Journal of System Simulation

Volume 33 | Issue 12 Article 8

1-13-2022

A Data-Driven Modeling Method for Game Adversity Agent

Zeng Bi
1. Beijing Simulation Center, Beijing 100854, China; ;2. Beijing Institute of Electronic System, Beijing 100854, China; ;

Fang Xiao
1. Beijing Simulation Center, Beijing 100854, China; ;

Deshuai Kong
3. China Aerospace Science and Industry Corporation Limited, Beijing 100048, China;

Xiangxiang Song
1. Beijing Simulation Center, Beijing 100854, China; ;

See next page for additional authors

A Data-Driven Modeling Method for Game Adversity Agent

Abstract

Abstract: Aiming at the problems of collaborative modeling of formation behavior and intelligent generation of decision-making in complex confrontation scenarios, *based on the serious game to simulate the confrontation scenarios of complex maritime equipment against the air, this paper proposes a data-driven modeling method for game agent and uses a distributed modeling technology of parallel adversarial scenarios and opportunistic decision making technology of smart targets to achieve agent modeling. It provides support for the further exploration of multi-objective collaborative modeling in complex confrontation scenarios.* The simulation results show that deep reinforcement learning algorithms can provide a basis for the modeling of agents dexterous strategies.

Keywords

deep reinforcement learning, data-driven, distributed training, opportunistic decision making

Authors

Zeng Bi, Fang Xiao, Deshuai Kong, Xiangxiang Song, Zhengxuan Jia, and Tingyu Lin

Recommended Citation

Zeng Bi, Fang Xiao, Kong Deshuai, Song Xiangxiang, Jia Zhengxuan, Lin Tingyu. A Data-Driven Modeling Method for Game Adversity Agent[J]. Journal of System Simulation, 2021, 33(12): 2838-2845.

第 33 卷第 12 期 系统仿真学报© Vol. 33 No. 12
2021 年 12 月 Journal of System Simulation Dec. 2021

一种数据驱动的对抗博弈智能体建模方法

曾贲[1,2]，房霄[1]，孔德帅[3]，宋祥祥[1]，贾政轩[1,2]，林廷宇[1,2]
(1. 北京仿真中心，北京 100854；2. 北京电子工程总体研究所，北京 100854；3. 中国航天科工集团有限公司，北京 100048)

摘要：针对复杂对抗场景下编队行为协同建模及决策智能生成等问题，*提出一种数据驱动的对抗博弈智能体建模方法，依托基于严肃游戏的复杂海上装备对空对抗模拟场景，通过基于并行对抗场景的分布式训练技术与基于灵巧目标的临机决策建模技术，结合空中目标、复杂海上装备等能力模型，实现人机混合增强的智能体建模，为后续深入开展复杂对抗场景下多目标协同建模研究提供了支撑。*实验结果表明：深度强化学习算法能够为智能体灵巧策略的建模提供基础。

关键词：深度强化学习；数据驱动；分布式训练；临机决策
中图分类号：TP391.9 文献标志码：A 文章编号：1004-731X (2021) 12-2838-08
DOI: 10.16182/j.issn1004731x.joss.20-FZ0532

A Data-Driven Modeling Method for Game Adversity Agent

Zeng Bi[1,2], *Fang Xiao*[1], *Kong Deshuai*[3], *Song Xiangxiang*[1], *Jia Zhengxuan*[1,2], *Lin Tingyu*[1,2]
(1. Beijing Simulation Center, Beijing 100854, China;
2. Beijing Institute of Electronic System, Beijing 100854, China;
3. China Aerospace Science and Industry Corporation Limited, Beijing 100048, China)

Abstract: Aiming at the problems of collaborative modeling of formation behavior and intelligent generation of decision-making in complex confrontation scenarios, *based on the serious game to simulate the confrontation scenarios of complex maritime equipment against the air, this paper proposes a data-driven modeling method for game agent and uses a distributed modeling technology of parallel adversarial scenarios and opportunistic decision making technology of smart targets to achieve agent modeling. It provides support for the further exploration of multi-objective collaborative modeling in complex confrontation scenarios.* The simulation results show that deep reinforcement learning algorithms can provide a basis for the modeling of agents dexterous strategies.

Keywords: deep reinforcement learning; data-driven; distributed training; opportunistic decision making

引言

面向未来高对抗、强博弈的复杂场景，对通过模拟训练加强复杂装备任务遂行能力的要求越来越高，传统的基于相对固定空中态势的仿真模拟已经难以满足模拟训练智能化发展的需要[1]，并存在与真实状态下目标模拟相差甚远、临机动态调整能力欠缺、对抗样式不足等问题，难以满足以提升实战化为目的的对抗训练。

传统的基于预先想定的建模方法。一般从目标运动特征、探测能力、决策能力以及防御能力等 4 个方面考虑：①目标运动特征模拟采用点迹建航法和六自由度建模法。点迹建航法主要思路为将目标抽象为一个质点，通过构建质点运动约束实现质点的运动模拟，考虑的约束一般包括升限、速度、转弯半径等条件。六自由度建模相比点迹建航法，能够更加精细地实现对目标运动特征的模拟[2]；②探测能力模拟主要模拟探测雷达威力。分为雷

收稿日期：2020-04-01 修回日期：2021-06-08
基金项目：国防基础科研(JCKY2018204C004)
第一作者：曾贲(1994-)，男，硕士，工程师，研究方向为深度强化学习技术。E-mail：duanting18@nudt.edu.cn

达威力包络模拟以及信号注入模拟等；③决策能力模拟主要模拟对抗中的指挥决策过程，在某些场景下常采用博弈论或者优化算法对指挥决策行为进行建模，如比较典型的粒子群优化算法[3]。但随着复杂海上装备数量及空中目标单元数量的递增，该优化问题的求解空间将逐渐增大至不可求解，而且极大消耗计算资源，很难适用于计算资源有限的复杂装备模拟训练中；④防御能力模拟主要采用轨迹拟合法，通过数据模型抽象出轨迹拟合公式进行模拟。

面向深度强化学习的智能体建模方法。近年来，在大数据、云计算、机器视觉等技术突飞猛进的基础上，人工智能得到了空前的发展，并逐步向着自主学习、数据驱动、虚实融合的方向演化，进而逐渐在应对多维度的复杂设计问题上实现了颠覆性的突破[4]，甚至在一些领域上超越了人类，如面向围棋/中国象棋/国际象棋[5]、DOTA2/星际争霸 II 等博弈对抗的系统设计上已经完美超越人类。特别地，2019 年，在更具挑战的即时策略游戏星际争霸 II 中，Deepmind 公司设计的 AlphaStar 又以 10:1 的战绩横扫世界顶尖职业玩家[6]，通过保持资源要素的合理调配、作战单元的临机决策为前提，短期、长期的目标规划，最终以精妙的战术规划、灵巧的进攻方式击败对手。类比到复杂海上装备模拟训练场景中，诸如不完备信息条件下的对抗博弈，长远规划策略学习以及大规模交战及决策空间求解等问题，已经在 AlphaStar 智能体上有所突破。随着智能化技术的发展，AI 智能越来越多地出现在美军的训练过程中，美军模拟训练正向实战化、智能化、体系化发展。

本文提出一种数据驱动的对抗博弈智能体建模方法，通过构建空中目标智能体自学习决策模型，充分赋予其智能特性，并基于平行对抗建模技术，让智能体并行地在不同对抗场景下与不同编队组合的复杂海上装备进行博弈，驱动智能体全方位地持续演进，孵化能够动态捕捉海上编队防御漏洞的灵巧智能体，形成多类别、适应性强的临机决策能力，进而在对抗强度、真实度足够的情况下实现对复杂装备边界能力的认知。

1 面向深度强化学习的智能体建模方法

1.1 场景定义

为简化问题求解，本文考虑基于严肃游戏的单/双机与复杂海上装备的对抗场景，在该场景下，空中目标按特定策略飞行靠近复杂海上装备，飞抵可打击区域，完成打击动作并成功脱离探测范围。而复杂海上装备则会发现目标，并对其进行反击对抗。

在此设定下，所需解决的问题可以抽象为：在考虑目标能力模型、复杂海上装备能力模型以及对抗条件模型约束下，对目标的对抗策略进行寻优。

1.1.1 空中目标的能力模型

为进一步简化问题求解，将目标模型考虑为质点模型。此外，考虑飞行性能以及打击能力的限制，对目标运动及打击模型采取如下限制：

(1) 运动特征模型

最小转弯半径设为 $R_{\min}$，即任意时刻的转弯半径 R 须满足 $R \geqslant R_{\min}$。飞行高度约束为 $H \in [H_{\max}, H_{\min}]$。加速度约束为单轴加速度 a_x, a_y, a_z 必须满足 $a_x, a_y, a_z \in [a_{\max}, a_{\min}]$，运动坐标系为北天东坐标系。飞行合速度限制在 $v \in [v_{\max}, v_{\min}]$ 范围内。

(2) 打击能力模型

设定在打击过程中需沿当前速度方向继续飞行 t s 以保持发射过程稳定，且与复杂海上装备间的夹角 θ 满足 $\theta \in [\theta_{\text{fire}}, \theta_{\min}]$ 方可完成打击动作。

1.1.2 复杂海上装备的能力模型

对复杂海上装备模型从探测模型、反击模型 2 个方面进行描述。

(1) 探测模型

探测模型主要用于模拟复杂海上装备发现跟

第 33 卷第 12 期 系统仿真学报 Vol. 33 No. 12
2021 年 12 月 Journal of System Simulation Dec. 2021

踪空中目标的能力。在海上编队北天东坐标系下，考虑探测半径约束，探测范围描述为

$$\begin{cases} x = r\cos\theta, z = r\sin\theta, r = \sqrt{k^2 y - y^2} \\ y \in \left[0, R_0^2/k^2\right], \theta \in [0, 2\pi] \\ x^2 + y^2 + z^2 \leqslant R_0^2, y \in \left[R_0^2/k^2, R_0^2\right] \end{cases} \tag{1}$$

(2) 防御模型

复杂海上装备防御反击采用简单策略实现：一旦探测到目标，复杂海上装备即对距离较近的目标进行围剿，其发射的飞行器预计飞行时间按其平均飞行速度，以及目标首次被探测到时复杂海上装备与目标之间的距离折算为 $t_{\text{intercept}}$，该值即为预计成功拦截的周期。通过与对抗条件模型中成功条件的比较，实现对目标的拦截，并且复杂海上装备具有一定的发射间隔 t_{interval}。同时，若飞行器未在目标逃出探测区时实现锁定，则一旦逃出探测区既判定拦截飞行器失效。

1.1.3 对抗的条件模型

空中目标由单/双机组成，其对抗条件模型主要涉及以下 2 个方面：

(1) 打击/防御条件

其距复杂海上装备的距离 $D_{\text{plane-ship}}$，满足 $D_{\text{plane-ship}} \leqslant D_{\text{fire}}$ 的条件；海拔满足 $H \in \left[H_{\max}, H_{\text{fire}}\right]$，且需满足速度方向与距复杂海上装备距离矢量的夹角 $\theta \in \left[\theta_{\text{fire}}, \theta_{\min}\right]$ 的条件，才可执行投放飞行器的动作，其运动模型由比例导引法构建。并且当其被飞行器锁定后，可以投放诱饵实现逃脱。

(2) 任务成功条件

基于复杂海上装备模型中的防御模型，考虑空中目标的生存时间是：在第一次被探测时，与复杂海上装备之间的距离除以飞行器的飞行速度进行近似。空中目标被探测到以后时间为 t_{detected}。其任务成功的条件为完成打击动作以后，投放的飞行器需要摧毁复杂海上装备，且每个空中目标必须逃离探测范围。

1.2 智能体建模方法

采用深度强化学习、联盟学习等新一代智能技术，构建空中目标智能体自学习决策模型，并面向并行对抗场景建模，充分生成不同初始状态下的对抗场景，让智能体并行地在不同对抗场景下与复杂海上装备进行博弈，进而认知足够多的对抗样式，从而寻找不同对抗场景下的防御突破点，形成满足各对抗条件下的最优决策集合，建模架构如图 1 所示。

1.2.1 基于灵巧目标的临机决策建模技术

本文采用深度强化学习算法完成空中目标智能体的建模过程，以提升其决策能力。框架如图 2 所示。智能体通过在对抗环境中不断地探索生成动作、感知状态和获得回报，从对抗数据中获得复杂因素的关联性和问题处理的完备性，加强其对复杂关联关系的拟合能力。

考虑常规强化学习的配置，其中智能体会与对抗场景产生互动。在每一个仿真间隔 t，智能体都会观测到一组态势信息 $s_t \in S$，分析判断之后，做出一组动作 $a_t \in A$，然后会收到环境反馈的奖励值 $r(s_t, a_t) \in R$，经过一段时间的迭代训练，智能体会形成一个决策集合 $\pi: S \rightarrow A$。

其中，每一个态势信息都对应智能体的一组动作。这样的一个态势信息与动作的映射函数反映出一种期望回馈，即依据每次获取到的态势信息 $s_t \in S$，从策略 π 中寻找最优的决策，直至对抗结束，所产生的所有累计奖励的值函数为

$$Q_\pi(s,a) = E\left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)\right] \tag{2}$$

式中：$\gamma \in [0, 1]$表示衰减因子。同样地，这个预期回馈也可以评估一个策略 π。因此，可以使用 Q_π 得到一种对 π 的更新方式，其目标为使 $J(\theta)$ 最大化，表示为

$$J(\theta) = E\left[Q_\pi(s, \pi_\theta(s))\right] \tag{3}$$

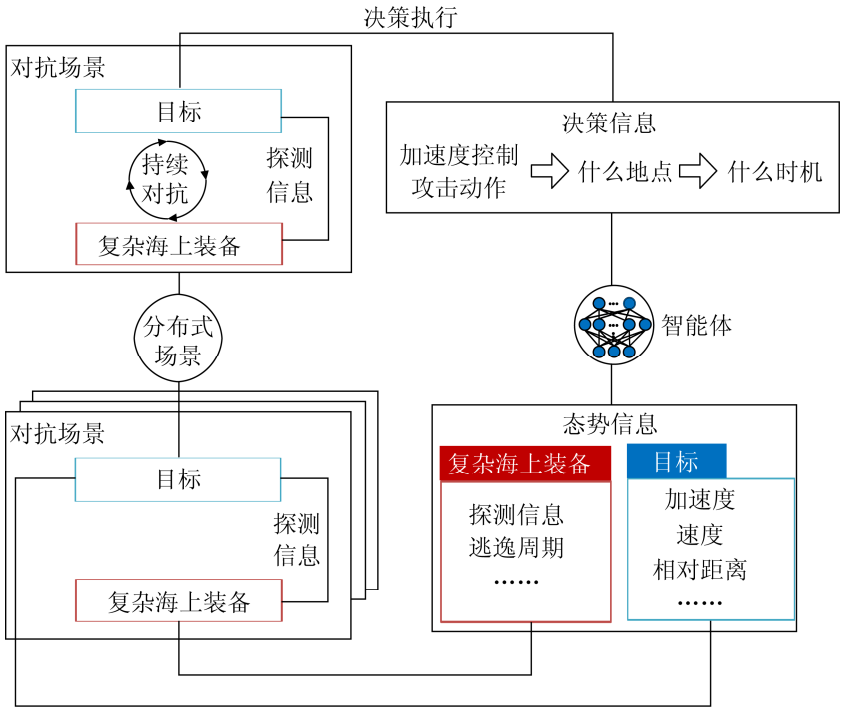

图 1 智能体建模架构

Fig. 1 Agent Modeling Architecture

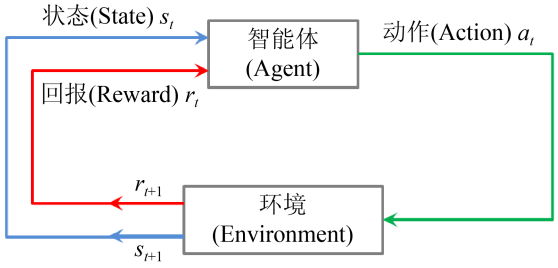

图 2 强化学习框架逻辑图

Fig. 2 Reinforcement learning framework

根据确定型策略梯度算法[7-8]可得策略 π_θ 的参数更新算法为

$$\nabla_\theta J(\theta)\approx E\left[\nabla_\theta \pi_\theta(s)\nabla_a Q_{\pi_\theta}(s,a)|_{a=\pi_\theta(s)}\right] \quad (4)$$

进而规定 π_θ 的更新方向，就能确定策略集合 π 的最终形态，既扮演决策执行者的身份，也称之为 actor 网络[9]。同时，为了更好地评价其 π_θ 的演进方向与真实叠加产生的 $Q_\pi(s, a)$之间的关系，可以设置一位评价者(critic 网络)，通过其观测、评估 actor 的决策质量，校正 actor 的演化方向[10]。使用 Bellman 方程，即

$$(\tau_\pi Q)(s,a)=r(s,a)+\lambda E\left[Q_\pi\left(s',\pi\left(s'\right)\right)\right] \quad (5)$$

式中：s' 表示下一次的态势信息。通过最小化 TD[11]误差的方式，修正值函数与 Bellman 方程推导出来的期望值之间的误差，即二者标准差，表示为

$$L(w)=E\left[\left(Q_w(s,a)-\left(\tau_{\pi_\theta},Q_{w'}\right)(s,a)\right)^2\right] \quad (6)$$

依据 Bellman 方程的更新方式，确实能够找到最优解，但事实上这种建模方式不够合理，单纯利用期望值进行迭代，从某种程度上来说损失了 Q_π 作为分布的信息[12]，因此，采用 N-Step 的分布 Bellman 方程，即

$$\begin{cases}\left(\tau_\pi^N Z\right)(s_0,a_0)=r(s_0,a_0)+\Delta r\\ \Delta r=E\left[\sum_{n=1}^{N-1}\lambda^n r(s_n,a_n)+\lambda^N Z\left(s_N,\pi\left(s_N\right)\right)|s_0,a_0\right]\end{cases} \quad (7)$$

式中：$Z(s, a)$为在状态 s 下执行动作 a 之后回报形成随机变量，具有概率分布的特性[13]，则上述推导出来的更新的方程修改为

$$\begin{cases}L(w)=E\left[d\left(\left(\tau_{\pi_\theta},Z_{w'}\right)(s,a),Z_w(s,a)\right)\right]\\ \nabla_\theta J(\theta)=E\left[\nabla_\theta\pi_\theta(s)E\left[\nabla_a Z_w(s,a)\right]|_{a=\pi_\theta(s)}\right]\end{cases} \quad (8)$$

式中：d 为分部之间的距离度量，采用交叉熵求取。

1.2.2 基于并行对抗场景的分布式训练技术

本文采用 Ring-AllReduce 分布式架构，所有智能体组成单向环形架构，既第 N–1 个智能体的梯度传输给第 N 个智能体，当所有智能体在其负责交互的仿真环境中收敛达到稳定，即可实现分布式训练，如图 3 所示。

在 Ring-AllReduce[14]的分布式智能体并行训练架构下，每一个智能体都会参与到初始条件有变化的仿真程序中进行实际对抗博弈，并且每一个仿真间隔实现一次智能体神经网络梯度采集。通过把第 N–1 个智能体的梯度传递给第 N 个智能体，组成单向环形梯度传递的链路，进而让智能体充分认知不同环境中的情况，实现各智能体梯度的快速收敛。

同时，各智能体采用异步更新的方式，即第 N–1 个智能体的梯度传递给第 N 个智能体后，第 N 个智能体会即刻进行梯度求解，并进行自身参数更新，而不需要同步等待各智能体梯度的目标传递。

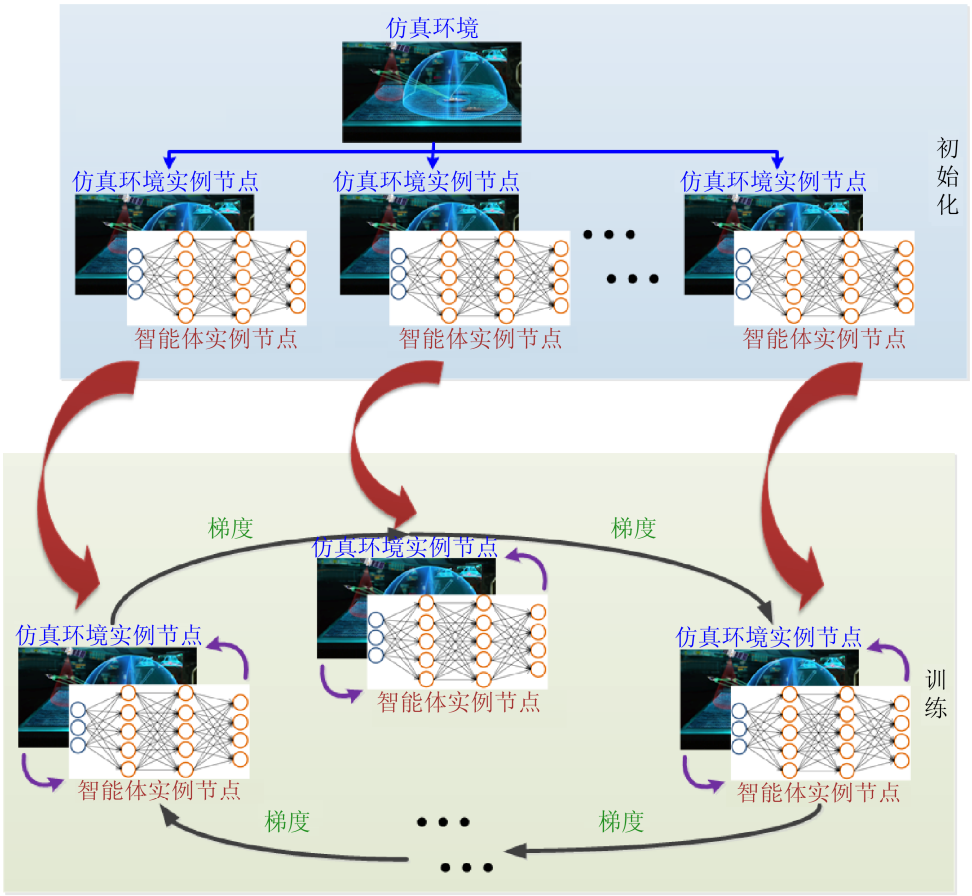

图 3 Ring-AllReduce 架构示意图

Fig. 3 Ring-AllReduce distributed architecture

1.3 算法流程

根据并行对抗场景构建和智能体建模，选取速度、距离、发射角度、是否被探测等信息，作为智能体每一时刻所获取的准确态势信息，其中，(v_x, v_y, v_z)表示目标的速度，(a_x, a_y, a_z)表示目标的加速度，H 表示目标距海平面的高度，θ_{fire} 表示目标投放的夹角，t_{detected} 表示目标被探测的时间总长，*approch* 表示目标是否达到打击的条件，*detected*, *fire*, *back* 均为标记变量，分别表示目标是否被探测、目标是否完成打击动作，以及目标是否脱离探测区域，

survival_rate 表示目标的诱饵与其被跟踪飞行器的数量比值。具体算法流程如图 4 所示。

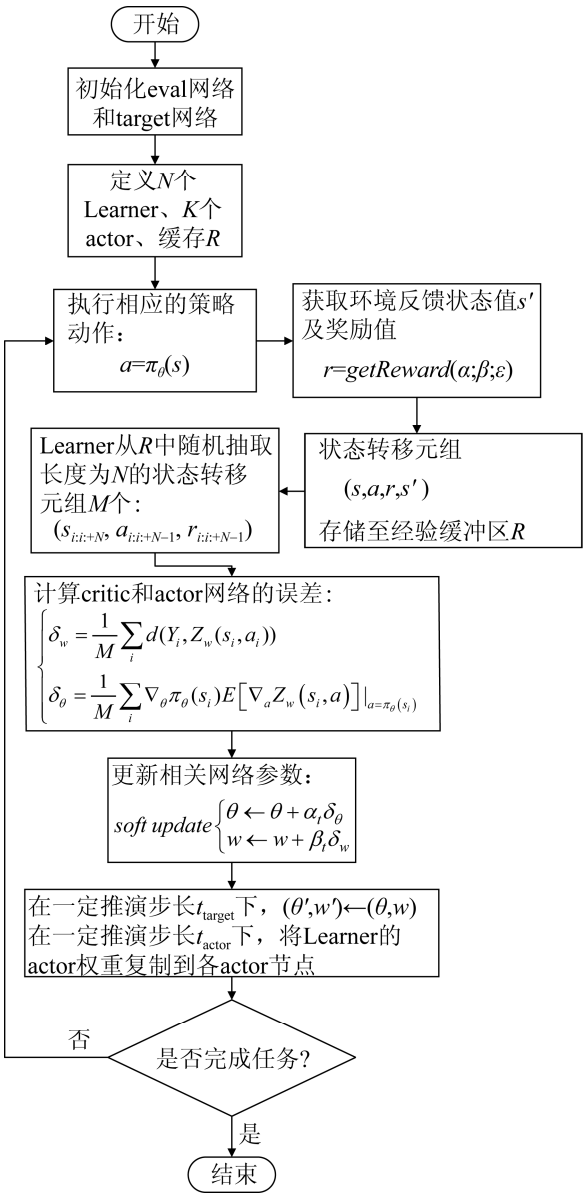

图 4 算法流程
Fig. 4 Algorithm process

2 结果分析

在初始目标位置、打击条件等可随机设置情况下，开展训练任务，经过一段时间的训练，得到目标智能体的收敛模型。为更方便地验证算法的稳定性，针对单机、双机场景开展验证说明，并随机选取打击条件：

$$\begin{cases} \theta_{\text{fire}} \leqslant \pi/6 \\ H_{\text{fire}} \geqslant 2\ \text{km} \\ distance \leqslant 90\ \text{km} \end{cases} \tag{9}$$

2.1 单机场景

选取其中一个对抗场景，不同目标智能体随机抽取的 14 条飞行轨迹如图 5, 6 所示。图中绿色轨迹表示在智能体能够完成任务时所生成的轨迹，其余各颜色的轨迹表示智能体训练不充分时决策出的飞行轨迹。从图中可以看出，智能体存在逐步进化的现象。

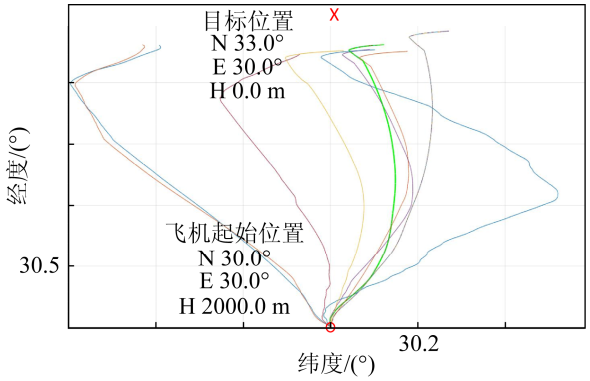

图 5 飞行轨迹对比图(地理坐标系俯视图)
Fig. 5 Comparison of flight trajectories (top view of geographic coordinate system)

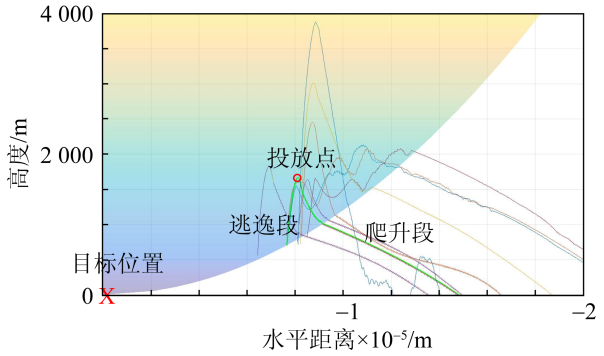

图 6 飞行轨迹对比图(复杂海上装备的坐标系 RH 图)
Fig. 6 Comparison of flight trajectories (coordinate system of complex maritime equipment)

对最终收敛结果进行详细分析，能够清楚看到目标自行迭代出的打击策略，在目标满足打击条件后尽早打击，在完成投放后迅速降高逃逸，以避免飞行器的打击。通过表 1 所列逃出探测区时间与生

第 33 卷第 12 期 系统仿真学报 Vol. 33 No. 12
2021 年 12 月 Journal of System Simulation Dec. 2021

存周期的对比可以看出，序号 14 逃逸的时间占比最少，也相对合理。

表 1 逃出探测区时间与生存周期对比
Tab. 1 Comparison of escape time and life cycle

序号	逃逸时间/s	生存周期/s	(逃逸时间/生存周期)/%
1	80	102.087 937	78.363 813
2	59	96.768 490 39	60.970 259 8
3	111	112.526 573	98.643 366 6
4	83	93.679 597 52	88.599 868 3
5	72	97.691 475 36	73.701 415 3
6	64	84.711 531 85	75.550 516 7
7	130	130.733 426 3	99.438 991
8	112	125.975 651 9	88.906 069
9	118	126.645 197 6	93.173 687
10	64	84.711 531 85	75.550 516 7
11	64	90.886 893 79	70.417 193 6
12	120	130.234 100 2	92.141 766 1
13	78	109.098 098 3	71.495 288 4
14	**52**	**91.600 691 48**	**56.768 130 4**

通过智能体飞行决策轨迹趋势能够直观看出，智能体能够通过降低高度躲避探测跟踪，并尽量深入到探测区内执行打击动作。同时为了确保生存，智能体在打击结束后会尽快降低高度，以躲避飞行器的打击。学习的结果收敛且基本满足预期。

2.2 双机场景

选取其中一个对抗场景，如图 7 所示。

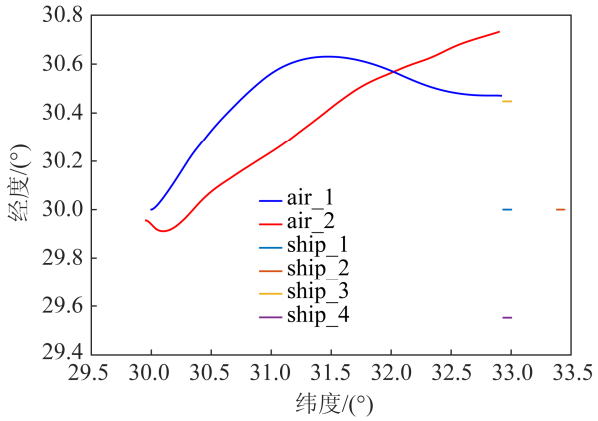

图 7 飞行轨迹对比图
Fig. 7 Comparison of flight trajectories

从图 7 可以清楚看到，因为复杂装备的反击策略是首先拦截距离较近的目标，因此飞机自行迭代出的打击策略是：通过蓝色目标 *air*_1 的前出诱骗，提供红色目标 *air*_2 的有效投放打击时间窗口与逃逸时间窗口。从整体策略达到的目的来看，初步具备了一定的协同性。

3 结论

针对传统模拟训练方法已经难以适应对抗场景日益复杂、协同决策日益明显等问题，本文面向基于严肃游戏的复杂海上装备对空方面的对抗场景，开展了并行分布式场景仿真和智能体模型的迭代训练。通过让智能体并行地在不同对抗场景下与复杂海上装备进行博弈，进而认知足够多的对抗样式，从而寻找不同对抗场景下的防御突破点，形成满足各对抗条件下的最优决策模型，初步实现了智能体灵巧策略的建模。后续将开展多智能体协作的建模，利用并行分布式的学习架构实现复杂对抗场景下的多目标协同建模。

参考文献：

[1] 钟华. 贴近实战的外军军事训练[J]. 国防科技, 2014, 35(4): 104-106.
Zhong Hua. Close to Actual Combat Military Training of Foreign Troops[J]. National Defense Science & Technology, 2014, 35(4): 104-106.

[2] 寇英信, 李战武, 李俊兵, 等. 现代战斗机作战任务管理与决策[M]. 北京: 国防工业出版社, 2017.
Kou Yingwin, Li Zhanwu, Li Junbing, et al. Modern Fighter Combat Mission Management and Decision-making[M]. Beijing: National Defense Industry Press, 2017.

[3] Poli R, Kennedy J, Blackwell T. Particle Swarm Optimization: An Overview[J]. Swarm Intelligence (S1935-3820), 2007(1): 33-57.

[4] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level Control Through Deep Reinforcement Learning[J]. Nature (S1476-4687), 2015, 518(7540): 529-533.

[5] Silver D, Huang A, Maddison C J, et al. Mastering the Game of Go with Deep Neural Networks and Tree Search[J]. Nature (S1476-4687), 2016, 529(7587):

484-489.

[6] Vinyals O, Babuschkin I, Czarnecki W M, et al. Grandmaster Level in StarCraft II Using Multi-agent Reinforcement Learning[J]. Nature (S1476-4687), 2019, 575(7782): 350-354.

[7] Silver D, Lever G, Heess N, et al. Deterministic Policy Gradient Algorithms[C]// International Conference on Machine Learning. PMLR, 2014: 387-395.

[8] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization[J]. arXiv preprint arXiv:1412.6980, 2014.

[9] Schulman J, Levine S, Abbeel P, et al. Trust Region Policy Optimization[C]// International Conference on Machine Learning. PMLR, 2015: 1889-1897.

[10] Lillicrap T P, Hunt J J, Pritzel A, et al. Continuous Control with Deep Reinforcement Learning[J]. arXiv preprint arXiv:1509.02971, 2015.

[11] Tesauro G. Temporal Difference Learning and TD-Gammon[J]. Communications of the ACM (S0001-0782), 1995, 38(3): 58-68.

[12] Bellemare M G, Dabney W, Munos R. A Distributional Perspective on Reinforcement Learning[C]// International Conference on Machine Learning. PMLR, 2017: 449-458.

[13] Barth-Maron G, Hoffman M W, Budden D, et al. Distributed Distributional Deterministic Policy Gradients[J]. arXiv preprint arXiv:1804.08617, 2018.

[14] Sergeev A, Del Balso M. Horovod: Fast and Easy Distributed Deep Learning in TensorFlow[J]. arXiv preprint arXiv:1802.05799, 2018.

https://dc-china-simulation.researchcommons.org/journal/vol33/iss12/8
DOI: 10.16182/j.issn1004731x.joss.20-FZ0532