Paper Summary
Genet: Automatic Curriculum Generation for Learning Adaptation in Networking

Sarthak | 2020CS10379

June 2024

Problem Statement

The paper addresses the challenge of training deep reinforcement learning (RL) models for network adaptation via curriculum learning and the problem of generation of a training curriculum that gradually increases the difficulty level of training environments.

Motivation

Traditional RL-based techniques in Networking face two main challenges :

- The asymptotic performance of the learned policies can be sub-optimal, especially when they are trained over a wide range of environments
- The trained RL policies may generalize poorly to unseen network environments

Key Idea

The core idea is to use curriculum learning to enhance RL-based network adaptation. Curriculum learning involves gradually increasing the difficulty of training environments. The paper introduces Genet, a framework that identifies rewarding environments where the RL model performs significantly worse than rule-based baselines, and focuses training on these environments.

Framework

The framework uses the performance gap between RL models and traditional rule-based algorithms to identify environments that can provide the most significant improvements. Key Points :

- Each round of Genet starts with training the current model for a fixed number of iterations
- After some iterations, the current RL model and a rule-based baseline are sent to a sequencing module to search for the environments where the current RL model has a large gap to baseline
- The expected gap-to-baseline over the environments created by configuration p : $Gap(p) = R(\pi^{rule}, p) - R(\pi_\theta^{rl}, p)$, where $R(\pi, p)$ is the average reward of a policy π over k environments. Since testing on all possible environments is expensive, Bayesian Optimization(BO) is used to search in the environment space for the configuration that maximizes $Gap(p)$
- RL model is updated by incorporating these prioritized environments into the training curriculum
- Genet interacts with an existing RL training codebase with two APIs, **Train** signals the RL to continue the training using the given distribution and returns a snapshot of the model after a few training iterations; **Test** calculates the average reward of a given algorithm over a specified number of environments drawn from the given distribution of configurations

Contributions

- Integrates curriculum learning into RL training for networking to address performance issues in varied environments
- Developed a novel method for automatically selecting training environments based on performance gaps between RL models and rule-based baselines (environments where the RL model shows significant room for improvement)
- Demonstrated Genet's effectiveness through case studies on adaptive video streaming, congestion control, and load balancing

Strengths

- Addresses performance degradation in RL models trained over broad ranges of network conditions
- Improves the ability of RL models to perform well in both seen and unseen environments
- Genet-trained RL policies have a much higher chance of out-performing various rule-based baselines specified during Genet-based RL training

Weaknesses

- The necessity of a good rule-based baseline can be a limitation in scenarios where such baselines are not well-established or easy to define
- Genet-trained RL policies may also achieve undesirable performance in environments beyond the training ranges, and does not guarantee adversarial robustness
- A small gap-to-baseline might also arise when the rule-based baseline has poor performance yet the RL model still has a large room for improvement. Genet ignores this case.