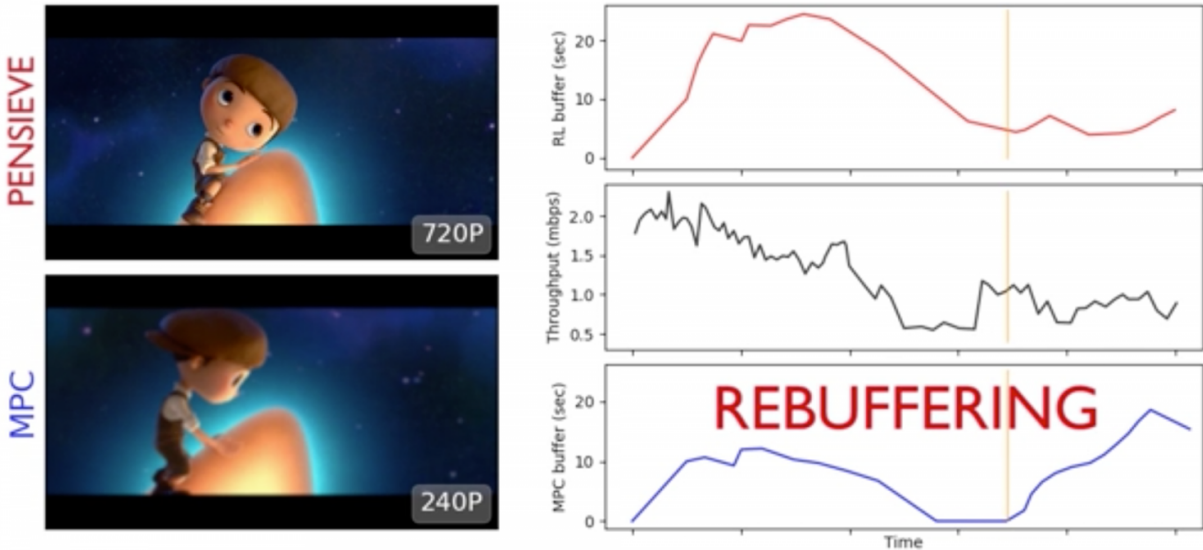


Machine-learning system enables smoother streaming that can better adapt to different network conditions

August 14 2017, by Adam Conner-Simons



In experiments, Pensieve could stream video with 10 to 30 percent less rebuffering than other approaches, and at levels that users rated 10 to 25 percent higher on key “quality of experience” metrics. Credit: Tom Buehler/MIT CSAIL

We've all experienced two hugely frustrating things on YouTube: our video either suddenly gets pixelated, or it stops entirely to rebuffer.

Both happen because of special algorithms that break videos into small chunks that load as you go. If your internet is slow, YouTube might make the next few seconds of video lower resolution to make sure you can still watch uninterrupted—hence, the pixelation. If you try to skip ahead to a part of the video that hasn't loaded yet, your video has to stall in order to buffer that part.

YouTube uses these adaptive bitrate (ABR) algorithms to try to give users a more consistent viewing experience. They also save bandwidth: People usually don't watch videos all the way through, and so, with literally 1 billion hours of video streamed every day, it would be a big waste of resources to buffer thousands of long videos for all users at all times.

While ABR algorithms have generally gotten the job done, viewer expectations for streaming video keep inflating, and often aren't met when sites like Netflix and YouTube have to make imperfect trade-offs between things like the quality of the video versus how often it has to rebuffer.

"Studies show that users abandon video sessions if the quality is too low, leading to major losses in ad revenue for [content providers](#)," says MIT Professor Mohammad Alizadeh. "Sites constantly have to be looking for new ways to innovate."

Along those lines, Alizadeh and his team at MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) have developed "Pensieve," an artificial intelligence (AI) system that uses machine learning to pick different algorithms depending on [network](#) conditions. In doing so, it has been shown to deliver a higher-quality streaming experience with less

rebuffering than existing systems.

Specifically, in experiments the team found that Pensieve could stream video with 10 to 30 percent less rebuffering than other approaches, and at levels that users rated 10 to 25 percent higher on key "quality of experience" (QoE) metrics.

Pensieve can also be customized based on a content provider's priorities. For example, if a user on a subway is about to enter a dead zone, YouTube could turn down the bitrate so that it can load enough of the video that it won't have to rebuffer during the loss of network.

"Our system is flexible for whatever you want to optimize it for," says PhD student Hongzi Mao, who was lead author on a related paper with Alizadeh and PhD student Ravi Netravali. "You could even imagine a user personalizing their own streaming experience based on whether they want to prioritize rebuffering versus resolution."

The paper will be presented at next week's SIGCOMM conference in Los Angeles. The team will also be open-sourcing the code for the project.

How adaptive bitrate works

Broadly speaking, there are two kinds of ABR algorithms: rate-based ones that measure how fast networks transmit data, and buffer-based ones that ensure that there's always a certain amount of future video that's already been buffered.

Both types are limited by the simple fact that they aren't using information about both rate and buffering. As a result, these algorithms often make poor bitrate decisions and require careful hand-tuning by human experts to adapt to different network conditions.

Researchers have also tried to combine the two methods: A [system out of Carnegie Mellon University](#) outperforms both schemes using "model predictive control" (MPC), an approach that aims to optimize decisions by predicting how conditions will evolve over time. This is a major improvement, but still has the problem that factors like network speed can be hard to model.

"Modeling network dynamics is difficult, and with an approach like MPC you're ultimately only going to be as good as your model," say Alizadeh.

Pensieve doesn't need a model or any existing assumptions about things like network speed. It represents an ABR algorithm as a neural network and repeatedly tests it in situations that have a wide range of buffering and network speed conditions.

The system tunes its algorithms through a system of rewards and penalties. For example, it might get a reward anytime it delivers a buffer-free, high-resolution experience, but a penalty if it has to rebuffer.

"It learns how different strategies impact performance, and, by looking at actual past performance, it can improve its decision-making policies in a much more robust way," says Mao, who was lead author on the new paper.

Content providers like YouTube could customize Pensieve's reward system based on which metrics they want to prioritize for users. For example, studies show that viewers are more accepting of rebuffering early in the video than later, so the algorithm could be tweaked to give a larger penalty for rebuffering over time.

Melding machine learning with deep-learning

techniques

The team tested Pensieve in several settings, including using Wifi at a cafe and an LTE network while walking down the street. Experiments showed that Pensieve could achieve the same video resolution as MPC, but with a reduction of 10 to 30 percent in the amount of rebuffering.

"Prior approaches tried to use control logic that is based on the intuition of human experts," says Vyaz Sekar, an assistant professor of electrical and computer engineering at Carnegie Mellon University who was not involved in the research. "This work shows the early promise of a machine-learned approach that leverages new 'deep learning'-like techniques."

Mao says that the team's experiments indicate that Pensieve will work well even in situations it hasn't seen before.

"When we tested Pensieve in a 'boot camp' setting with synthetic data, it figured out ABR algorithms that were robust enough for real networks," says Mao. "This sort of stress test shows that it can generalize well for new scenarios out in the real world."

Alizadeh also notes that Pensieve was trained on just a month's worth of downloaded video. If the team had data at the scale of what Netflix or YouTube has, he says that he'd expect their performance improvements to be even more significant.

As a next project his team will be working to test Pensieve on virtual-reality (VR) [video](#).

"The bitrates you need for 4K-quality VR can easily top hundreds of megabits per second, which today's networks simply can't support," Alizadeh says. "We're excited to see what systems like Pensieve can do

for things like VR. This is really just the first step in seeing what we can do."

More information: Paper: Neural Adaptive Video Streaming with Pensieve: [people.csail.mit.edu/hongzi/co ... nsieve-Sigcomm17.pdf](https://people.csail.mit.edu/hongzi/co...nsieve-Sigcomm17.pdf)

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Citation: Machine-learning system enables smoother streaming that can better adapt to different network conditions (2017, August 14) retrieved 7 September 2024 from <https://techxplore.com/news/2017-08-machine-learning-enables-smoother-streaming-network.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.