The Seventeenth International Conference on Advances in System Modeling and Simulation

Generalizable Spatiotemporal Reinforcement Learning Model for Maritime Search Path Planning

Authors: **Pengcheng Yang**, Yingying Gao, Jing Xu, Qingqing Yang

College of Systems Engineering, National University of Defense Technology, China

Contact email: **yangpengcheng@nudt.edu.cn**

Personal Biography

PengCheng Yang

received the bachelor's degree in Geographic Information Science from Shandong University of Science and Technology, China in 2019. He is currently a master's student majoring in Management Science and Engineering at the College of Systems Engineering, National University of Defense Technology.

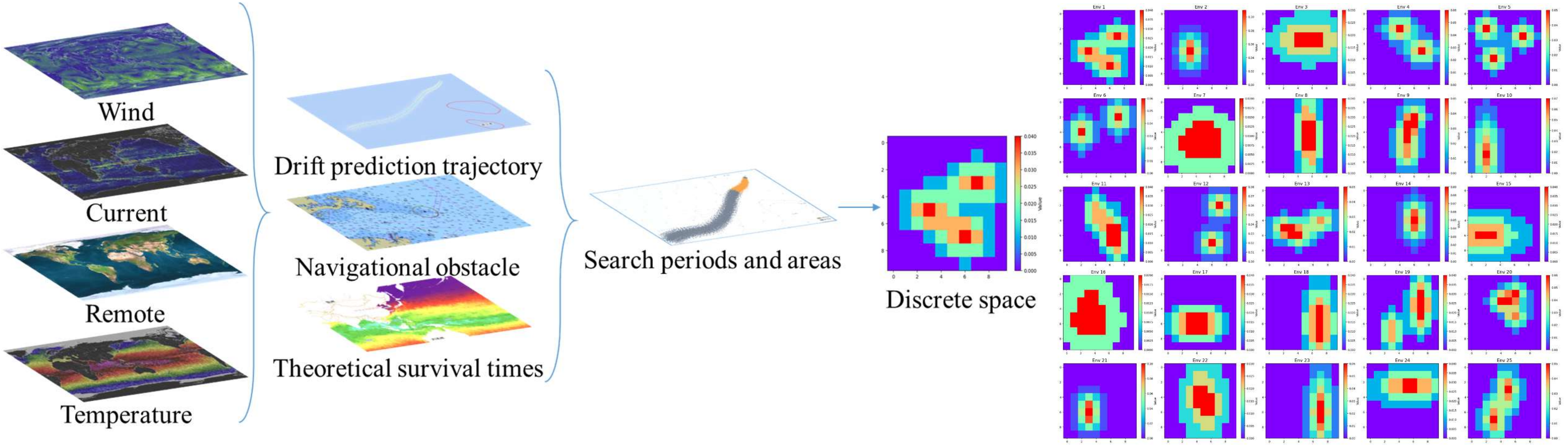
His research interest focuses on reinforcement learning, maritime search and rescue, emergency management, and intelligent decision-making.

Problems

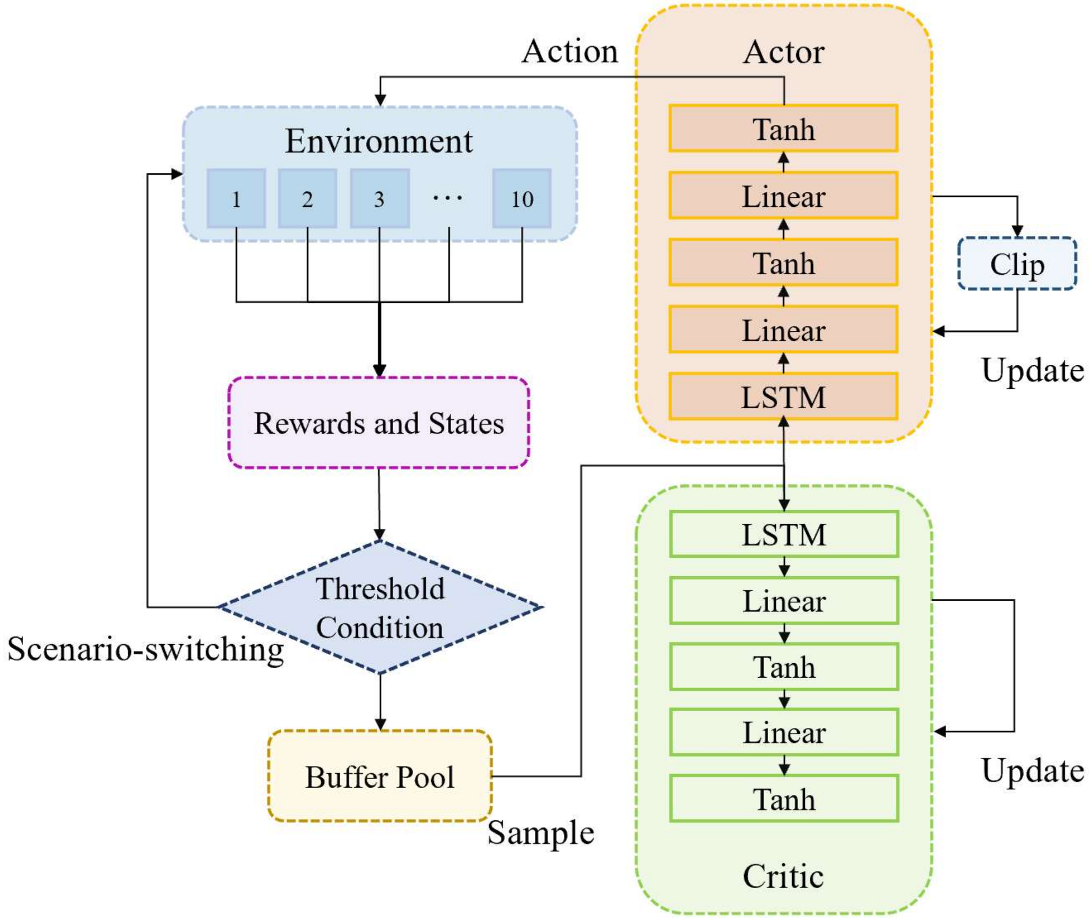
- Maritime accidents are increasing, making search and rescue very important. The search phase takes the most time; we need to find targets quickly.
- But existing path planning methods are **not flexible**, are **slow**, and **cannot adapt to new scenes well**. We use reinforcement learning, but its generalization ability is still weak.
- So, the core problem is: **How to improve the generalization and efficiency of path planning models**?

Scenario Generation Framework

- Current mainstream method uses **Monte Carlo simulation** with a large number of particles to simulate all possible drift routes near the point of fall, and the final result conforms to a **Gaussian distribution**.
- So, we directly used random Gaussian distributions to efficiently generate **1,000 environments**, with 900 for model training and 100 for testing purposes.



Improvement



- **State Space**
  Local observation+Time remaining+Current position
- **Action Space**
  Move to 4 adjacent grid cells (East/South/West/North)
- **Reward Function**
  Positive reward for new discoveries
  Penalty for repeated searches

- **LSTM Enhancement**
  We integrated LSTM into PPO to capture temporal patterns, significantly improving spatiotemporal feature extraction and sequential decision-making for path planning.

- **Training Mechanism**
  A threshold-based mechanism only switches training scenes when performance reaches 90% of optimal, ensuring stable learning and accelerated convergence.

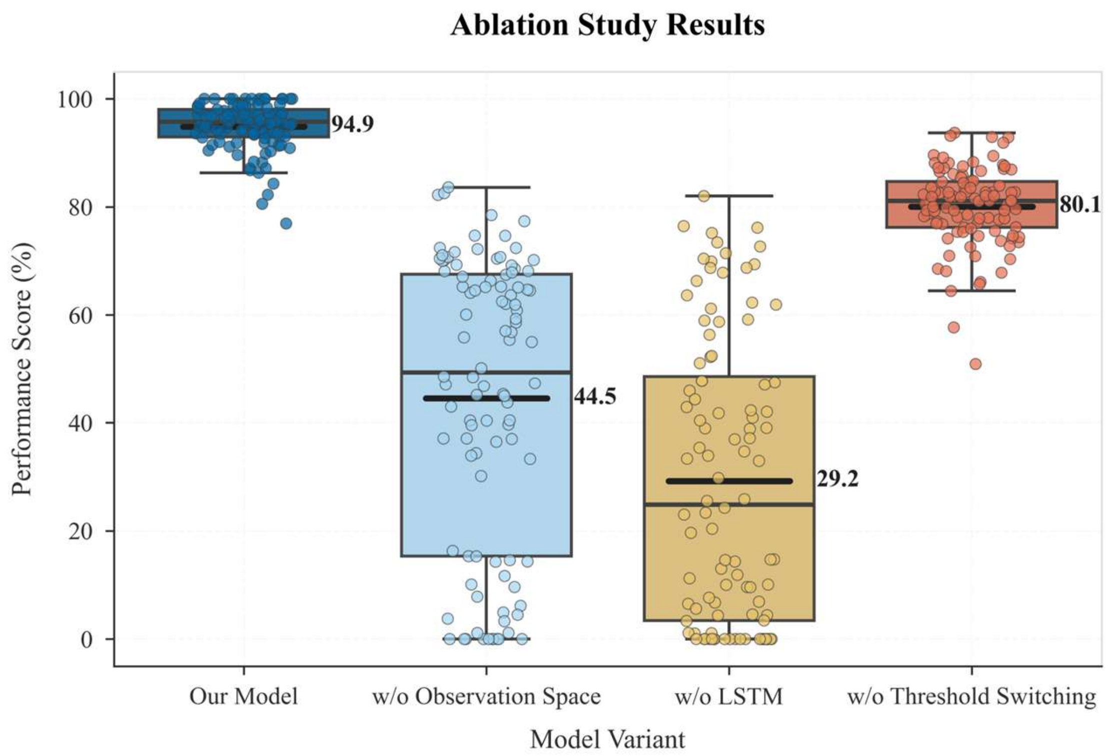
Results

➢ Generalization Results

Our method achieved **94.87%** performance on unseen scenarios, demonstrating strong adaptability across different target distributions.

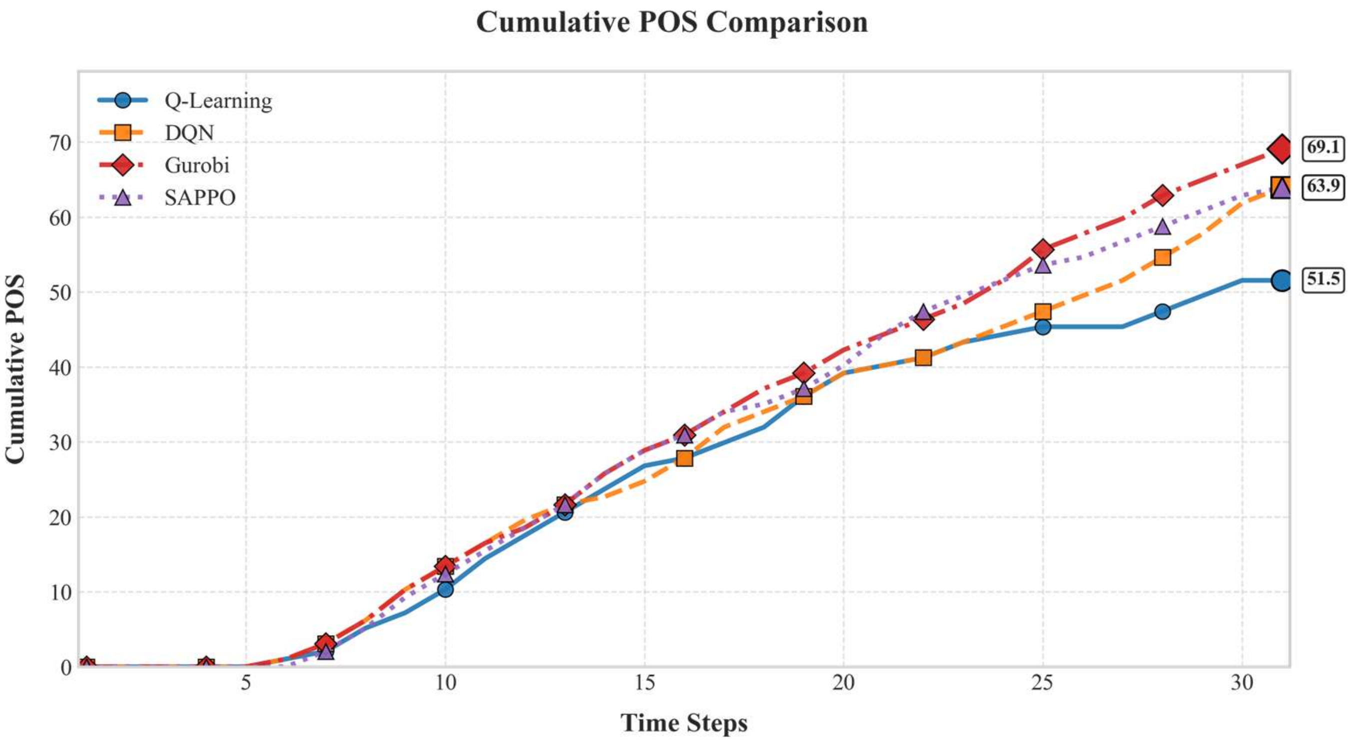
➢ Ablation Findings:

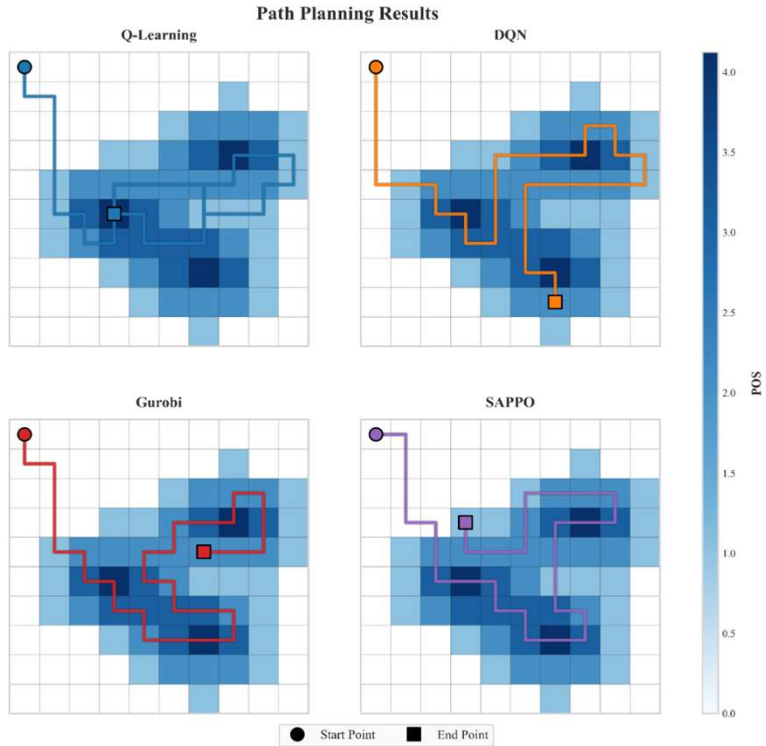
Removing the observation window, LSTM, or the threshold mechanism all caused significant performance drops, proving that all three components are essential for stable and efficient learning.



Real Case Validation

We validated our method using a real maritime distress case in the Pearl River Delta. Our algorithm achieves similar search performance to the optimal solver while being **ten times faster**, and significantly outperforms Q-Learning and Deep Q Network algorithms by avoiding redundant paths and focusing on high-probability areas.





Conclusion And Future Work

- **Conclusion**
  Our proposed algorithm effectively addresses maritime search path planning by integrating spatiotemporal awareness through LSTM, demonstrating strong generalization (94.87% performance) and high efficiency (10× faster than traditional methods) in real-case validation.

- **Future Work**
  We will explore multi-agent collaborative search systems and extend our method to more complex environments with dynamic obstacles and multi-target tracking capabilities.

The Seventeenth International Conference on Advances in System Modeling and Simulation

Generalizable Spatiotemporal Reinforcement Learning Model for Maritime Search Path Planning

Authors: **Pengcheng Yang**, Yingying Gao, Jing Xu, Qingqing Yang

College of Systems Engineering, National University of Defense Technology, China

Contact email: **yangpengcheng@nudt.edu.cn**