Neural Adaptive Video Streaming with Pensieve

Hongzi Mao, Ravi Netravali, Mohammad Alizadeh, SIGCOMM ’17

Pensieve is a machine-learning approach to generating adaptive bitrate (ABR) algorithms for optimizing video playback quality that is reflective of the network deployment environment. Specifically, Pensieve uses reinforcement learning (RL), automatically creating an ABR algorithm using neural networks and observations about network performance and quality of experience (QoE) measurements. Existing approaches to ABR algorithms make assumptions using rate-based (network throughput estimation) or buffer-based (playback buffer occupancy measurements) approaches, which Pensieve claims gives a limited and even inaccurate model of the deployment environment. Instead, Pensieve uses a neural network (NN) to train the ABR algorithm to predict the best bitrate for *future* video chunks using the following state inputs: past throughput measurements (x), past download times (tau), available sizes for next chunk (n), current buffer level (b), and then number of chunks remaining in video (c) and bitrate of last chunk (l). The NN chooses an action to take based on these input states, and then returns the observed quality of experience metrics as the reward value. The RL agent then continues to try to maximize the expected cumulative reward (i.e. the quality of experience metric) using policy gradient training.

In our class discussion, there were a few major points.

First, there was a good deal of discussion on online training versus offline training for RL. Pensieve was trained entirely offline using a simulated environment and unmodified after deployment. This raised the question whether Pensieve's results met their claim to be truly reflective of the environment without making assumptions of the environment like other algorithms. The point was raised that true online training is difficult to deploy for most systems, unless the problem task is quite simple. In the discussion section of the paper, the authors note that while actual online training would be extreme and perhaps unnecessary, their work could be extended to allow periodic updates.

There was also discussion on the RL methods chosen, as well as the training and testing methodology. Pensieve's choice of neural network implementation was questioned, since overly complex neural networks are more susceptible to overfitting, presenting better accuracy results initially during training but performing poorly in deployment. Additionally, the authors seemed to almost arbitrarily choose the A3C algorithm for training the RL agent. They justify their choice by pointing out that A3C has been successfully used in other learning scenarios, but this is a generic statement. Additionally, they write that A3C supports online training because of its asynchronous parallel training capabilities. However, designing Pensieve for online training is not a focus of the paper. In order to justify the algorithm choices, the authors could have trained and tested variants of Pensieve in comparison to other existing ABR algorithms.

According to the authors, Pensieve outperforms the other existing state-of-the-art ABR algorithms they tested, with average quality of experience improvements of 12%-25%. However, we noted that other studies have outperformed Pensieve and that Pensieve's claims of algorithm generalizability have not been fulfilled.