ITU KALEIDOSCOPE
SANTA FE 2018

Machine learning for a 5G future

Double Sarsa Based Machine Learning to Improve Quality of Video Streaming over HTTP through Wireless Networks

Dhananjay Kumar
Department of Information Technology
Anna University, MIT Campus, Chennai, India
Email: dhananjay@annauniv.edu

26-28 November
Santa Fe, Argentina

Outlines

- Motivation
- Machine Learning Approach
  - Client Server Architecture
- Sarsa and Double Sarsa Learning Methods
  - Softmax and ε-greedy policy
- The ITU-T P.1203 model
- Development of Algorithms
- Implementation Aspects
- Results and Discussions
- Acknowledgements

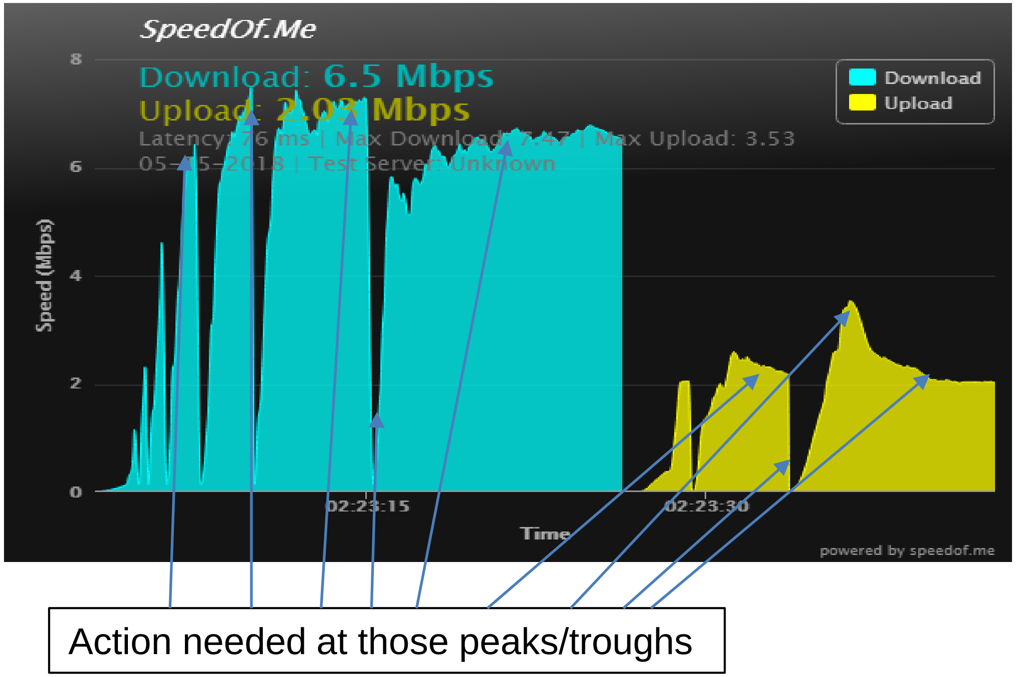
Motivation

Bit Rate Fluctuation in Internet over Cellular Wireless Networks

- ✓ Internet service based on Best-effort model
- ✓ Received signal strength fluctuation in wireless channel

Possible Solution for Adaptive Video Streaming

- ➢ Link bit rate prediction at receiver
- ➢ Adaptive QoS approach
- ➢ Improving QoE using machine learning techniques



Machine Learning in Client-Server Environment

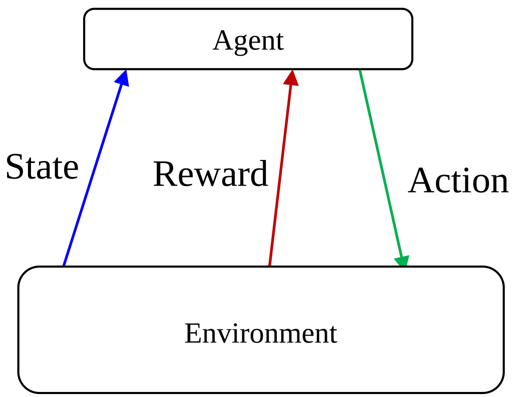


Fig.1. A Q-learning (reinforcement) approach

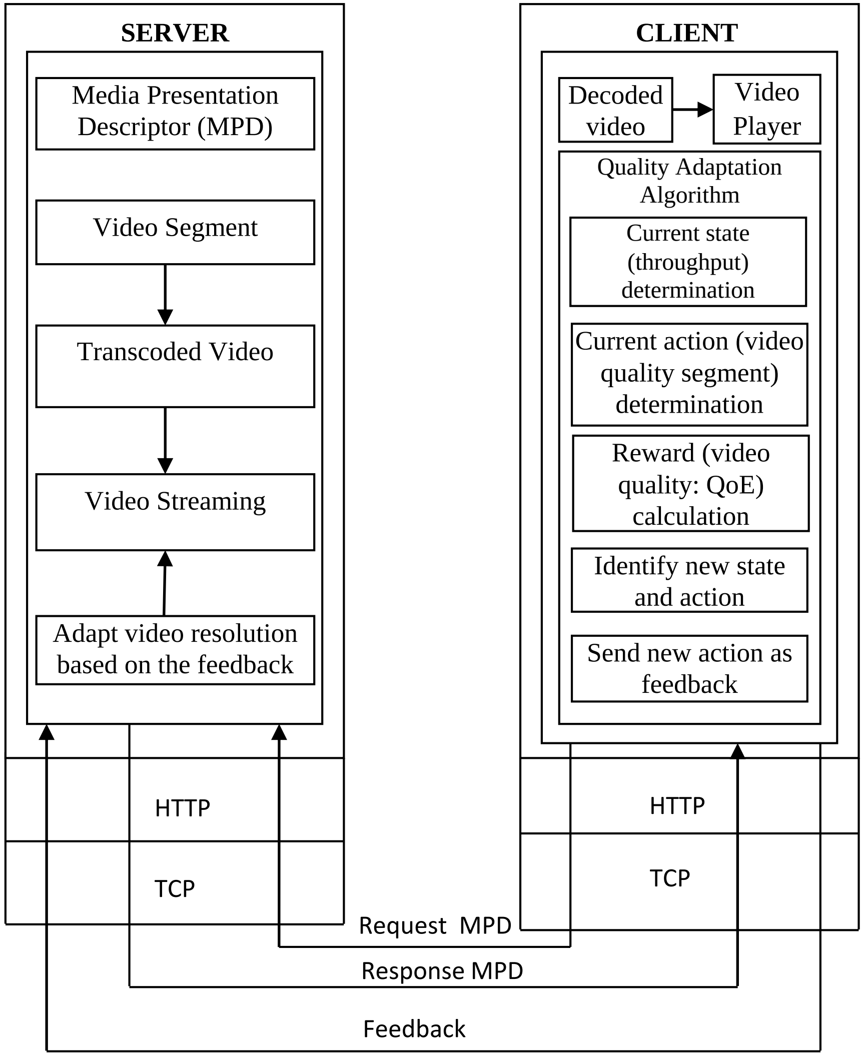


Fig.2. Client-Server model for video streaming

State Action Reward State Action (Sarsa)

The Q-value represents the learned value that the system will acquire by taking the action ***a*** in state ***s*** formulated as

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma\, Q(s',a') - Q(s,a)] \tag{1}$$

where $\boldsymbol{\alpha}$ represents learning rate and $\boldsymbol{\gamma}$ the discount factor, $Q(s', a')$ the Q-value resulting from new action $\boldsymbol{a}'$ in state $\boldsymbol{s}'$.

Double Sarsa

In Double Sarsa method, two action-value estimates $Q^A(s,a)$ and $Q^B(s,a)$ are defined in improving the performance of Sarsa in stochastic scenarios.

The update rule for Double Sarsa is given as follows.

$$Q^A(s,a) \leftarrow Q^A(s,a) + \alpha[r + \gamma\, Q^B(s',a') - Q^A(s,a)] \tag{2}$$

Softmax Policy: Action-selection is based on probabilities determined by ranking the value-function estimates using a Boltzmann distribution given by

$$\pi(a/s) = \frac{e^{\frac{Q^A(s,\mathrm{a})+Q^B(s,\mathrm{a})}{\tau}}}{\sum_b e^{\frac{Q^A(s,\mathrm{b})+Q^B(s,\mathrm{b})}{\tau}}} \tag{3}$$

where $\boldsymbol{\tau}$ is a positive parameter called temperature, and $\boldsymbol{b}$ represents all possible actions.

ε-greedy policy: It uses the average of the two tables to determine the greedy action as follows

$$\pi(a/s) = \begin{Bmatrix} 1-\varepsilon, if\ \mathrm{a} = argmax_{a' \varepsilon\, A(s)}[Q^A(s,a') + Q^B(s,a')] \\ \frac{\varepsilon}{N_{a},-1},\ otherwise \end{Bmatrix} \tag{4}$$

where $\boldsymbol{\pi(a/s)}$ is the probability of taking action $\boldsymbol{a}$ from states, and $\boldsymbol{N_a}$ is the number of actions that can be taken from state $\boldsymbol{s}$.

The ITU-T P.1203 Video Quality Estimation Model

The **ITU P.1203 Core model**

- Quantization degradation (Dq)
- Upscaling degradation (Du)
- Temporal degradation (Dt)

Integration

D = max(min(Dq+Du+Dt,100),0)

The estimated **Mean Opinion Score (MOS)** is calculated as:

$$\widehat{\mathrm{MOS}} = \begin{cases} \widehat{\mathrm{MOSq}}, & \text{if Du} = 0 \text{ and Dt} = 0 \\ \mathrm{MOSfromR}(\widehat{\mathrm{Q}}), & \text{otherwise} \end{cases}$$

where $\widehat{\mathrm{MOS}}$ and $\widehat{\mathrm{Q}}$ are the estimated video encoding qualities on two different scales: $\widehat{\mathrm{MOS}} \in [1,5]$ and $\widehat{\mathrm{Q}} \in [0,100]$.

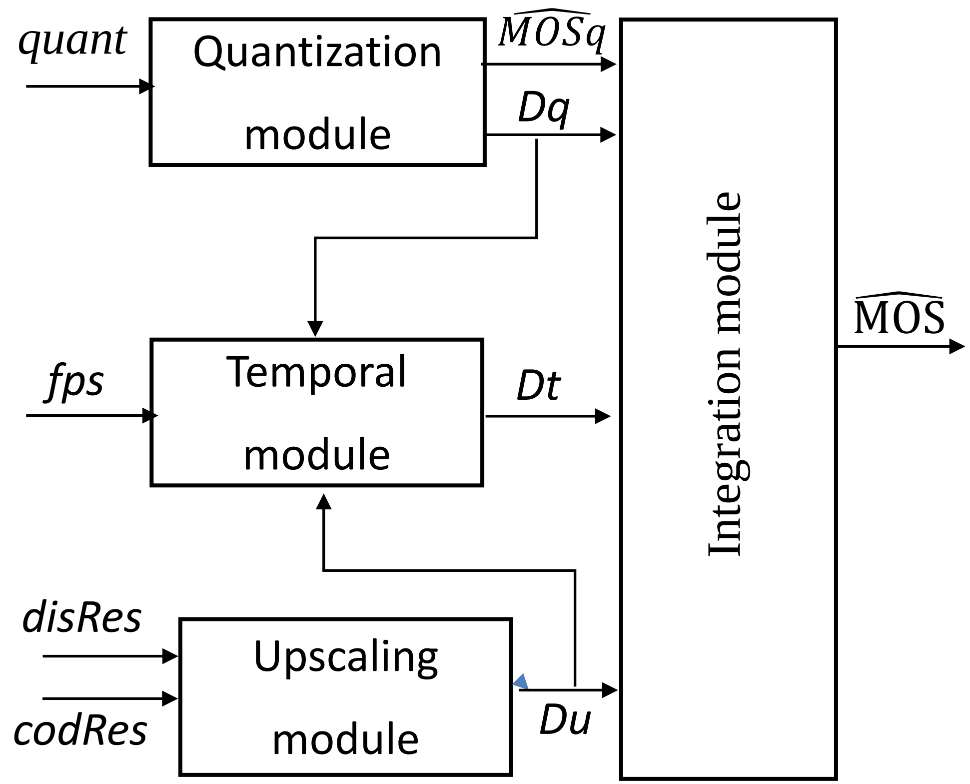


Fig. The Video Quality Model (P_v module) in the ITU-T P.1203

Algorithm Development

- **Double Sarsa Based Adaptation Algorithm**
  - Exploration Policy
    - ***Softmax(Q,s):*** Double Sarsa – Softmax (DS-S)
    - ***ε-greedy(Q,s):*** Double Sarsa – Greedy (DS-G)
- **QoE Driven Video Streaming Strategy with Future Information**
  - A Probabilistic Bandwidth Prediction Model Based Approach
    - QoE Driven Strategy (QD-S)**

** L. Yu, T. Tillo, J. Xiao, “QoE-Drien Dynamic Adaptive video Streaming Strategy with Future Information,” IEEE Trans. on Broadcasting Society, vol. 63, no. 3, pp. 523 – 534, Sep. 2017.

Implementation Aspects:
State and Action Mapping

Throughput Range	State
0-199	0
200-399	1
400-599	2
600-899	3
900-1199	4
1200-1499	5
1500-1799	6
1800-2099	7
2100-2399	8
>2400	9

Action	Resolution		Frame Per Second
	Width	Height	
1	426	240	24
2	426	240	27
3	426	240	30
4	640	360	24
5	640	360	27
6	640	360	30
7	854	480	24
8	854	480	27
9	854	480	30
10	1280	720	24
11	1280	720	27
12	1280	720	30
13	1920	1080	24
14	1920	1080	27
15	1920	1080	30

Test Parameters as per ITU Recommendations

Standards	Parameters	Metrics
ITU-T J.247	Transmission	Errors with packet loss
	Frame rate	5fps to 30fps
	Video codec	H.264/AVC (MPEG-4 part10), VC-1, Windows Media9, Real Video (RV10), MPEG-4 Part 2
	Temporal errors	Maximum of 2 seconds
ITU-T P.1203.1	Input video length	20 seconds
	Video resolution / bitrate	240p: 75-150 kbps 360p: 220-450 kbps 480p: 375-750 kbps 720p: 1050-2100 kbps 1080p: 1875-12500 kbps

Experimental Results

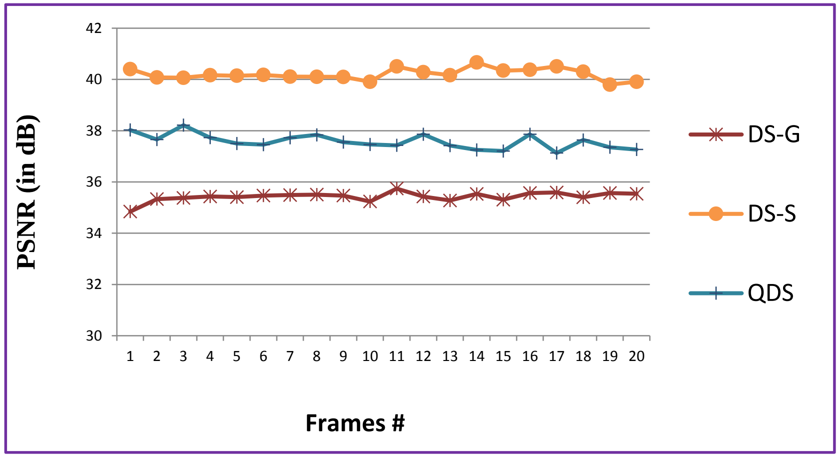


Fig.1. The PSNR measurement

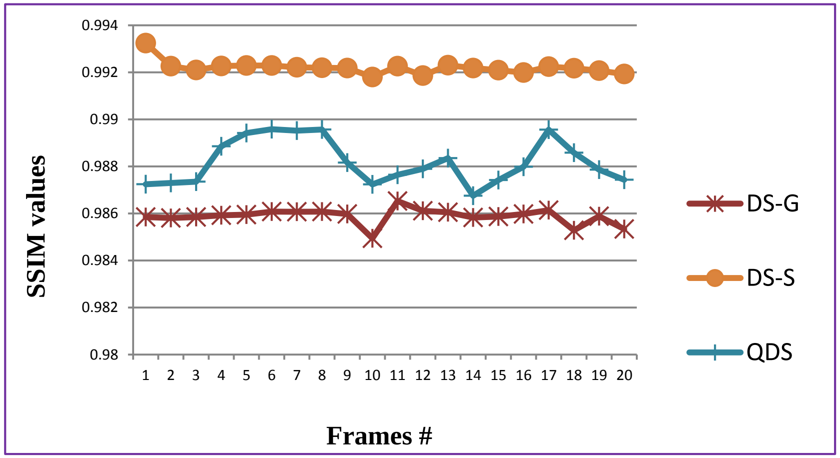


Fig.2. The SSIM index

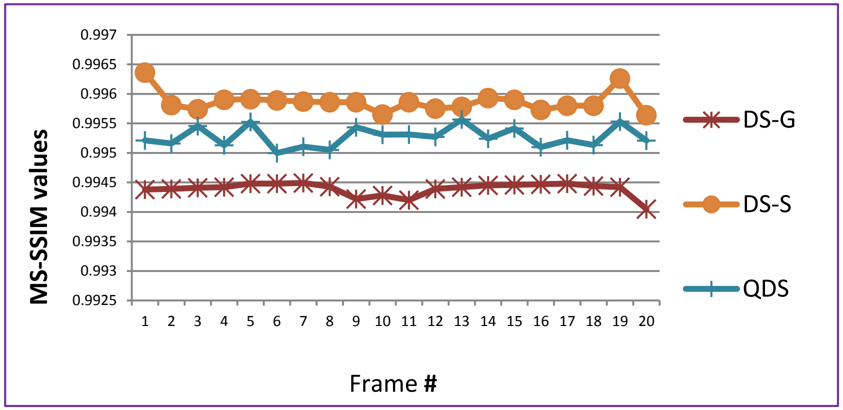


Fig.3. The MS-SSIM index

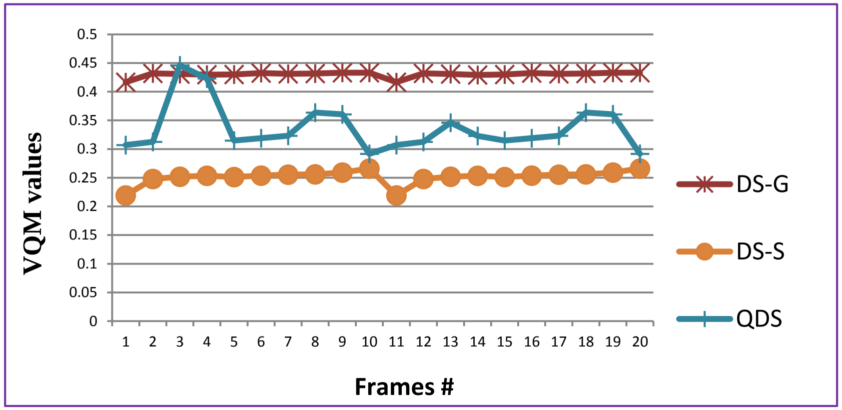


Fig.4. The VQM Observation

Experimental Results cont.

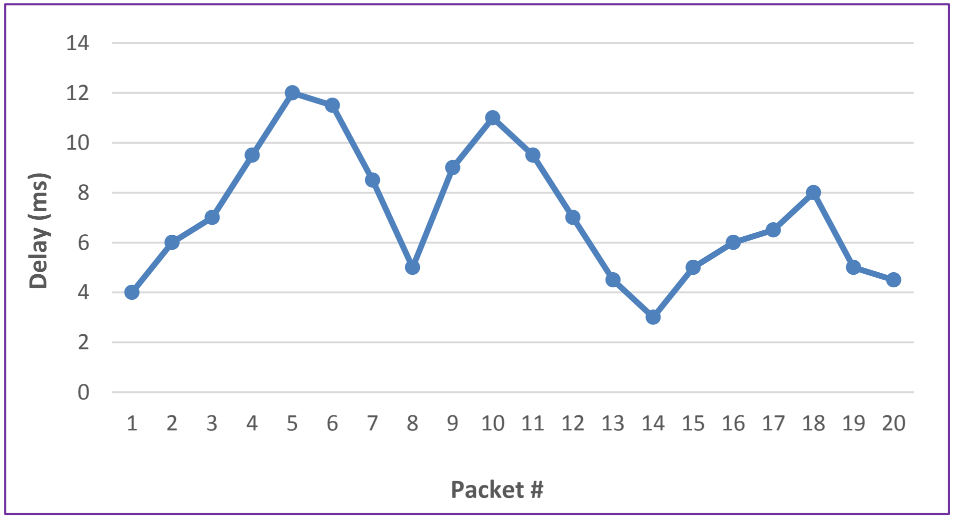


Fig.5. Inter packet delay

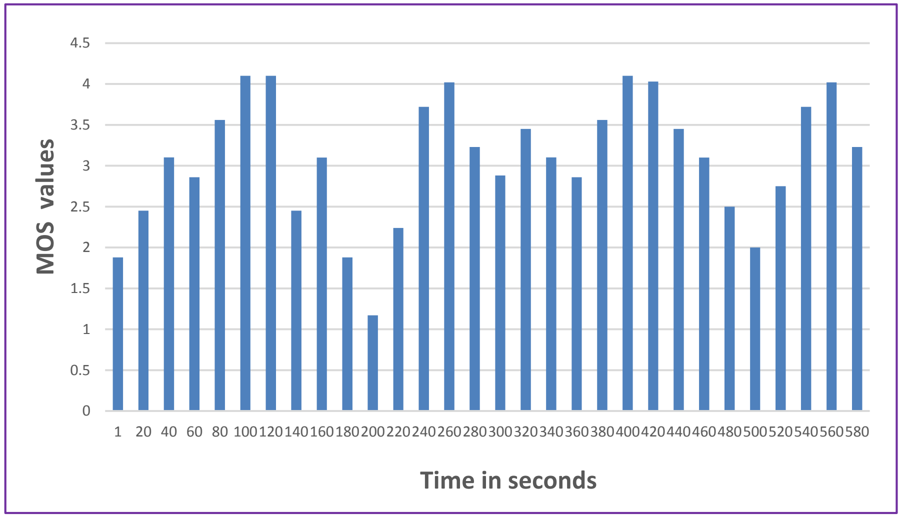


Fig.6. MOS values

Sample Snapshots

Fig.1a. Live Stream:
Original Frames

Fig.1b. Live Stream:
Decoded Frames

Fig.2a. *Stored Video Clipping of "Big Buck Bunny"* :
Original Frames

Fig.2b. *Stored Video Clipping of "Big Buck Bunny"* :
Decoded Frames

ITU KALEIDOSCOPE
SANTA FE 2018

Thank you

Acknowledgements

The University Grants Commission (UGC), New Delhi, India

Anna University, Chennai, India