

- Scheme fetches the next higher segment when μ is greater than a switch-up factor
- This scheme cannot avoid frequent changes in Video Quality

[1] **Liu:2011**.

Related Work

Existing Schemes based on Buffer Occupancy

- Le et al.[1] proposed the BAHS (Buffer-based Adaptation for Adaptive HTTP Streaming) to improve the QoE
- If Buffer occupancy is low then the BAHS decreases the number of selectable Video Quality level
- Slowly adapts to network bandwidth variation, it may cause buffer underflow when the network bandwidth decreases rapidly

[1] **le2013buffer**.

Related Work

Existing Schemes based on both Bandwidth and Buffer Occupancy

- Kim et al.[1] proposed VAS which adapts Video Quality based on bandwidth measurement and buffer occupancy of client
- Increases the Video Quality when condition (2) and (3) are meet

$$Th_{seg}[i] > R_{curr+1} \tag{2}$$

$$B_{curr}[i] > B_{up} \tag{3}$$

where $Th_{seg}[i]$ = *Segment Throughput,* B_{up} = *Buffer Threshold*

- Fails to minimize the frequent changes of Video Quality in short-term network bandwidth variation

[1] **kim2016video**.

Limitations

- Fails to avoid unnecessary video quality switches
- Buffer Underflow Conditions

Motivation

Objectives of the Proposed Work

- Minimize the frequency of Video Quality Switches.
- Minimize the Buffer Underflow events
- Maximize the Average Video Quality

Problem Definition

Proposed framework based on the Buffer occupancy and Throughput on the client side.

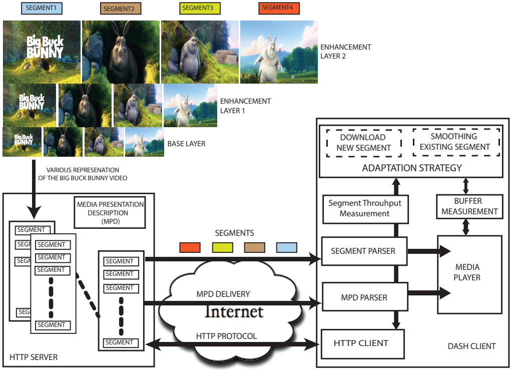


Figure: Proposed Video Quality Adaptation Framework

DASH using SVC

Advantages of using SVC DASH

- A Video Stream at the j^{th} Enhancement level EL_j contains all the substreams present in EL_{j-1} along with one additional substream
- SVC DASH provides the flexibility of enhancing the already downloaded video segment

Rate Adaptation Strategy

The proposed framework employs a two phase mechanism:

- Download Phase
  - New segments are fetched when

$$(B_{cur} < Th_{safe}) \tag{4}$$

B_{cur} is current buffer size, Th_{safe} is the safety threshold

- Smooth-out Phase
  - Already downloaded segments are repeatedly upgraded when

$$(B_{cur} > Th_{safe}) \tag{5}$$

Downloading Phase

Downloading a New Video Segment

- An upgradation is allowed *iff*:

$$(B_{cur} > V_{up}) \wedge (\bar{r}(t) > R_{cur+1}) \tag{6}$$

V_{up} is calculated as:

$$V_{up} = \frac{R_{cur+1}}{R_L} Th_{safe} \tag{7}$$

where $\bar{r}(t)$ is the instantaneous bandwidth, R_L is bit-rate of the highest enhancement level

Downloading Phase

Downloading a New Video Segment

- Enhancement level for downloads is degraded *iff* :

$$(B_{cur} < V_{down}) \vee (\overline{r}(t) < R_{cur}) \tag{8}$$

V_{down} is calculated as:

$$V_{down} = \frac{R_{cur}}{R_L} Th_{safe} \tag{9}$$

Smooth-out Phase

Smoothing-out Already Downloaded Segments

- Selecting a Candidate Set

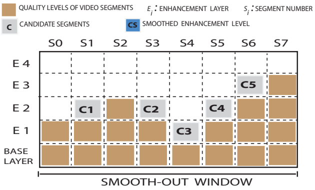


Figure: Selecting a Candidate Set

- $\langle C1 = 1, C2 = 3, C3 = 4, C4 = 5, C5 = 6 \rangle$ are the selected candidates.

Smooth-out Phase

Smoothing-out Already Downloaded Segments

- Determining the Best Candidate Segment

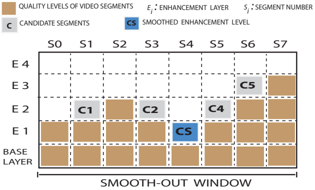


Figure: Determining the Best Candidate Segment

- The switching factor Ψ_{ci} for a candidate S_{ci} is defined as:

$$\Psi_{ci} = |\uparrow l_{ci} - l_{ci-1}| + |\uparrow l_{ci} - l_{ci+1}| \qquad \forall i \in C \tag{10}$$

Experimental Setup

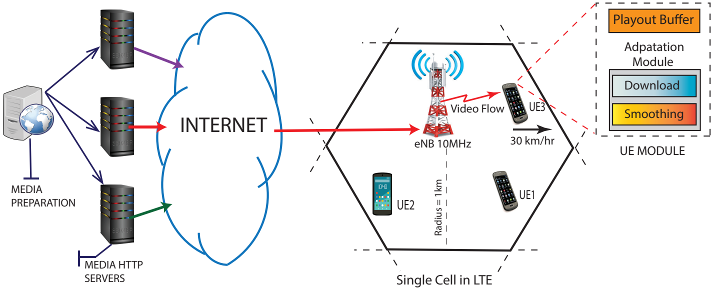


Figure: Scenario for Experimental Setup in a Single Cell

Comparative Results for Switching Instability

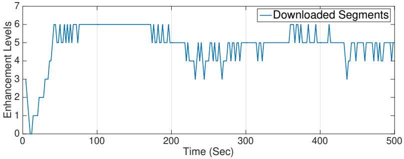


Figure: Delivered Video Quality Vs. Time for VAS[1]

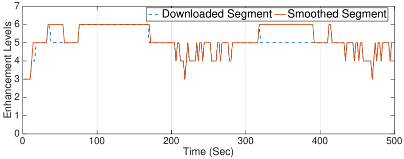


Figure: Delivered Video Quality Vs. Time for VQAF

[1] **kim2016video**.

Comparative Results for Switching Instability

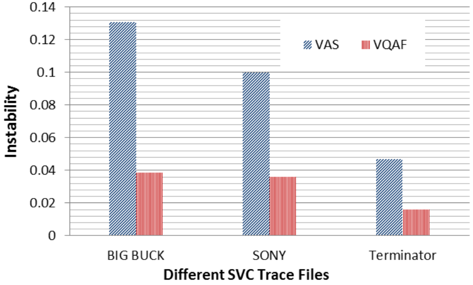


Figure: Comparison of Switching Instability

Comparative Results for Aggregate Video Quality:

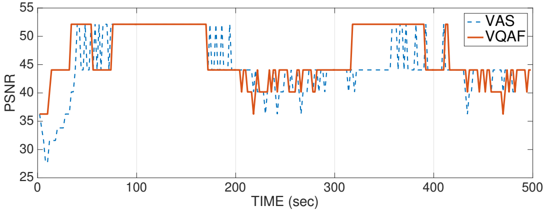


Figure: Comparative Results for Aggregate Video Quality (in PSNR)

Comparative Results for Aggregate Video Quality:

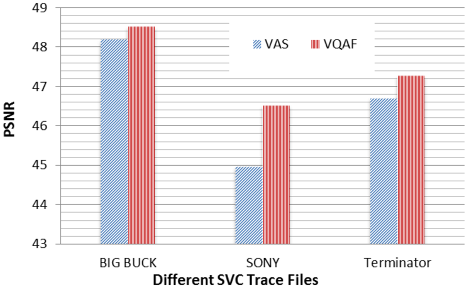


Figure: Comparative Results for Aggregate Video Quality (in PSNR)

Comparative Results for Buffer Occupancy

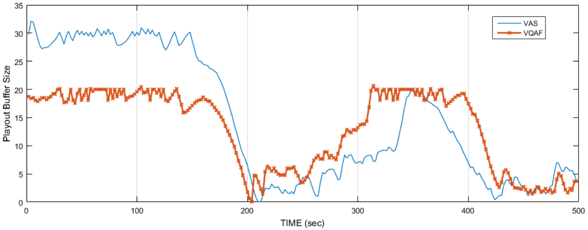


Figure: Comparison of Buffering Time

Comparative Results for Buffer Occupancy

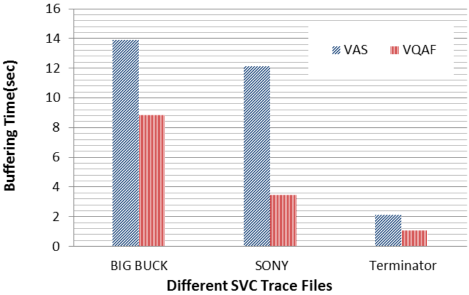


Figure: Comparison of Buffering Time

Video Bit-rate Prediction

- Learn the sequence of video bit-rates affected by various QoE parameters.
- Different Machine Learnign techniques are available for prediction.
- LSTM (Long Short Term Memory) is a well known model.

Long Short-Term Memory Network (LSTM)

- Three types of gates: Forget, Input and Output

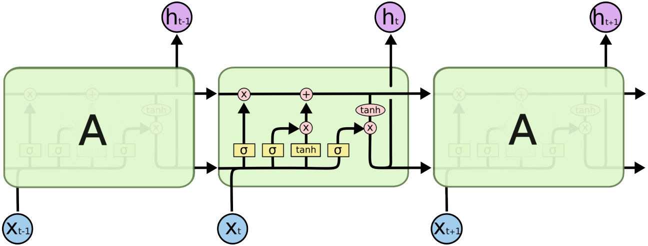


Figure: Repeating module in an LSTM containing four interacting layers

Related Work

Machine Learning based Prediction Model

- Chien et al.[2] proposed a Machine Learning based Rate Adaptation strategy, to improve the QoE.
- Existing Rate Adaptation Strategy used to label the best video rate corresponding to the feature.
- Random Forest used to train the model and predict the bit-rate.
- Unable to optimise all the QoE parameters.

[2] **chien2015machine**.

Objectives

- Capturing the complex features that affect video quality.
- Accurate predictions of video bit-rate.
- Optimizing the concerned QoE metrics simultaneously.

Problem Definition

Proposed framework for Model Training

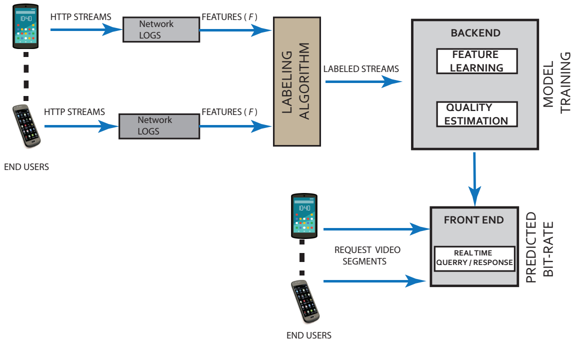


Figure: VBPM Sytem Architecture

Critical Features Selection

- Selecting the Critical Features set $\mathcal{F}$ affecting the QoE of a client

$$\mathcal{F} = (time, BL, BT, IS) \tag{11}$$

- $\mathcal{F}$ consist of the following features:
  - BL: The current Buffer Level of the client
  - BT : Buffering Time
  - IS: Instability Metrics
- $\mathcal{F}$ corresponds to the video bit-rate $\mathcal{V}$

$$\mathcal{F}_i(time_i, BL_i, BT_i, IS_i) = \mathcal{V}_i \tag{12}$$

VBPM System Architecture

The proposed framework consists of three phases:

- Rate Labeling Phase
- Model Training Phase
- Video Bit-rate Predictions

Trace based Evaluation

Evaluation of Training Model

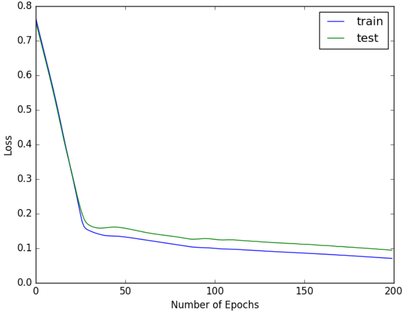
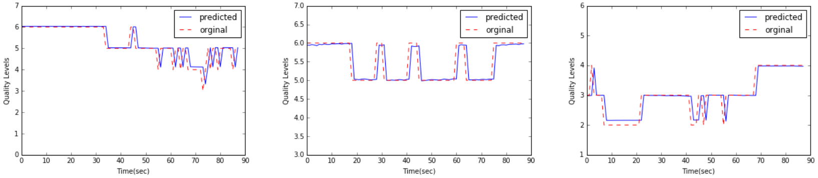


Figure: Train and Test loss from the LSTM model during Training

Trace based Evaluation

Video Bit-rate Predictions



(a) Sony Demo (b) Terminator (c) Big Buck Bunny

Figure: Observed Vs Predicted bit-rate requested video segment using the LSTM model

Trace based Evaluation

Table: Performance of VBPM prediction model using LSTM

Test Video Sequences	*VBPM (using LSTM)*			
	Quality Levels	Quality Levels(True)	RMSE	Correlation
Sony Demo	5.29	5.25	0.42	0.82
Terminator	5.49	5.48	0.297	0.82
Big Buck Bunny	3.03	3.01	0.34	0.87

Trace based Evaluation

Table: Comparative Result for QoE metrics

Test Video Sequences	*VAS*			*VBPM (using LSTM)*		
	Quality Levels	Buffer Status	Instability	Quality Levels	Buffer Status	Instability
Sony Demo	4.95	18.74	0.065	5.29	14.14	0.04
Terminator	5.31	28.06	0.07	5.49	19.11	0.019
Big Buck Bunny	2.7	5.49	0.17	3.03	10.31	0.04

Trace based Evaluation

Table: Comparison of Different Machine Learning Models

Test Video Sequences	LSTM		Random Forest		Linear Regression	
	RMSE	Cor	RMSE	Cor	RMSE	Cor
Sony Demo	**0.42**	**0.82**	0.46	0.78	0.69	0.6
Terminator	**0.297**	**0.82**	0.31	0.78	0.47	0.4
Big Buck Bunny	0.34	**0.87**	**0.32**	0.86	0.5	0.66

Thank You!