

Graph Attention Reinforcement Learning for Structured Interaction in Cooperative Decision Networks

Mikkel Jensen¹, Anders Nielsen^{2*}, Rasmus Pedersen³

Department of Electrical Engineering, Technical University of Denmark (DTU), 2800 Kgs. Lyngby, Denmark

**Corresponding author: anders.nielsen@dtu.dk*

Abstract.

Modeling interaction structure is essential for scalable coordination in cooperative decision systems. This work explores a graph attention-based reinforcement learning approach in which agents are represented as nodes and communication dependencies are encoded as dynamic edges. Policy updates are performed using actor-critic learning combined with graph attention networks (GAT) to capture context-aware influence among agents. The model is evaluated on 8,500 collaborative decision instances derived from logistics routing and distributed resource allocation tasks. Compared with fully connected communication baselines, the proposed framework achieves a 21.7% improvement in task efficiency and reduces communication overhead by 18.3%. Moreover, performance gains remain stable as the number of agents increases from 5 to 20, demonstrating improved scalability. These results highlight the importance of structured interaction modeling in reinforcement-driven coordination systems.

Keywords:

Graph attention networks;
Reinforcement learning;
Multi-agent systems; Actor-critic; Coordination

INTRODUCTION

Cooperative decision networks are increasingly used in logistics routing, distributed resource allocation, and networked control, where multiple agents must coordinate under uncertainty and changing operating conditions. Reinforcement learning (RL) has attracted substantial attention in these settings because it can learn decision policies directly from interaction and adapt to dynamic environments without relying on hand-crafted coordination rules. Recent studies indicate that multi-agent reinforcement learning (MARL) can improve collaborative performance in transportation systems, communication networks, and resource management tasks, particularly when decision quality depends on continuous interaction among multiple agents [1]. This advantage is especially relevant in complex decision processes that require structured information exchange and adaptive inference over multiple steps, where efficient coordination depends not only on local action selection but also on how agents organize and interpret shared information [2]. Even so, existing MARL methods still face persistent difficulties in scalability, communication cost, and performance stability as the number of agents grows [3]. These limitations become more

serious in large cooperative systems, where dense interaction may overwhelm the learning process and weaken the practical value of coordination models. A major limitation of many early MARL approaches is their reliance on centralized training or shared global information. Although such designs often improve coordination during training, they usually assume that all agents can communicate with one another continuously. This assumption is difficult to satisfy in large-scale systems, where unrestricted communication increases transmission cost, introduces redundant information, and may even reduce learning efficiency by obscuring the most useful interaction signals [4,5]. As a result, recent work has shifted toward learning communication structure, in which agents decide what information should be exchanged and with whom, according to task requirements and local context [6,7]. This direction reflects an important change in perspective: communication is no longer treated as a fixed background mechanism, but as part of the decision process itself. For cooperative systems operating in uncertain and distributed environments, the ability to form selective interaction patterns is often as important as the ability to optimize actions. Graph-based methods provide a natural framework for modeling such structured interaction. In these approaches, agents are represented as nodes and their relationships are encoded as edges, allowing the model to focus on localized dependencies rather than unrestricted global communication. This representation is well suited to cooperative decision networks because real interaction structures are often sparse, time-varying, and dependent on task context. Recent studies show that graph neural networks can improve coordination efficiency while reducing unnecessary communication overhead [8,9]. Learnable communication graphs have been shown to outperform fixed interaction structures, while hierarchical graph-based designs can preserve coordination quality as the number of agents increases [10,11]. Related findings also suggest that graph-based MARL is particularly effective in large-scale systems where communication resources are limited and the relevance of neighboring agents changes over time [12]. These results support the view that structured interaction is not simply a computational convenience, but a key mechanism for scalable coordination. Attention mechanisms further strengthen graph-based coordination models by enabling agents to assign different weights to different neighbors. In cooperative tasks, the value of incoming information is rarely uniform. Some neighboring agents provide highly relevant signals for current decisions, whereas others contribute little and may even introduce noise. Attention-based communication helps the model distinguish critical interactions from less informative ones, thereby improving both efficiency and decision quality [13]. Recent work has shown that attention mechanisms are useful for handling partial observations, adaptive communication, and higher-level interaction patterns in cooperative multi-agent environments [14,15]. By combining graph structure with attention-based weighting, the model can capture both who should interact and how much each interaction matters under a given task state. This is particularly important for large decision networks, where the quality of coordination increasingly depends on selective information filtering rather than unrestricted message passing. Graph-enhanced reinforcement learning has already been applied to a range of practical problems, including vehicle routing, structured multi-agent coordination, robust decision-making under system changes, and large-scale resource allocation [16,17]. These studies demonstrate that structured interaction can improve coordination effectiveness in realistic application settings. However, most existing work remains focused on a single domain or a narrow class of tasks, which limits the broader understanding of how graph-based communication behaves across heterogeneous environments. In many cases, reported improvements are tied closely to a specific scenario design, making it difficult to determine whether the same interaction mechanism remains effective when task structure, agent density, or communication demand changes substantially. As

a result, the general value of graph-attention coordination for cooperative decision networks has not yet been fully established. Despite recent progress, several important gaps remain. Many existing methods still rely on fixed or densely connected communication structures, which do not match real systems where interactions are sparse, selective, and continuously changing. Experimental evaluation is also often confined to limited task settings, reducing confidence in cross-domain generalization. In addition, communication overhead and scalability are not always examined in sufficient detail, even though these factors directly affect whether a method can be deployed in larger cooperative systems. Another limitation is that many empirical studies involve only a small number of agents, leaving open the question of how coordination quality changes when system size increases and interaction patterns become more complex [18]. These issues indicate that a more adaptive and scalable communication framework is needed, one that can preserve coordination quality while controlling the cost of information exchange. This study addresses these issues by developing a graph attention reinforcement learning framework for structured interaction in cooperative decision networks. In the proposed framework, agents are modeled as nodes and their communication links are represented as dynamic edges that change with the task state. A graph attention network is integrated with an actor-critic learning scheme to capture context-dependent interaction among agents and to strengthen selective communication during decision-making. This design allows the model to focus on useful coordination signals while reducing unnecessary information exchange, thereby improving both efficiency and robustness in large multi-agent environments. The purpose of this study is to examine whether structured interaction can provide a more effective basis for scalable coordination in cooperative decision systems. The framework is evaluated on 8,500 decision tasks drawn from logistics routing and resource allocation domains, with particular attention to task efficiency, communication overhead, and performance stability as the number of agents increases. The significance of this work lies in showing how graph-based representation and attention-guided interaction can be combined within reinforcement learning to support more efficient coordination under sparse and dynamic communication conditions. By moving beyond dense communication assumptions and single-scenario validation, this study aims to provide stronger empirical evidence for scalable cooperative decision modeling and to contribute a practical interaction mechanism for reinforcement-driven multi-agent systems operating in complex real-world environments.

2. Materials and Methods

2.1. Samples and Study Scope

The experiments were carried out on 8,500 cooperative decision cases from two task domains: logistics routing and distributed resource allocation. Each case included a set of agents, a state description, an action space, and task constraints. In routing tasks, agents represented delivery units working on networks with changing demand, travel cost, and time limits. In resource allocation tasks, agents handled limited resources under competing requests and deadline constraints. The number of agents ranged from 5 to 20 so that model performance could be tested under different system sizes. All cases were generated in controlled simulation environments to cover different network forms, constraint levels, and interaction patterns. The full dataset was divided into training, validation, and test sets, and no case appeared in more than one set.

2.2. Experimental Design and Baseline Settings

The proposed method combined a graph attention network with an actor-critic reinforcement learning framework. Its performance was compared with two baseline methods. The first baseline used full communication, where each agent exchanged information with all other agents at every step. The second baseline used a fixed graph, where the communication links were set in advance

and did not change during training. All methods used the same network size, learning rate, and training length to keep the comparison fair. The experiments focused on three aspects: task efficiency, communication cost, and performance stability under different numbers of agents. Each setting was repeated with several random seeds, and the final results were averaged across runs.

2.3. Measurement and Quality Control

Model performance was evaluated with three measures. Task efficiency was calculated by the average reward in each episode. Communication cost was measured by the average number of active links used during one decision step. Scalability was examined by tracking how performance changed as the number of agents increased. To improve result reliability, each experiment was repeated five times with different initial settings. Mean values and standard deviations were reported for all measures. Unstable runs were screened with the interquartile range rule. Key hyperparameters, including learning rate, batch size, and discount factor, were selected on the validation set and then kept unchanged in all test experiments. Training quality was checked by monitoring reward curves and policy update behavior throughout the learning process.

2.4. Data Processing and Model Formulation

All input features were normalized before training so that variables from different tasks had a similar scale. Agent states were encoded as node features, and communication relations were written as an adjacency matrix. The graph attention layer assigned a weight to each connected neighbor according to

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T[W h_i \parallel W h_j]))}{\sum_{k \in N(i)} \exp(\text{LeakyReLU}(a^T[W h_i \parallel W h_k]))}$$

where h_i and h_j are the features of agents i and j , W is the projection matrix, a is a learnable vector, and $N(i)$ is the neighbor set of agent i . Policy learning followed the actor-critic objective

$$L_{\pi} = -E[\log \pi(a_t | s_t) \hat{A}_t]$$

where $\pi(a_t | s_t)$ is the policy output and $\hat{A}_t$ is the estimated advantage at step t . The value branch was trained by minimizing the mean squared error between predicted returns and target returns.

2.5. Implementation and Evaluation Procedure

The model was built on a graph-based actor-critic structure with shared parameters across agents. Training used mini-batch updates over multiple episodes. At each step, agents exchanged information through the attention layer, and the communication graph was updated according to the learned attention scores. In the test stage, the model was evaluated on unseen cases with different network structures and different numbers of agents. Performance was measured both before extra updates and after a small number of adaptation steps. All methods were trained and tested under the same hardware and software settings to avoid differences caused by the computing environment. Final results were reported as mean values with standard deviations over repeated runs.

3. Results and Discussion

3.1. Overall task performance

The proposed graph attention reinforcement learning method achieved higher performance than the baseline methods in both logistics routing and resource allocation tasks. Compared with the fully connected communication setting, task efficiency increased by 21.7%. This result shows that the model can focus on useful interactions instead of using all communication links. As shown in Fig.1, the proposed method kept better performance across both domains, which indicates that the improvement is stable rather than limited to a small number of cases. Previous studies also reported

that graph-based methods can improve coordination, but many of them were tested on simpler environments [19,20]. The present results extend this finding to more complex decision tasks.

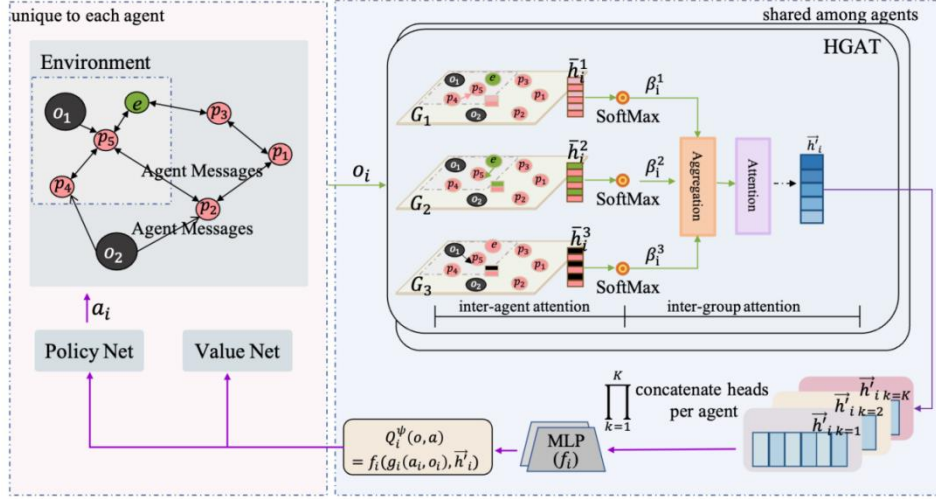


Figure 1 Performance comparison across methods, showing task efficiency in routing and resource allocation tasks.

3.2. Communication cost and efficiency

The proposed method also reduced communication cost by 18.3%. This means that fewer communication links were used while task performance still improved. In cooperative systems, full communication often leads to repeated or unnecessary information exchange. The graph attention mechanism reduces this problem by assigning higher weights to important neighbors and lower weights to less useful ones. As shown in Fig.2, this advantage becomes clearer as the system grows. This result supports the idea that structured and selective communication is more effective than full communication in large systems. Similar observations have been reported in recent graph-based reinforcement learning studies [21,22].

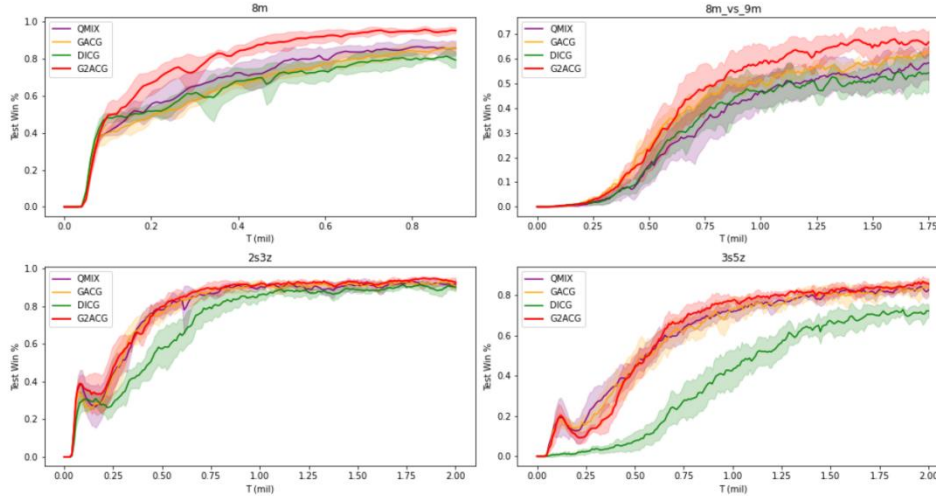


Figure 2 Comparison of communication cost, measured by the number of active links used during decision steps across different system sizes.

3.3. Scalability with increasing agents

The model remained stable when the number of agents increased from 5 to 20. The proposed method kept its performance level, while the fully connected baseline showed a clear decrease. This result indicates that structured interaction helps control the increase in communication

complexity. In full communication settings, the number of connections grows quickly with the number of agents, which can affect learning and decision quality. In contrast, the attention-based graph keeps only the most useful connections. Previous work also showed that graph-based methods can improve scalability, but most studies used smaller systems [23,24]. The current results show that this advantage remains in larger and more complex settings.

3.4. Comparison and practical implications

Compared with previous studies, this work has two main advantages. First, it tests the method on two different task domains, which provides a broader evaluation. Many earlier studies focus on a single type of problem, such as routing or simulation benchmarks. Second, this study examines both task performance and communication cost, which gives a more complete view of model behavior. The results suggest that graph attention is useful for improving both efficiency and scalability in cooperative systems. However, some limitations remain. The experiments were conducted in simulated environments, and all agents used the same model structure. Real systems may have more complex conditions, such as delayed communication or different agent roles. Future work should test the method in real-world settings and consider more flexible model designs.

4. Conclusion

This study investigated a graph attention reinforcement learning method for structured interaction in cooperative decision networks. The framework combined a graph attention mechanism with an actor-critic learning scheme to model how agents share information during decision making. The results showed higher task efficiency, lower communication cost, and stable performance as the number of agents increased. These findings suggest that selective interaction can improve coordination while reducing unnecessary communication. The main contribution is the use of a graph-based interaction model to support scalable learning in multi-agent systems, together with an evaluation on routing and resource allocation tasks. The method can be applied to systems where many agents operate with limited communication, such as logistics networks and distributed resource management. However, the experiments were conducted in simulated environments, and all agents used the same model structure, which may limit performance in more complex settings. Future work should include tests in real systems and explore model designs that can better handle heterogeneous agents and changing communication patterns.

References

- Louati, A., Louati, H., Kariri, E., Neifar, W., Hassan, M. K., Khairi, M. H., ... & El-Hoseny, H. M. (2024). Sustainable smart cities through multi-agent reinforcement learning-based cooperative autonomous vehicles. *Sustainability*, 16(5), 1779.
- Xu, D., Liu, H., Qiu, D., & Ma, Q. (2026). Structured Modeling and Representation Methods for Post-Retrieval Inference Processes in Large Video Language Models.
- Hady, M. A., Hu, S., Pratama, M., Cao, Z., & Kowalczyk, R. (2025). Multi-agent reinforcement learning for resources allocation optimization: a survey. *Artificial Intelligence Review*, 58(11), 354.
- Zhang, Y., Gu, W., & Wang, J. (2026). Construction of Wind Farm Asset Health Index Based on Multi-Dimensional Indicators and Analytic Hierarchy Process and Its Correlation with Operational Performance. *Authorea Preprints*.
- Charalambous, T., Pappas, N., Nomikos, N., & Wichman, R. (2025). Toward goal-oriented communication in multi-agent systems: An overview. *arXiv preprint arXiv:2508.07720*.
- Liu, S., Liu, X., & Feng, H. (2025, November). Research on AI-Driven Visual Design and Immersive Interactive Experiences Based on Multimodal Cognition and User. In

- Proceedings of the 2025 International Conference on Digital Society and Intelligent Computing (pp. 734-740).
- Maldonado, D., Cruz, E., Torres, J. A., Cruz, P. J., & Benitez, S. D. P. G. (2024). Multi-agent systems: A survey about its components, framework and workflow. *IEEE Access*, 12, 80950-80975.
- Qiu, D., Xu, D., & Yue, L. (2025, December). Reinforcement Learning-Augmented LLM Agents for Collaborative Decision Making and Performance Optimization. In *2025 7th International Conference on Frontier Technologies of Information and Computer (ICFTIC)* (pp. 1337-1342). IEEE.
- Tam, P., Song, I., Kang, S., Ros, S., & Kim, S. (2022). Graph neural networks for intelligent modelling in network management and orchestration: a survey on communications. *Electronics*, 11(20), 3371.
- Zhang, Y., Gu, W., & Wang, J. (2025). Research on First Article Inspection (FAI)-Driven Quality Assurance Methods for Wind Turbine Installation and Operation & Maintenance and Their Effect on Reliability Improvement. Available at SSRN 6094206.
- Charalambous, T., Pappas, N., Nomikos, N., & Wichman, R. (2025). Toward goal-oriented communication in multi-agent systems: An overview. *arXiv preprint arXiv:2508.07720*.
- Gao, G., Gao, R., Gao, R., & Zhou, H. (2026). Engineering Analysis and Quantitative Research on the Platform-Based Evolution of Enterprise Communication Systems.
- Kumar, S., & Thakur, M. K. (2025). Attention dynamics: Evolution of attention-based view and its implications. *Management Review Quarterly*, 75(4), 3727-3772.
- Jiao, Y., Zhao, B., Wang, A., & Shi, T. (2026). Construction and Empirical Study of a Modularized Teaching System for Art Courses Based on a Unified Training Pathway.
- Low, D., & Zhou, Y. (2025). Cooperative multi-agent reinforcement learning for robotic systems: A review. *Multiagent and Grid Systems*, 21(2), 96-123.
- Xu, D., Chen, H., & Gui, H. (2026). Unified Online Estimation Method for SOC, SOH, and Power Capacity Considering Safety Boundary Consistency in Battery Management Systems.
- Singh, P., Hazarika, B., & Huang, W. J. (2025). Graph-based Federated Multi-Agent DRL for Semantic and Intent-Aware V2X Communication. *IEEE Internet of Things Journal*.
- Wang, Y., Yin, X., Chen, J., & Wang, Y. (2026). Evidence-Based Study on Low-Burden Digital Phenotyping for Precision Screening of Oral Anti-Obesity Drug Efficacy.
- Okubo, T., & Takahashi, M. (2023). Multi-agent action graph based task allocation and path planning considering changes in environment. *IEEE Access*, 11, 21160-21175.
- Jiao, Y., Wang, A., Zhao, B., & Shi, T. (2026). The Impact of Visual Language Strategies in Public Art Creation on Community Spatial Perception and Public Behavior.
- Munikoti, S., Agarwal, D., Das, L., Halappanavar, M., & Natarajan, B. (2023). Challenges and opportunities in deep reinforcement learning with graph neural networks: A comprehensive review of algorithms and applications. *IEEE transactions on neural networks and learning systems*, 35(11), 15051-15071.
- Xu, D., Gui, H., & Chen, H. (2026). Research on Layered Control and Fault Recovery Mechanisms for Fast Charging Safety Diagnosis of High Voltage Battery Systems Under Charging Network Interoperability Conditions.
- Pasham, S. D. (2024). Scalable Graph-Based Algorithms for Real-Time Analysis of Big Data in Social Networks. *The Metascience*, 2(1), 92-129.
- Gao, G., Ma, X., Lu, C., & Gao, R. (2026). Reliability Analysis and Application Research of SMS Communication Systems in Medical Notification Scenarios.