

# 基于图神经网络技术的水下无人系统智能决策研究

冯振宇, 彭 倍, 王 刚

(电子科技大学机械与电气工程学院, 四川 成都 611731)

**摘 要:** 人工智能辅助决策, 是实现水下无人系统集群作战应用智能化所面临的关键问题。在实际作战应用中, 水下无人系统集群存在装备异构性、约束动态性、任务不确定性等问题。传统的人工智能方法难以解决状态及约束要素动态变化所导致的模型不确定性问题。图神经网络技术是基于认知科学的连接主义人工智能方法——关系型强化学习的一种。通过构建决策图, 用决策图的顶点表示无人系统集群智能决策状态及约束要素属性, 用决策图的边表示各决策要素之间的逻辑推理关系属性, 通过强化学习方法训练整个决策图的顶点属性、边属性以及决策图全局属性, 从而通过提取决策图的属性信息实现水下无人系统的人工智能辅助决策。本文旨在探究将图神经网络技术应用在水下无人系统智能辅助决策中的可行性, 对水下无人系统智能辅助决策进行图神经网络建模, 构建智能辅助决策推理算法伪代码, 研究基于图神经网络技术的水下无人系统智能决策的技术实现。

**关键词:** 水下无人系统; 人工智能; 智能辅助决策; 图神经网络

**中图分类号:** TP3-05; TP301.6 **文献标识码:** A

**文章编号:** 1672-7649(2020)12-0063-04 **doi:** 10.3404/j.issn.1672-7649.2020.12.012

## Research on intelligent decision making of underwater unmanned system based on graph neural network technology

FENG Zhen-yu, PENG Bei, WANG Gang

(School of Mechanical and Electrical Engineering, University of Electronic Science and  
Technology of China, Chengdu 611731, China)

**Abstract:** Artificial intelligent decision support is a key problem to realize intelligent application of underwater unmanned system cluster operation. In practical operational applications, underwater unmanned system clusters are confronted with such problems as equipment heterogeneity, constraint dynamics and mission uncertainty. The traditional artificial intelligence method is difficult to solve the model uncertainty. The graph neural network technology is one of the connectionist artificial intelligence methods based on cognitive science, relational reinforcement learning. By building decision graph, by the decision graph vertex unmanned systems cluster state and constraint factors attribute of intelligent decision with the decision graph to represent the logic relations between various decision-making factors attribute, through reinforcement learning method to train the whole decision graph vertices attributes, attribute and decision graph of global properties, thereby graph by extracting decision attribute information to realize the underwater unmanned systems artificial intelligence decision-making. The purpose of this paper is to explore the graph neural network technology application in underwater unmanned systems intelligent aided decision-making feasibility of exploratory on underwater unmanned intelligent aided decision-making system for the graph neural network modeling, building intelligent auxiliary decision-making reasoning algorithm pseudo-code, research based on the technique of graph neural network underwater unmanned system implementation of the intelligent decision technique.

**Key words:** underwater unmanned system; artificial intelligence; intelligent decision making; graph neural network

## 0 引 言

图神经网络 (graph neural network, GNN) 是一种

流行的图形数据学习表示工具, 包括但不限于社交网络、分子图和知识图<sup>[1]</sup>。GNN 比传统的决策树的逻辑推理更具有效性<sup>[2-4]</sup>。基于图神经网络的水下无人系统

收稿日期: 2020-08-18

基金项目: 四川省国际合作基金资助项目 (2019YFH0017)

作者简介: 冯振宇 (1990-), 男, 博士研究生, 研究方向为人工智能强化学习、图神经网络、无人系统集群智能决策。

Electronic Publishing House. All rights reserved. <http://www.cnki.net>

智能决策,是针对水下无人系统集群实际应用的特殊需求,如通信受限、多任务、复杂多变环境等约束,为了满足水下无人系统集群智能的决策需求,采用认知推理理论,运用图神经网络方法解决水下无人系统集群智能实现过程中存在的智能决策问题。集群智能决策主要是基于“约束—集群—环境”的认知推理,并在实时交互环境中实现集群的任务分配策略、任务执行策略的推理决策。

人工智能强化学习(reinforcement learning, RL)领域是基于知识表示、认知学习实现智能推理决策的,当前的强化学习方法在关系型问题推理和约束动态推理上存在很多不足<sup>[5]</sup>。所以,根据领域最新的研究成果,基于图神经网络的强化学习方法是实现关系型问题推理和约束动态推理的研究方向之一<sup>[6]</sup>。

强化学习的基本思想是智能体(Agent)在与环境交互的过程中根据环境反馈得到的奖励不断调整自身的策略以实现最佳决策,主要用来解决决策优化类的问题。其基本要素有策略、回报函数、值函数、环境模型,学习过程可以描述为如图1所示的马尔科夫决策过程。强化学习基本学习模型首先智能体感知当前状态 $S$ ,从动作空间 $A$ 中选择动作 $a_t$ 执行;环境根据智能体做出的动作来反馈相应的奖励 $r_{t+k}$ ,并转移到新的状态 $S_{t+k}$ ,智能体根据得到的奖励来调整自身的策略并针对新的状态做出新的决策。强化学习的目标是找到一个最优策略 $\pi^*$ ,使得智能体在任意状态和任意时间步骤下,都能够获得最大的长期累积奖赏:

$$\pi^* = \arg \max_{\pi} E \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} | S_t = S \right\},$$

$$\forall S_t \in S, \forall t \geq 0. \quad (1)$$

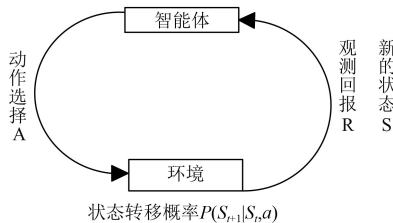


图1 强化学习基本框架

Fig. 1 The basic framework of Reinforcement learning

其中 $\pi$ 表示智能体的某个策略, $\gamma \in [0, 1]$ 为折扣率, $k$ 为未来时间步骤, $S$ 为状态空间。

图神经网络算法理论是基于人脑认知推理决策的认知学科的推理决策模式<sup>[7]</sup>。基于GNN强化学习智能决策算法,是运用人工智能GNN算法来实现集群Agents的智能决策策略的求解,通过构建带有属性的

图[顶点,边],继而通过顶点到边,边到顶点,边和顶点到全局图属性的迭代计算实现智能决策推理学习<sup>[3]</sup>。

图2为针对无人系统集群智能辅助决策系统作战应用的具体场景。首先,进行认知建模,将决策影响因素抽象成实体与关系的图。然后,根据GNN算法原理构建实体、关系,进行推理决策参数训练。

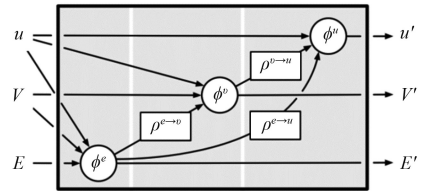


图2 GNN理论模型

Fig. 2 GNN theoretical model

基于图神经网络的无人系统集群智能强化学习研究,是在传统集群智能的基础上,将人工智能-强化学习技术融合进去,主要采用的是连结主义核心思想(见表1),采用最新的图神经网络算法理论,实现无人系统集群的智能的推理决策、模型训练、经验学习,进一步提高无人系统集群智能程度<sup>[8]</sup>。

表1 人工智能实现方法主要流派

Tab. 1 The main schools of artificial intelligence implementation methods

	认知主义	连结主义	具身-生成进路
隐喻	表征与计算	神经网络	具身动力系统
突出成果	工作记忆模型冯诺依曼结构	神经科学深度学习	类脑芯片具身机器人
对应AI流派	符号学派	连接学派	行为学派
例子	计算、理性思维	记忆、刻意训练	象似性、生态知觉
学习观	获得知识	构建联系	创造意义

## 1 水下无人系统智能决策模型建立

航行器节点属性矩阵 $U_{i,t}$ 为描述单个航行器平台固有属性和状态属性的矩阵。该矩阵能够描述航行器性能和当前位置等状态并实时更新,从而在决策图中作为顶点来进行决策图全局属性的迭代计算;任务节点属性矩阵 $T_{i,t}$ 为描述水下无人系统集群在一次任务中需要完成的一个或者多个特殊任务属性的矩阵,该矩阵包含需要执行任务的类型,任务信息描述(区域,范围等);约束节点属性矩阵 $C_{i,t}$ 为描述任务执行过程中的约束条件矩阵,该矩阵包含一次任务过程中水下无人航行器集群会面临的时间约束,能量约束,复杂环境约束等信息;队形节点属性矩阵 $F_{i,t}$ 为描述水下无人

航行器集群任务执行、行进过程中的队形的矩阵, 包含集群需要保持的队形信息; 全局属性矩阵  $G_{i,t}$  为描述决策图所有顶点及其之间对应关系边所构成的决策结果描述矩阵, 包含该次决策结果的衡量和描述信息。

顶点属性更新边的属性, 边是有方向的, 接收顶点矩阵与发出点矩阵通过对应回报计算函数给当前有向边属性进行赋值, 表示当前逻辑连接关系的回报。通过对应的预先设定的回报计算函数来进行任务  $\leftarrow f(T, U) \rightarrow$  航行器, 约束  $\leftarrow f(C, U) \rightarrow$  航行器回报, 队形  $\leftarrow f(F, U) \rightarrow$  航行器的对应边关系进行边属性回报值计算:

$$G_t = f(U_{t-1}, T_{t-1}, C_{t-1}, G_{t-1})。 \quad (2)$$

图 3(a) 为基于 GNN 的水下无人系统智能决策的决策表示, 决策图中顶点分别表示在一次智能决策中所有的任务、约束、航行器、队形等需要进行决策和影响决策的信息。图 3(b) 为在初始化的决策图的基础上, 通过决策算法 1 的强化学习迭代求解对决策图的顶点属性、边属性、全局图属性进行了一定程度的更新, 直至最终算法迭代终止, 完成一次基于 GNN 的水下无人系统智能决策的强化学习, 并根据强化学习结果给出对应的最优智能决策的策略。

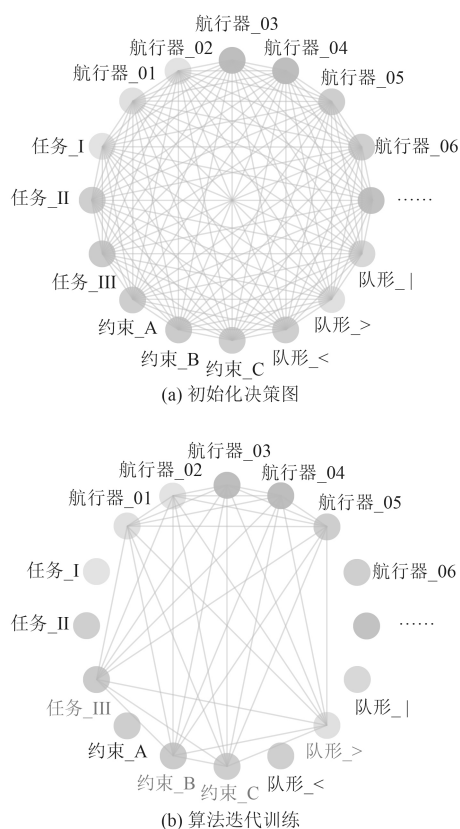


图 3 决策训练  
Fig. 3 Decision training

## 2 水下无人系统智能决策仿真

基于 GNN 的水下无人系统决策仿真试验验证, 对 1 个任务、1 个约束、1 个队形约束、4 台不同类型的水下无人航行器进行智能决策 GNN 图的强化学习, 目的是从 4 台无人水下航行器中选择几台来执行满足该约束与队形的任务。图 4 为 MDP 强化学习的仿真 GNN 决策图。其中节点 1 表示搜索任务, 节点 2 表示能量约束, 节点 3 表示任务对应的队形约束, 节点 4~节点 7 表示可以选择来完成航行的航行器, 每个航行器的最大速度、续航能力等都不相同。如果能够满足任务需求该航行器代表的顶点与约束顶点的边属性即强化学习回报为 1, 否则回报为 0。同理, 决策时能够满足任务需求的边属性回报值为 1, 否则为 0。

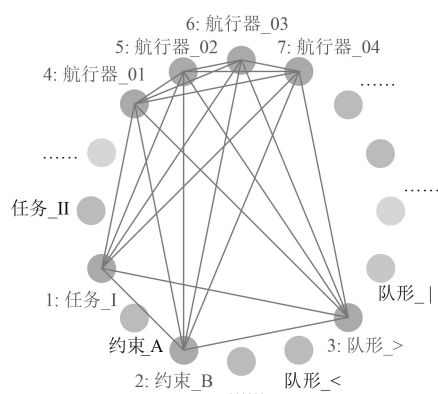


图 4 MDP 仿真模型 GNN 决策图  
Fig. 4 Simulation GNN Decision Graph of MDP

图 5 为基于图神经网络技术的水下无人系统智能决策 Matlab 仿真试验结果, 其中纵坐标是决策图全局属性的总回报, 横坐标是决策图进行强化学习训练的迭代步数。由图可知, 在进行 100 次训练时就可以通过 GNN 决策图输出可以满足任务执行需求的决策结果。最优的决策图全局策略回报值为 4, 如果强化学习对决策图的训练结果总回报为 4 时即表示策略成功。

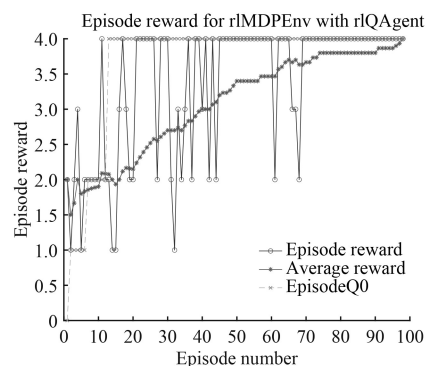


图 5 智能决策强化学习仿真结果  
Fig. 5 Intelligent decision RL simulation results

最终基于 GNN 的水下无人系统智能决策强化学习仿真给出的决策策略结果如图 6 所示。从备选 UUV 中选择航行器 02, 航行器 03, 航行器 04 即可顺利完成节点 2、节点 3 约束下的节点 1 任务。

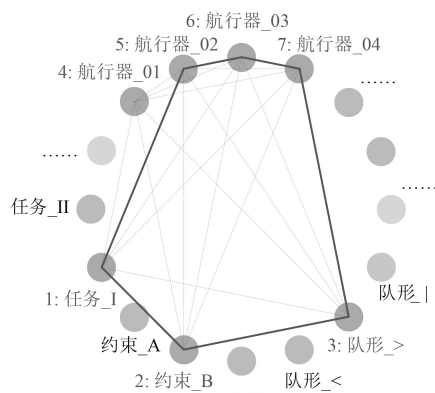


图 6 智能决策策略结果

Fig. 6 Intelligent Decision Policy Result

所以, 从该决策仿真试验的结果来看, 水下无人系统能够基于图神经网络技术与人工智能强化学习方法有效结合, 在较短的时间内提出智能决策策略, 从而为指挥人员提供参考。

### 3 结 语

本文的研究表明基于图神经网络智能决策方法能够满足水下无人系统智能决策动态任务, 动态约束, 动态集群需求的智能决策, 并且能够快速训练出最优的决策策略, 为指挥人员提供智能决策建议。但是,

本文的研究简化了决策图顶点属性及顶点间边属性的回报值计算。在后续的研究中, 将结合实际应用场景进行决策边属性的更新计算, 并探索不同决策图之间共性顶点的经验学习。

### 参考文献:

- [1] HAMRICK J B, ALLEN K R, BAPST V, et al. Relational inductive bias for physical construction in humans and machines[J]. arXiv: Learning, 2018.
- [2] WANG T, LIAO R, BA J, et al. NerveNet: Learning structured policy with graph neural networks[C]. International Conference on Learning Representations, 2018.
- [3] BATTAGLIA P W, HAMRICK J B, BAPST V, et al. Relational inductive biases, deep learning, and graph networks[J]. arXiv: Learning, 2018.
- [4] ZAMBALDI V, RAPOSO D, SANTORO A, et al. Relational deep reinforcement learning[J]. arXiv: Learning, 2018.
- [5] SILVER D, HUBERT T, SCHRITTWIESER J, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play[J]. *Science*, 2018, 362(6419): 1140–1144.
- [6] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. *Nature*, 2015, 518(7540): 529–533.
- [7] ZHANG Y, CHEN X, YANG Y, et al. Can graph neural networks help logic reasoning[J]. arXiv: Learning, 2019.
- [8] VINIYALS O, BABUSCHKIN I, CZARNECKI W M, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning[J]. *Nature*, 2019, 575(7782): 350–354.