Deep Reinforcement Learning for Communication Networks

Raffaele Galliera

Department of Intelligent Systems and Robotics - The University of West Florida (UWF)
The Institute for Human and Machine Cognition (IHMC)
Pensacola, FL, USA
rg101@students.uwf.edu

Abstract

This research explores optimizing communication tasks with (Multi-Agent) Reinforcement Learning (RL/MARL) in Point-to-Point and Group Communication (GC) networks. The study initially applied RL for Congestion Control in networks with dynamic link properties, yielding competitive results. Then, it focused on the challenge of effective message dissemination in GC networks, by framing a novel game-theoretic formulation and designing methods to solve the task based on MARL and Graph Convolution. Future research will deepen the exploration of MARL in GC. This will contribute to both academic knowledge and practical advancements in the next generation of communication protocols.

Introduction

The inherent need to share and receive information has always been a fundamental process in human societies, evolving from primitive acoustic and visual signals, such as drums and smoke, to the sophisticated digital networks of today. In modern communication systems, efficient and reliable group interaction is crucial to support critical operations across domains like disaster response and autonomous vehicles.

To ensure seamless data transfer, foundational components of Point-to-Point communication such as Congestion Control (CC), are experiencing a rapid evolution to respond to these new scenarios. However, these methods are usually self-centric and have significantly limited, or zero, visibility of third-party traffic sources present in the network. This lack of coordination often leads to the inefficient usage of networking resources, such as a shared bottleneck link.

In networking scenarios where collaboration is not negotiable, the methods used for point-to-point communication become inefficient or even deleterious. Dynamic networks such as Vehicular Ad-hoc Networks (VANETs) require constant coordination driven by the necessity to disseminate information about the nodes participating, e.g. identity and status, and crucial events happening in the network. Group Communication (GC) protocols are designed to enable such coordination and find a natural application in these modern networking scenarios. Nevertheless, these systems can be characterized by congestion-prone networks and/or different resource constraints with nodes observing partial segments of the entire network. Disseminating information that a group of nodes (or all nodes) must receive becomes considerably expensive if not adequately managed.

Research Questions

The world of communication protocols is dominated by (often distributed) heuristics to optimize networking tasks. In the case of GC protocols, **cooperation** between the nodes is often enabled by the exchange of "control messages" which are then exploited to perform decisions. These heuristics require careful tuning depending on the scenario, especially if deployed in networks where the topology or its properties drastically vary. Because of the sensitivity to their parameters and the networking scenario where they are deployed, I investigated how to learn multi-agent strategies by leveraging the underlying communication protocols to optimize communication tasks. The goal of my research can be summarized as "*how* to learn to communicate in GC tasks, determining *what* information should be communicated and *when* it should be communicated". My curiosity is driven by three main questions:

Q1) **How can we exploit the mechanisms underlying communication protocols to learn both what should be communicated and what actions should be taken to optimize a given task?**

In GC protocols, control messages can assume different forms, some might include local topological information, and others might even deliver "orders" to nodes regarding what to do in a certain scenario. These communication mechanisms can enable learned, distributed strategies.

Q2) **How can agents learn how to synchronize their behavior and their exchange of messages?**

Timing and synchronizing control messages is a crucial aspect of GC protocols. The frequency of their dispatching is often set as a parameter upon deployment.

Q3) **How can agents learn distributed strategies able to cope with environments characterized by different types of constraints?**

In varied networking environments, adaptability to resource constraints—be it bandwidth, storage, or computation—is essential.

Related Work

Recently, researchers have considered learning communication protocols (Foerster et al. 2016) with Multi-Agent Reinforcement Learning (MARL) to encourage the emergence of more efficient strategies. In cooperative tasks, agents interact within a shared environment and often find the need to exchange information to optimize their collective performance. This has led to the development of communication protocols that are learned rather than pre-defined.

Different contributions have focused on the communication of local encodings of the agents' observations. Methods such as CommNet (Sukhbaatar, Szlam, and Fergus 2016) allow agents to share a distilled version of their perspectives, enabling more informed collective decision-making. Approaches such as ATOC (Jiang and Lu 2018) have ventured into attention mechanisms to determine which agents to communicate with and what information to share. Yet another approach, as exemplified by Graph Convolutional Reinforcement Learning (Jiang et al. 2020), harnesses the power of Graph Neural Networks (GNNs) to model interactions, relations, and communications between the agents.

Current Progress

My research is characterized by a progressive exploration of the role of Reinfocement Learning (RL) in communication networks. First, I investigated the application of RL to address CC and the challenges of training and deploying RL agents in real networks, particularly, when third-party traffic sources of different natures are competing on the usage of a shared bottleneck link (Galliera et al. 2023a). I further refined the agent and its environment to more constrained network conditions, enhancing the adaptability of the agent, particularly during transitions between varying link properties (Galliera et al. 2023c). Transitioning from a single-agent perspective, I shifted into the realm of MARL, aiming to optimize message dissemination in broadcast networks (Galliera et al. 2023b). This contribution is characterized by a game-theoretic formulation of message dissemination based on communication mechanisms found in broadcast protocols like Optimized Link State Routing (OLSR) and it culminated in the development of two initial methods, namely L-DGN and HL-DGN, rooted in the basics of Graph Convolutional RL. I performed an extensive evaluation of the learned strategy on static graphs, demonstrating the capability of the agents to outperform a widely used heuristic.

Future Research Plan

Having laid down this groundwork, my research focus is now on pursuing the study of MARL for GC, investigating different aspects concerning the agents' communication and their learned strategies, designing new methods to improve cooperation, and integrating them in real GC protocols.

- **01-2024)** Provide richer representations of multi-agent systems in dynamic networks and develop novel methods to handle the complexities that such dynamism brings. For example, the observation of variable graph structures.
- **Mid-Late 2024)** Enable the agents to decide when to communicate information, such as their features and/or their latent representations, by enhancing their actions space and crafting novel learning algorithms to cope with the dualism of deciding *what* and *when* to communicate.
- **2025)** Focus on critical network properties such as bandwidth consumption and latency in real-world, resource-constrained, networking scenarios. In this regard, the initial research made on learning CC will be pivotal. Agents will need to jointly optimize their primary tasks and the utilization of shared resources. Evaluate the trained agents on real broadcast protocols.

Anticipated Thesis Contribution. The contributions of my thesis will constitute an initial step towards merging the world of GC protocols and MARL, providing a first model for designing MARL agents suitable to these environments. The integration of real-world evaluations will further solidify the practical relevance and applicability of the devised solutions, delivering a research compendium that not only enriches academic literature but also equips the industry with actionable insights and methodologies for the next generation of communication protocols.

References

Foerster, J. N.; Assael, Y. M.; de Freitas, N.; and Whiteson, S. 2016. Learning to Communicate with Deep Multi-Agent Reinforcement Learning. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, 2145–2153. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781510838819.

Galliera, R.; Morelli, A.; Fronteddu, R.; and Suri, N. 2023a. MARLIN: Soft Actor-Critic based Reinforcement Learning for Congestion Control in Real Networks. In *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, 1–10.

Galliera, R.; Venable, K. B.; Bassani, M.; and Suri, N. 2023b. Learning Collaborative Information Dissemination with Graph-based Multi-Agent Reinforcement Learning. arXiv:2308.16198.

Galliera, R.; Zaccarini, M.; Morelli, A.; Fronteddu, R.; Poltronieri, F.; Suri, N.; and Tortonesi, M. 2023c. Learning to Sail Dynamic Networks: The MARLIN Reinforcement Learning Framework for Congestion Control in Tactical Environments. In *(To Appear) MILCOM 2023 IEEE Military Communications Conference (MILCOM)*.

Jiang, J.; Dun, C.; Huang, T.; and Lu, Z. 2020. Graph Convolutional Reinforcement Learning. In *International Conference on Learning Representations*.

Jiang, J.; and Lu, Z. 2018. Learning Attentional Communication for Multi-Agent Cooperation. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, 7265–7275. Red Hook, NY, USA: Curran Associates Inc.

Sukhbaatar, S.; Szlam, A.; and Fergus, R. 2016. Learning Multiagent Communication with Backpropagation. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, 2252–2260. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781510838819.