Learning to Disseminate Information with Graph-based Multi-Agent Reinforcement Learning

Raffaele Galliera, Kristen Brent Venable, Matteo Bassani, Niranjan Suri

Department of Intelligent Systems and Robotics, The University of West Florida (UWF)
The Institute for Human and Machine Cognition (IHMC)

rg101@students.uwf.edu, {bvenable, nsuri}@uwf.edu,
{rgalliera, bvenable, mbassani, nsuri}@ihmc.org

Abstract

Decentralized, efficient, and collaborative information dissemination is a critical aspect for effective applications in disaster response and autonomous vehicles. Here, we focus on **Multi-Agent Reinforcement Learning (MARL)** employing **Graph Attention Networks (GATs)** to learn dissemination strategies while being compatible with current standard broadcast protocols such as Optimized Link State Routing Protocol (OLSR).

Overview

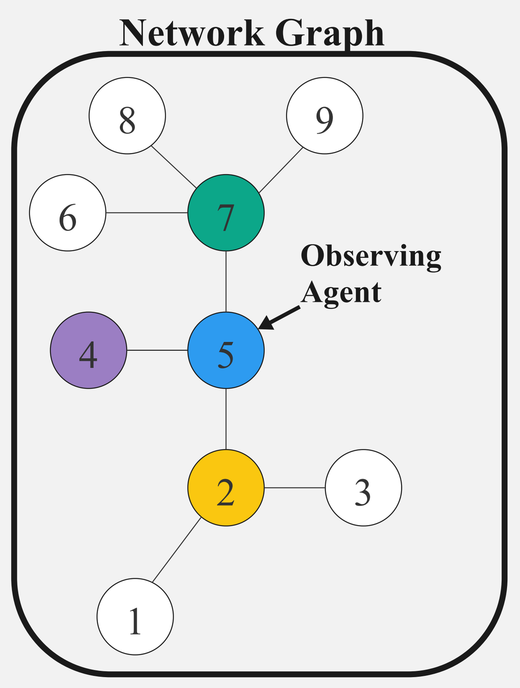
- **Current Limitations:** Standard Multi-Point Relay (MPR) solutions require careful parameter tuning.
- **Opportunities:** Broadcast protocols allow the exchange of control messages between neighboring nodes to enable cooperation.
- **Our idea:** Use MARL to learn dissemination strategies and Graph Neural Networks (GNNs) to exchange learned latent representations between the agents.

Key Contributions

Our key contributions include:

- Developing a **Partially Observable Stochastic Game** (POSG) formulation for information dissemination with reduced 2-hop knowledge compared to MPR.
- Proposing **two methods** based on Graph Convolutional Reinforcement Learning characterized by different levels of communication overhead.
- Evaluating our methods, demonstrating efficiency in network coverage and message optimization when **compared to a well-known heuristic** (MPR) and a MARL baseline (DGN-R).

Network Example



The observing node (5) is not aware of its two-hop neighborhood structure, which is a more constrained observation compared to standard MPR heuristics.

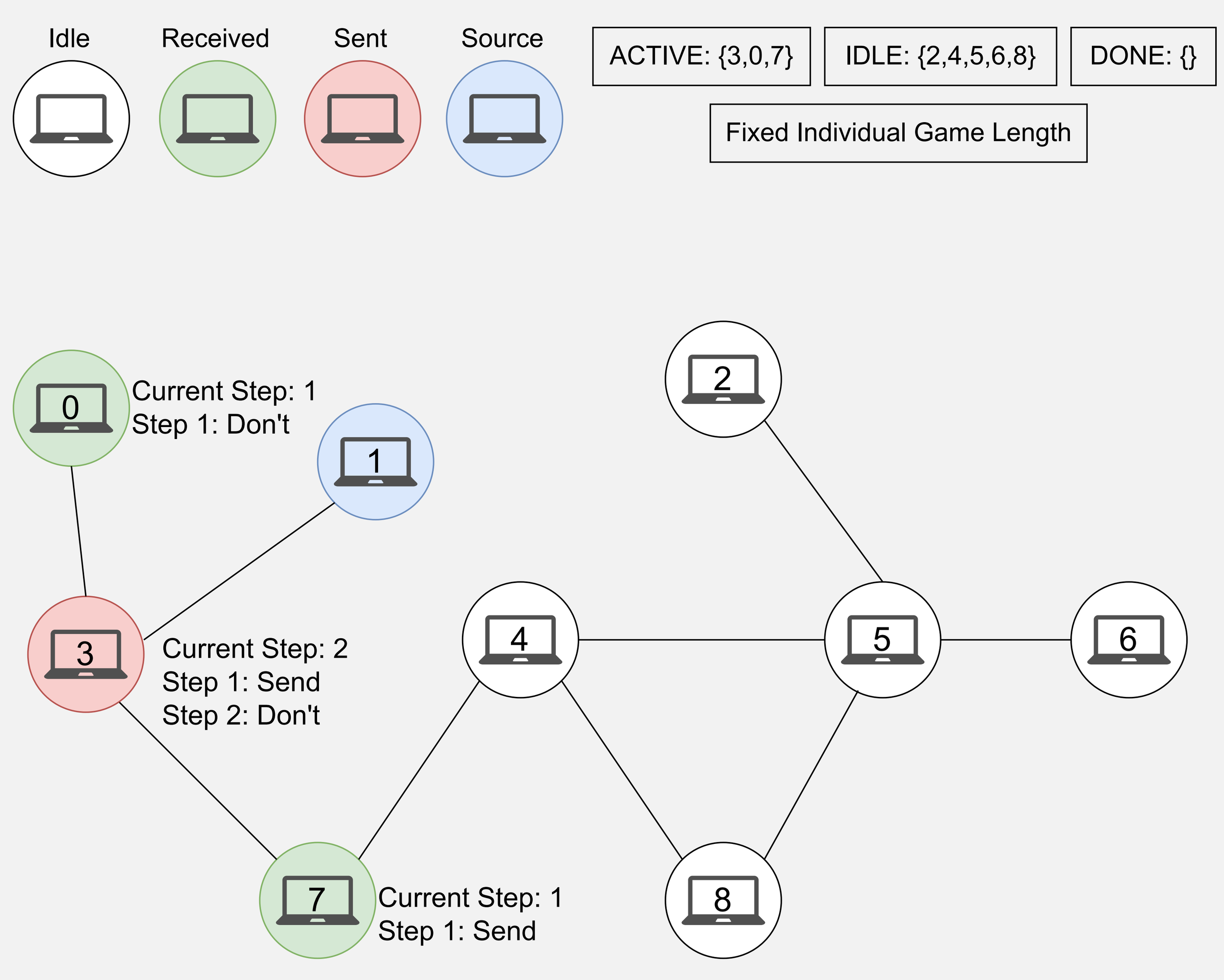
Method

- We encapsulate the participation of a node in the dissemination process within a **local horizon**, limiting the active involvement to a defined number of steps.
- **Observation:** Agents identify one-hop neighbors and their degree (i.e. number of neighbors), aligning with a constrained observation space compared to the Neighborhood Discovery Protocol in OLSR.
- **Binary actions: to forward the message or not.**
- The **reward** mechanism for agents encourages 2-hop coverage efficiency, incorporating penalties based on forwarding behavior and the unexploited coverage potential. The reward function is defined as:

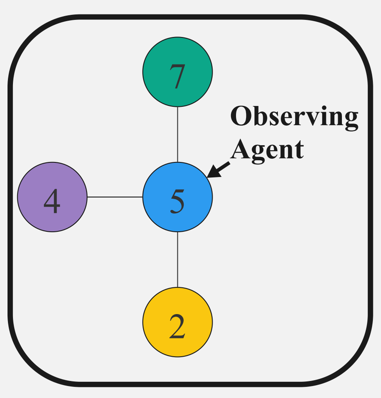
$$r_{i,t} = \frac{\upsilon(\mathcal{M}_i, t)}{|\mathcal{M}_i|} - p(i,t),$$

where $\upsilon(\mathcal{M}_i, t)$ accounts for the two-hop coverage of agent i, and $p(i,t)$ is a penalty function, depending on the agent's actions:
  - If agent i has forwarded the message, a *neighborhood shared transmission cost* is applied, based on the total number of messages sent within its neighborhood.
  - If agent i has not forwarded the message, a *coverage potential* penalty is applied, reflecting the unexploited opportunity to reach uncovered neighbors.

Dynamic Agents Participation Based on Message Reception

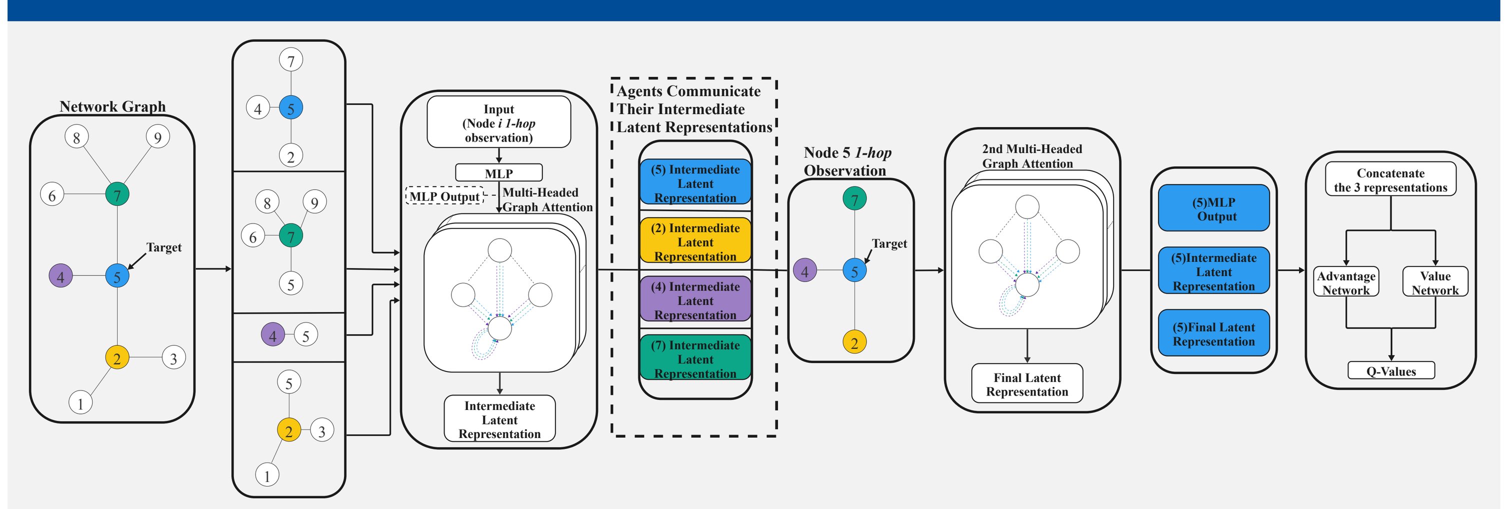


Observation Example with Node Features

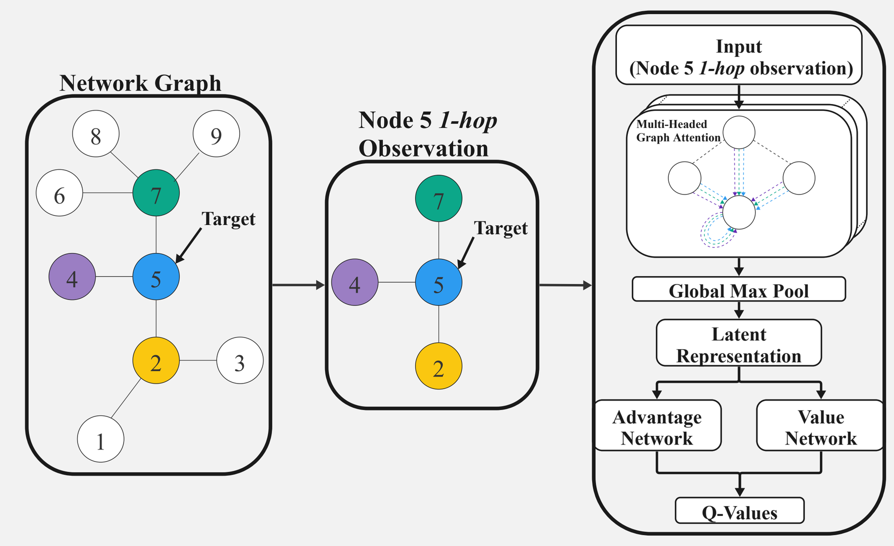


Node	# Neighbours	Data Messages	A_1	A_2	A_3	A_4
2	3	0	0	0	0	0
4	1	1	1	0	0	0
5	3	1	0	0	1	0
7	4	0	0	0	0	0

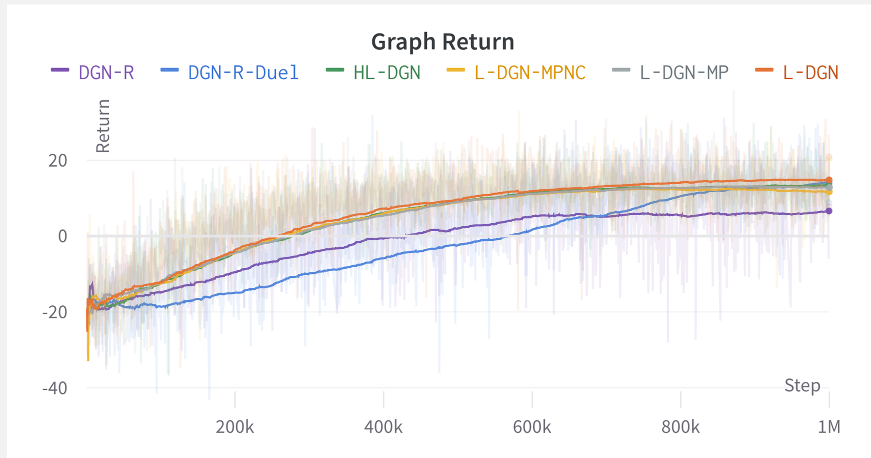
L-DGN Architecture



HL-DGN Architecture



Ablation Study



- *DGN-R-Duel*: Merges aspects of L-DGN and DGN-R with a dueling network.
- *L-DGN-MP*: Removes the second GAT layer from L-DGN, replacing it with a global max pool operator.
- *L-DGN-MPNC*: Eliminates both the second GAT layer and concatenation of encoding stages from L-DGN. Performance decrease noted, but serves as a basis for the simplified and efficient HL-DGN architecture.

Results

Nodes	Method	Coverage	Data Messages	Bootstrap Control Overhead	Two-Hop Anonymity
20					
	MPR	100%	12.05	60	No
	DGN-R	100%	21.06	60	Yes
	L-DGN	99.95%	11.84	60	Yes
	HL-DGN	100%	13.17	40	Yes
50					
	MPR	100%	30.8	150	No
	DGN-R	99.98%	60.65	150	Yes
	L-DGN	93.3%	25.42	150	Yes
	HL-DGN	100%	35.1	100	Yes

- **Bootstrap Control Overhead:** Number of control messages needed to be exchanged beforehand.
- **Data Messages:** Number of forwarding actions performed by the agents.

Conclusion

- Shown the **effectiveness of MARL in optimizing information dissemination in broadcast networks**, comparing our solutions with traditional heuristics and a MARL baseline on needed control overhead and data message efficiency.
- Our **future work** will explore dynamic graphs and reduce control overhead.
- **Extension of the MARL approach** to real-world network scenarios and broader domains like social networks and computational social choice.