# Alloy: A Fusing Algorithm for Adaptive Video Streaming with Meta-Reinforcement Learning

Wangyu Choi and Jongwon Yoon*
Hanyang University

## Abstract

In this work, we present Alloy, an ABR algorithm that fuses the exiting algorithms to cover a wide variety of environments. We train Alloy in various simulation-based environments to obtain a transferable model that combines the strengths of the exiting algorithms via meta-reinforcement learning. At the end of simulation-based training, Alloy adapts well to new/various environments with a few update steps and provides high QoE to all users.

## Ⅰ. Introduction

In this work, we propose Alloy, an ABR algorithm that fuses existing algorithms to cover a wide variety of environments. Alloy provides high QoE to all users by combining the strengths of various algorithms. To do this, we find transferable parameters between multiple algorithms via meta-updates. Finally, with transferable parameters, Alloy covers a wide range of environments, providing different policies in different environments with just a few update steps.

## Ⅱ. System Design

**Step 1: configuring the environment.** We create various and vast simulation-based environments before training our model. First, we enumerate the properties that make up video streaming environment. Then we configure the environments by sampling the values of their characteristics. We use the following properties to configure the environments: network bandwidth, video, and client's configuration (buffer size). Finally, each property is combined to form a different video streaming environment.

**Step 2: finding transferable parameter.** Now, we find a parameter Θ that allows fast adaptation to the new environments by training in different environments. We adopt MAML [1] algorithm that combines the models. Our training process consists of two phases. An internal adaptation phase aims to compute the updated parameter $\theta_i'$ for the task $\mathcal{T}_i$ sampled from the distribution for the tasks $p(\mathcal{T}_i)$. For an internal update, we adopt PPO [2], a state-of-the-art policy gradient algorithm. In the external adaptation phase, the parameter $\theta$ is updated according to the training trajectories of the PPO agent. Both phases alternate to update the parameter $\theta$ that produces the highest reward after the PPO agent is trained on the $N$ update steps.

**Step 3: adapting to new real-world environments.** After obtaining the transferable parameter $\theta$, the model is easily adapted to the new environment with only $N$ update steps after deployment. The training agent updates $\theta$ to $\theta^*$ to adapt to unexperienced environments in just $N$-step updates.

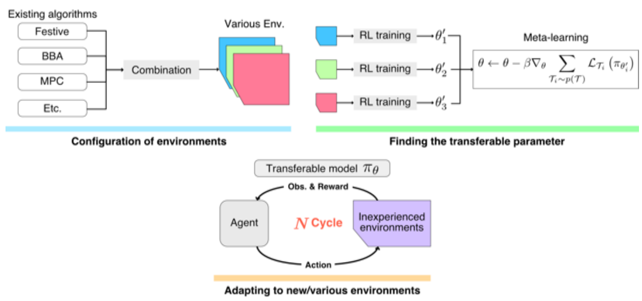


*Figure 1: Overview of the proposed system, Alloy.*

## Ⅲ. Evaluation

We implement Alloy using MPEG-DASH and create the following the scenario for evaluation.

**Coping with new environment and video.** Figure 2 shows the average QoE per segment using public Wi-Fi (coffee shop and campus), and Bug Buck Bunny video (i.e., new environments). We use BBA [3], RobustMPC [4], and AccurateMPC as state-of-the-art algorithms for comparison. AccurateMPC is a variant of RobustMPC with accurate future bandwidth.

We can clearly see that Alloy can provide better streaming services in highly variable Wi-Fi environments where bandwidth drops periodically and suddenly.

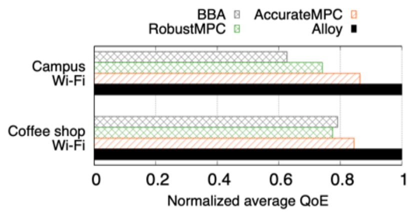


*Figure 2: Alloy outperforms state-of-the-art algorithms.*

## ACKNOWLEDGMENT

This work was supported by Basic Science Research Program through the National Research Foundation of South Korea (NRF) funded by the Ministry of Education NRF-2022R1A2C1008743 and the MSIT(Ministry of Science and ICT), Korea, under the Grand Information & Communication Technology Research Center support program (IITP-2022-2020-0-101741) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

## REFERENCES

[1] Finn et al., “Model-agnostic meta-learning for fast adaptation of deep networks,” ICML (2017).

[2] Schulman et al., “Proximal Policy Optimization Algorithms,” CoRR abs/1707.06347 (2017).

[3] Huang et al, “A Buffer-Based Approach to Rate Adaptation: Evidence from a Large Video Streaming Service,” SIGCOMM (2014).

[4] Yin et al., “A Control-Theoretic Approach for Dynamic Adaptive Video Streaming over HTTP,” SIGCOMM (2015).