Zhongyang MAO, Zhilin ZHANG, Faping LU, Xiguo LIU, Yaozong PAN, Jiafang KANG, Yang YOU, 2025. Dynamic joint resource allocation in maritime wireless communication networks: a meta-reinforcement learning approach based on knowledge embedding. *Frontiers of Information Technology & Electronic Engineering*, 26(12):2672-2687. https://doi.org/10.1631/FITEE.2500007

Dynamic joint resource allocation in maritime wireless communication networks: a meta-reinforcement learning approach based on knowledge embedding

Key words: Marine wireless communication; Resource allocation; Knowledge embedding; Meta-reinforcement learning

Corresponding author: Zhilin ZHANG
E-mail: zzl19970811@163.com
ORCID: https://orcid.org/0009-0006-1442-3735

Motivation

- As human exploration of the ocean expands, the demand for continuous, high-quality, and ubiquitous maritime communication is steadily increasing.
- However, the dynamic nature of the marine environment and resource constraints present significant challenges for traditional heuristic resource allocation methods, complicating the balance between high-quality communication and limited network resources. This results in suboptimal system throughput and an over-reliance on specific problem structures.

Main idea

1.We integrate meta-reinforcement learning with the OFDMA framework to optimize information energy efficiency per unit power for autonomous decision-making in offshore base stations. We propose a joint resource allocation method based on knowledge-embedding meta-reinforcement learning. This method improves the meta-DRL approach's ability to generate action combination strategies for joint resource allocation by aligning the distribution of multiple agent actions through dynamic transfer mapping.

2. By incorporating knowledge embedding, we design a domain knowledge-based physical guidance loss function to guide the meta-DRL model in allocating power and spectrum in accordance with the physical world's known rules, thereby reducing the model's dependency on large amounts of data.

Method

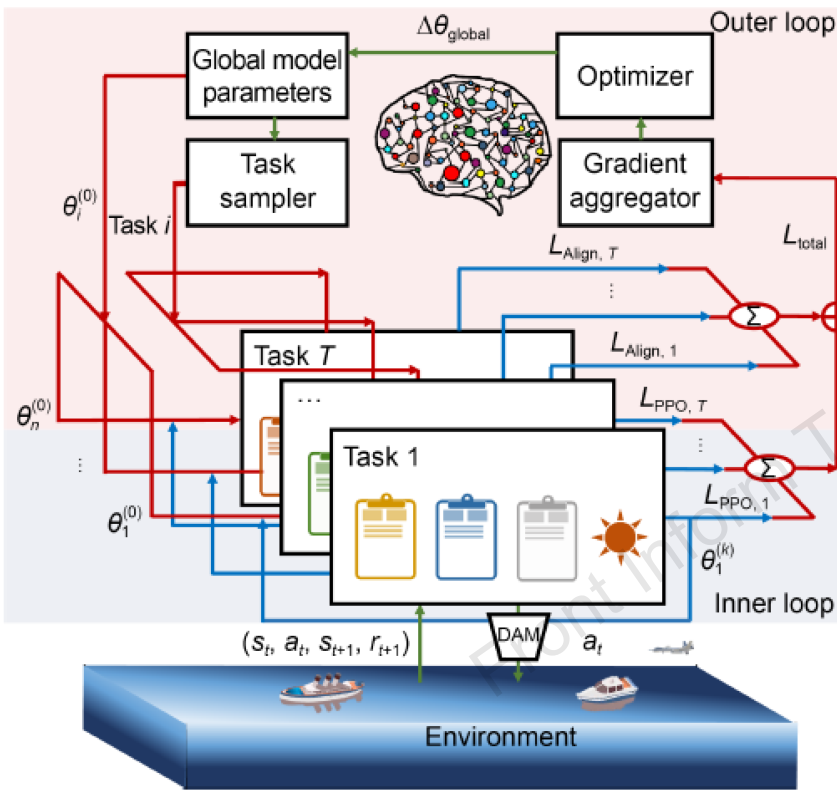

Fig. 3 Dynamic joint resource allocation method for maritime wireless communication based on knowledge embedding (DAM: distribution alignment module)

1. We propose a joint resource allocation method, knowledge-embedding model-agnostic meta-learning (KE-MAML), based on meta-reinforcement learning and knowledge embedding.

2. We introduce a universal action distribution alignment module. This module uses various distribution mapping techniques based on the level of conflict within action combinations, enabling agents to generate actions that adhere to real-world constraints.

Major results

1. Table 3 presents the average algorithm performance for tasks with the same channel condition intervals.

Table 3 Algorithm performance comparison

Algorithm	Total throughput (Mbit/s)			Average throughput for number-independent nodes (Mbit/s)			Fairness index		
	Channel 1	Channel 2	Channel 3	Channel 1	Channel 2	Channel 3	Channel 1	Channel 2	Channel 3
DAM-MAML-PPO	140.90	98.52	39.71	45.42	29.05	9.80 ↑	0.35	0.39	0.50 ↑
Loss_MAML-PPO	118.10	74.52	32.76	28.81	18.57	9.20	0.38	0.36	0.31
MAML-PPO	123.29	76.67	36.77	35.41	18.76	7.96	0.31	0.35	0.45
RL^2	88.77	53.25	30.59	16.73	13.26	6.30	0.43 ↑	0.42 ↑	0.48
KE-MAML	173.37 ↑	108.50 ↑	44.28 ↑	52.99 ↑	29.19 ↑	9.17	0.35	0.31	0.41

↑ indicates the optimal parameter

Major results

2. Comparison of the total throughput between the proposed method and the comparative method under different training data sizes

Table 4 Comparison of total throughput under different training data sizes

Training data size	KE-MAML			MAML-PPO		
	Channel 1	Channel 2	Channel 3	Channel 1	Channel 2	Channel 3
5×10	175.71	108.94	43.79	92.00	61.40	31.27
10×5	167.37	108.55	44.04	115.41	78.57	35.49
10×10	175.19	100.36	45.16	99.10	58.23	31.60
30×20	173.37	108.50	44.28	123.29	76.67	36.77
Mean squared error	3.83 ↓	4.16 ↓	0.60 ↓	14.41	10.39	2.76

↓ indicates the optimal parameter

Conclusions

1. Simulation results show that KE-MAML serves as an effective communication resource allocation strategy, sacrificing fairness in favor of maximizing system throughput in resource-limited and unfamiliar environments.

2. Compared to various benchmark methods, KE-MAML significantly enhances system throughput across diverse channel conditions, further validating the effectiveness of our joint resource allocation meta-reinforcement learning method grounded in knowledge embedding.

Zhongyang Mao received the Ph.D. degree from Naval Aviation University, China, in 2011. He is currently a Professor in Naval Aviation University, Yantai, China. His research areas include modern communication system and intelligent networking.

Zhilin Zhang received the M.D. degree from Naval Aviation University, China, in 2021. He is currently pursuing his PhD degree in Naval Aviation University, Yantai, China. His research areas include wireless communication network and resource allocation.