

# MARLIN: Soft Actor-Critic based Reinforcement Learning for Congestion Control in Real Networks

Raffaele Galliera<sup>†,\*</sup>, Alessandro Morelli<sup>†</sup>, Roberto Fronteddu<sup>†</sup>, Niranjan Suri<sup>†,\*,‡</sup>

<sup>†</sup>*Florida Institute for Human & Machine Cognition (IHMC)*

<sup>\*</sup>*Department of Intelligent Systems & Robotics - The University of West Florida (UWF)*  
Pensacola, FL, USA

<sup>‡</sup>*US Army Research Laboratory (ARL)*

Adelphi, MD, USA

{rgalliera, amorelli, rfronteddu, nsuri}@ihmc.org

**Abstract**—Fast and efficient transport protocols are the foundation of an increasingly distributed world. The burden of continuously delivering improved communication performance to support next-generation applications and services, combined with the increasing heterogeneity of systems and network technologies, has promoted the design of Congestion Control (CC) algorithms that perform well under specific environments. The challenge of designing a generic CC algorithm that can adapt to a broad range of scenarios is still an open research question. To tackle this challenge, we propose to apply a novel Reinforcement Learning (RL) approach. Our solution, MARLIN, uses the Soft Actor-Critic algorithm to maximize both entropy and return and models the learning process as an infinite-horizon task. We trained MARLIN on a real network with varying background traffic patterns to overcome the sim-to-real mismatch that researchers have encountered when applying RL to CC. We evaluated our solution on the task of file transfer and compared it to TCP Cubic. While further research is required, results have shown that MARLIN can achieve comparable results to TCP with little hyperparameter tuning, in a task significantly different from its training setting. Therefore, we believe that our work represents a promising first step towards building CC algorithms based on the maximum entropy RL framework.

**Index Terms**—Computer Networks, Communications Protocol, Machine Learning, Congestion Control, Reinforcement Learning, Soft Actor-Critic.

## I. INTRODUCTION

Network communications are the backbone of an ever-increasingly distributed digital world. Each day, technologies such as software, platform, and infrastructure as a service over Cloud/Edge/Fog computing systems, the Internet of Things, 5G networks, space communications and networks, peer-to-peer networks, blockchain, ad-hoc and mesh networks enable countless organizations and people to run their business and perform tasks that have become part of our daily routine with just a tap on a screen. The different characteristics of those technologies, combined with the requirements of very diverse services built on top of them, present unique challenges for network protocol designers and researchers.

Transport protocols that can achieve high network resource utilization while maintaining congestion, i.e., end-to-end queueing delays, low in presence of ever-changing network conditions and high churn of connections are the key enabling technology that underlies modern distributed

systems. Within transport protocol design, this dual goal (high channel utilization and low congestion) is the objective of Congestion Control (CC) algorithms. The critical impact of CC on the performance of distributed applications and services has led researchers and engineers to design many variants, each one with different target scenarios and trade-offs in terms of aggressiveness, responsiveness to loss and congestion, fairness, and friendliness [1, 2, 3]. However, the challenge of designing a generic CC algorithm that can provide near-optimal performance in a broad range of network and traffic scenarios is still open.

Recent advances in computational capacity, made possible by novel CPU, GPU, and Application-Specific Integrated Circuit (ASIC) architectures, along with the availability of low-cost, yet powerful, versions of such hardware accelerators, have led to an explosion of successful applications of Machine Learning (ML) technology in several domains, coming to meet the requirements of resource constrained environments such as the edge [4]. This has sprung renewed research interest towards the development of ML models, Reinforcement Learning (RL) agents, algorithms, and applications that can face the unique challenges of the real-world [5, 6, 7, 8].

The difficulty of designing an efficient generic CC algorithm, in contrast with the relative ease of collecting data on the communication performance and the small size of the action space, which typically focuses solely on adjusting the size of the Congestion Window (CWND), has prompted researchers to investigate the use of ML for CC optimization [9, 10]. While many studies have already shown promising results, they still underperform when compared to existing TCP CC algorithms in many scenarios. Reasons may include: difficulties to train RL agents and/or ML models using real networks and hardware; low fidelity of simulation and emulation environments; sub-optimal environment representation (state), action space, and/or reward function design; and delayed action effects on the state due to communications delay. Additional research is needed to address these problems.

This paper introduces our initial work on MARLIN, a RL agent for CC. Our approach is based on Soft Actor-Critic (SAC) [11], an off-policy, entropy-regularized RL algorithm, that has seen successful application in numerous real-world