

《DRL360: 360-degree Video Streaming with Deep Reinforcement Learning》

CUC - 19 数媒技 杨雪婷

2021.05.24

汇报内容介绍

1. 背景
2. 结构
3. 实验
4. 评估

Part One : 背景

问题：难以确定哪些tile应该被分配高、低码率

- 动态变化的大量的相关特征（输入向量大）
- 目前大多数是启发式的方法，需要用不同的方式来优化各种QoE指标

DRL-360:

根据客户端视频播放器收集到的数据，**自适应**地为未来视频帧的贴图**分配码率**。(将QoE指标组成一个动态特征组)

分配码率的两点原则：

- 视口外的贴图应该低于视口内的贴图
- 视口内的贴图应该保持一样的码率，防止有边界感

Part Two : 结构

整体结构：

- LSTM: predict bandwidth & viewport, 高精度的带宽预测和基于RNN的视点预测
learn the long-term QoE reward
- DRL: 揭示特征与QoE目标之间的关系, LSTM 和 Actor-Critic 优化QoE目标来适应网络条件和视口的未来变化

1、Features

- *Viewpoint and viewport*: in viewport $\rightarrow V = 1$; or $V = 0$
- *Video-content independent features*: 比如视口内的总速率
- *Bandwidth*: 计算下载两个连续块的间隔时间
- *Buffer Occupancy*

2、QoE metrics

- *Average Viewport Quality [Mb]*
- *Rebuffer*
- *Average Viewport Temporal Variations*

将这三点加权求和后得到QoE指标

3、基本Definitions:

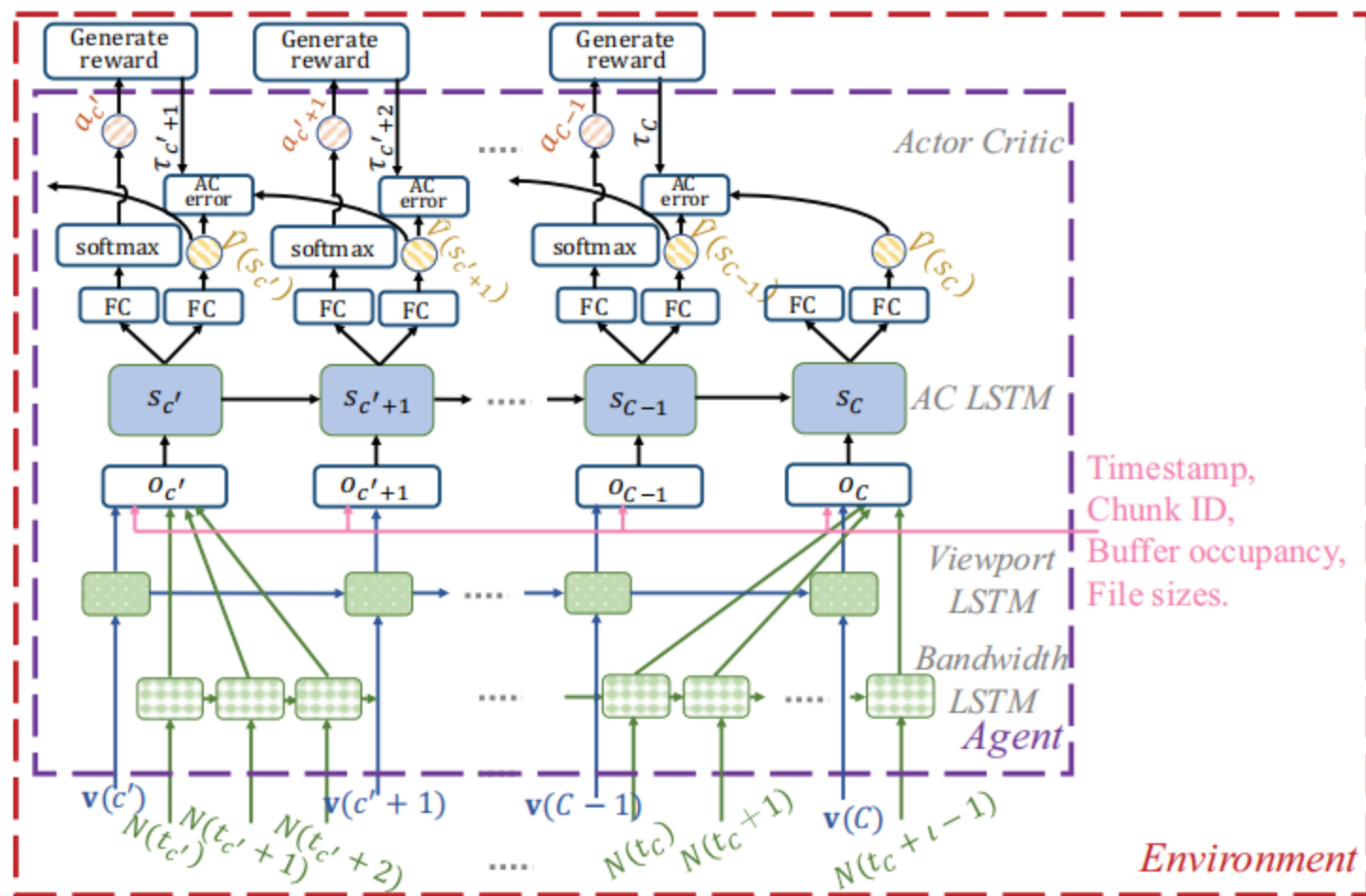
State: chunk ID-- C , viewport-- $v(c)$, bandwidth-- $N(tc)$, Buffer occupancy-- B_c , 下一个chunk的贴图大小-- $dc, ij(.)$, 预测的viewport-- $V'(c+1)$, 预测的带宽-- $N'(t+*)$

Action: 基于当前state, 视口内的第 $c+1$ 块chunk的比特率选择

Reward: 客户端记录每个块的视图中的平均速率, 并监控重缓冲时间, QoE指标

Policy: 神经网络 -> 最大化累计折扣奖励的期望值。 $\pi_{\theta}(s_c, a_c)$ 是给定当前状态 s_c 时, 动作 a_c 被采取的概率。算出 V 后可以确定有最大值的动作和选择的chunk的码率。

4、算法流程：



当请求一个新chunk时：

1. 环境将分别向预测LSTM模型提供以前的带宽记录和视口观察。
2. LSTM模型生成未来的viewport和bandwidth值。
3. 结合环境得到当前的observation经过AC LSTM后得到输出。
4. 输出经过两个FC神经网络更新state并做出action
5. 产生的动作会影响环境，而环境会带回奖励和新的观察结果，这些可以用来更新训练过程中的参数

Part Three : 实验

实验参数：

视频信息：

16个快、慢节奏的视频，25帧级，48条viewport轨迹。

两个video是测试集，其余是训练集。

码率：5种，分别为：1Mbps (360p), 5Mbps (720p), 8Mbps (1080p), 16Mbps (2K) and 35Mbps (4K)

segment 长度：1 s

tile in chunk：4 * 6

带宽信息：

2Mbps to 15Mbps, 16条, 每条至少400s

算法比较：

将三种算法进行比较：Action-Critic, Reinforce, Reinforce-Baseline

我们比较了提议的DRL360系统与六种最先进的360度视频流媒体系统的性能。

QoE指标：四种不同偏好的QoE指标。

Part Four : 评估

比较预测bw和vp的精度

比较预测模型的性能：LSTM, LR, CNN, KNN, 检查未来三个连续贴图的平均精度。
(带宽是检查未来三秒的带宽)

论文中都是CDF曲线 (分布函数)

RL算法的性能评估

优点：更看重某个QoE指标时可以更改权重参数