

## 基于生成对抗网络的无人机博弈对抗技术\*

刘小峰<sup>1,2</sup> 王欢<sup>1</sup> 王子洋<sup>1</sup>

(1 河海大学物联网工程学院 江苏常州 213022)

(2 江苏省特种机器人技术高校重点实验室 江苏常州 213022)

**摘要:** 针对无人机(UAV)对抗环境,研究了基于生成对抗网络(GAN)的无人机博弈对抗技术。首先,确立了无人机智能体的六自由度模型及动作生成框架;然后,针对无人机对抗环境下的态势信息的缺失,给出了相应的态势信息重建算法;最后,针对无人机博弈对抗决策问题,提出了基于生成对抗网络的博弈对抗决策生成方法。

**关键词:** 博弈;无人机;生成对抗网络

**中图分类号:** V279 **文献标志码:** A **文章编号:** 1674-909X(2021)05-0001-05

## UAV Game Confrontation Technology Based on Generative Adversarial Network (GAN)

LIU Xiaofeng<sup>1,2</sup> WANG Huan<sup>1</sup> WANG Ziyang<sup>1</sup>

(1 College of IoT Engineering, Hohai University, Changzhou 213022, Jiangsu, China)

(2 Jiangsu Key Laboratory of Special Robots Technology, Changzhou 213022, Jiangsu, China)

**Abstract:** Aimed at the unmanned aerial vehicle (UAV) confrontation environment, the UAV game confrontation technology based on the generative adversarial network (GAN) is studied. Firstly, a 6-DOF (degree of freedom) model and an action generating framework of the UAV agents are established. Then, for the lack of situation information in the UAV confrontation environment, the corresponding situation information reconstruction algorithm is given. Finally, aimed at the decision-making problem for the UAV game confrontation, a game confrontation decision-making method is proposed based on the GAN.

**Key words:** game; unmanned aerial vehicle (UAV); generative adversarial network (GAN)

### 0 引言

现代战场的空战是敌我双方各自作战平台在己方地面指挥所、预警指挥机、电子干扰机和雷达等支援下进行的协同作战行动,战场形势复杂,态势多变。因此,现代空战战场环境是高智能强动态的复杂环境,对于战场信息进行准确获取、分析和预测,

进而做出最优决策,是决定胜负的关键所在<sup>[1-2]</sup>。

目前,国外已开展了大量无人机作战方面研究工作。JS McGrew等<sup>[3]</sup>提出一种基于近似动态规划的空战对抗策略;Perez-Leon等<sup>[4]</sup>研究了基于4D弹道作战背景下的大规模无人机冲突问题;Gaertner U等<sup>[5]</sup>开发了一种两级马尔可夫过程进行空战模拟;Nowak D J等<sup>[6]</sup>将仿生和遗传算法等技术融合,提出

\* 基金项目:国家重点研发项目(2018AAA0100800)、江苏省重点研发项目(BK20192004B和BE2018004)、广东省林业科技创新项目(2020KJCX005)和常州市国际合作与交流项目(CZ20200035)资助项目。

收稿日期:2021-08-10

引用格式:刘小峰,王欢,王子洋.基于生成对抗网络的无人机博弈对抗技术[J].指挥信息系统与技术,2021,12(5):1-5.

LIU Xiaofeng, WANG Huan, WANG Ziyang. UAV game confrontation technology based on generative adversarial network (GAN)[J]. Command Information System and Technology, 2021, 12(5): 1-5.

一种无人机群自主空战决策方法。

国内无人机作战研究近年来得到了蓬勃发展。陈侠等<sup>[7]</sup>提出一种基于模糊动态博弈的无人机空战决策方法;Luo等<sup>[8]</sup>提出采用混合自适应遗传算法解决空战中协同多目标攻击的决策问题,优化了无人机群空战的目标分配;杨啟明等<sup>[9]</sup>使用将部分可观测马尔可夫过程理论引入无人机路径规划问题,提高了路径规划的效率;MAO<sup>[10]</sup>等提出一种自适应启发式动态规划方法,提升了无人机对抗性能。上述方法均未能考虑无人机空战中的态势数据缺失问题,且不能利用已有的专家空战经验数据。

因此,本文针对无人机空战场景,提出了基于生成对抗网络(GAN)的无人机博弈对抗技术思路。首先,完成了无人机智能体建模;然后,将GAN<sup>[11]</sup>结合LSTM<sup>[12]</sup>用于战场缺失态势数据的重建;最后,提出采用生成对抗模仿学习(GAIL)方法从已有专家无人机对抗数据中学习专家策略,自动生成对抗决策方法<sup>[13-18]</sup>。

## 1 挑战与问题

在人工智能时代,数据量的大规模增长、机器智能的巨大飞跃和人机对抗等应用的飞速发展分别形成了信息大数据、算法高智能和对抗强动态的复杂环境。实现高智能、强动态环境下的无人机对抗决策面临以下挑战和问题:

1) 无人机智能体基本模型确立:为适应态势动态变化、障碍威胁多且信息获取不完全的动态环境,智能体需具备足够适用的动作策略和动作规则以提升自身在对抗过程中的生存力。因此,需确定智能体基本模型,建立完备的智能体动作库。

2) 战场缺失态势信息数据重建:战场环境具有高复杂性和瞬时性,如探测雷达受到电磁和高山障碍物等干扰,导致对抗双方获得的信息出现缺失异常,影响对敌方态势的预测,需根据对抗环境中量测数据的时序特性和相关性,对缺失数据信息进行重建。

3) 博弈对抗决策生成:战场态势变化莫测,要求智能体具有迅速的反应和决策能力。传统基于规则的决策方法需要设计者具有较强的专业背景,在高智能、强动态环境下,难以设计适应自主空战决策要求的对抗决策生成方法。因此,需设计一种能对环境自适应的对抗决策方法,能对战场态势做出准确预测并做出最优决策动作。

## 2 基于GAN的无人机博弈对抗

### 2.1 无人机智能体建模

本文在空战战场高智能、强动态的战场背景下,确立了无人机智能体六自由度模型,优化动作决策过程。首先,需以无人战斗机为对象,建立无人机智能体的运动学模型。建立无人战斗机模型是实现其空战决策规划的基础,无人战斗机的运动轨迹可以通过构造飞机质心运动学方程组进行描述。飞机质心运动学方程用于描述飞机质心在空间的位置与质心运动速度的关系,用来确定飞机质心在空间的位置随时间的变化。为了方便研究,作如下假设:

- 1) 无人机飞行中不考虑飞行阻力,速度仅由飞机机动动作决定;
- 2) 无人机飞行过程中不存在侧滑,侧滑角度为0;
- 3) 飞机运动时不计风速;
- 4) 飞机质量及重力加速度保持不变;
- 5) 视地球为惯性系统,忽略地球自转和公转的影响,即视地球为静止的;
- 6) 假设地球是平面大地,忽略地球曲率;

基于以上假设,可以得到简化的飞机质心运动学方程组如下:

$$\begin{cases} dv/dt = g(n_x - \sin\mu) \\ d\mu/dt = g(n_t \cos\gamma - \cos\mu)/v \\ d\varphi/dt = gn_t \sin\gamma/(v \cos\mu) \\ dx_g/dt = v \cos\mu \cos\varphi \\ dy_g/dt = v \cos\mu \sin\varphi \\ dz_g/dt = -v \sin\mu \end{cases} \quad (1)$$

其中, $v$ 为速度; $t$ 为时间; $g$ 为重力加速度; $n_x$ 为沿飞行速度方向切向的过载; $n_t$ 称为法向过载; $\mu$ 为俯仰角; $\gamma$ 为偏航角; $\varphi$ 为滚转角; $x$ 、 $y$ 和 $z$ 分别为无人机在以地面为参考系的三维坐标空间中的坐标值, $x$ 为正北方向, $y$ 为正东方向, $z$ 为垂直地面向上方向。

基于简化的无人机质心运动方程,本文设计了相应的动作库,包括水平机动和俯冲机动等。无人机智能体对抗动作生成框架如图1所示。为了优化决策层对智能体的控制过程,本文设计并建立了战术规划层、机动动作库与六自由度运动控制指令相互间的映射关系,可将战术规划层产生的机动动作组合转换成六自由度无人机的控制层指令形式,作为输入信息输入至自动驾驶仪回路,由此构成了无人机智能体对抗动作的生成框架。无人机智能体的

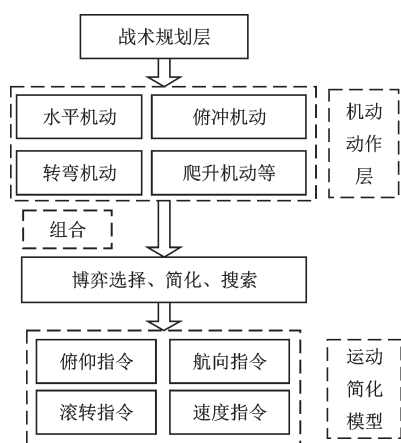


图1 无人机智能体对抗动作生成框架

每个战术动作均可简化为包含上述动作库中若干动作的组合序列。

## 2.2 基于LSTM-GAN的态势数据信息重建方法

实际的对抗博弈环境下,受到各种外部因素干扰及硬件设施的局限性,对于敌方无人机飞行态势数据的记录易出现缺失。例如,对于飞行轨迹相关数据(包括位置、高度、速度和航向等),正确的轨迹数据顺序为A-B-C,受到敌方仪器的干扰,雷达被屏蔽,数据出现缺失,仅采集到数据A-?-?,这对于后续对抗决策的生成可能产生一定影响。

GAN在图像修复方面已具备非常好的表现,但在时序数据修复方面应用相对较少。在GAN中,生成器 $G$ 试图生成欺骗判别器 $D$ 的数据,而判别器 $D$ 试图区分真实数据和生成的样本数据,网络根据判别器 $D$ 的输出进行更新,直至 $G$ 能够生成逼真的真实数据。考虑到飞行轨迹相关数据特有的时间相关性,将长短期记忆人工神经网络(LSTM)和GAN结合在一起对无人机缺失的态势数据进行填补。无人机的飞行态势数据包括无人机的位置坐标、姿态和速度等,其中位置坐标最关键,飞行姿态和速度等数据也具有一定的时间相关性,可用于辅助模型提高修复的精确度。LSTM能确保生成飞行数据的序列相关性,GAN保证了修复数据与已有数据的分布相似性,而相对于图像修复中庞大的图像数据,本文对抗场景中需处理的态势数据少,耗时短。因此,融合LSTM与GAN对态势信息数据进行修复具有可行性。态势数据信息重建方法结构如图2所示。

生成器 $G$ 输入条件为 $0 \sim T$ 时刻数据 $(x_0, x_1, \dots, x_T)$ ,通过LSTM提取时间特征,再经过全连接层后生成结果,即 $T+1$ 时刻数据 $x_{T+1}$ ;判别器 $D$ 输入条件为 $x_{T+1}$ ,经LSTM和全连接层处理后,输出对生成器 $G$ 生成结果的评估;根据评估结果

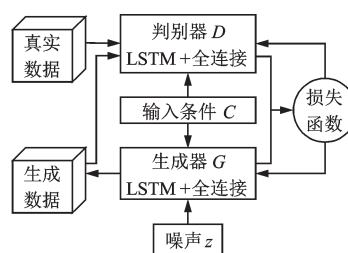


图2 态势数据信息重建方法结构

对网络更新,反复迭代收敛到纳什均衡,最终生成器 $G$ 生成接近真实样本的数据,得到最终训练模型。当需要进行态势信息重建时,将存在缺失态势信息的数据作为输入传入模型,得到对应的生成数据,即可完成对缺失部分态势数据的修复。

## 2.3 博弈对抗决策生成

针对高智能、强动态环境下博弈对抗决策生成问题,如何借鉴已有的专家博弈对抗经验快速完成博弈决策。本文提出将基于GAN的GAIL融入决策生成方法,用于从专家经验中学习相应的博弈对抗策略。该方法通过生成模型和判别模型的迭代训练,得到最终的生成模型,结合了深度强化学习和对抗学习的方法,利用现有的人类专家经验数据对模型进行训练,借鉴已有的无人机对抗策略,生成人机对抗博弈决策。训练中将训练目标分配到2个子网络上,生成器 $G$ 以重建后的态势信息作为输入生成决策数据,用判别器 $D$ 以区分在对抗环境历史数据下的生成策略是否符合真实专家数据特性,生成器 $G$ 需要不断改进生成数据来提升生成数据的逼真程度,判别器 $D$ 则需要不断提高自身的判别能力从而更准确地区分专家数据与生成数据,直至两者收敛达到纳什均衡。

根据上述思路,本文提出基于GAN的博弈对抗决策生成方法,该方法以GAIL方法为基础构建。GAIL是一种以GAN为基础的模仿学习方法。该方法可从现有专家数据中学习隐含的专家策略,能够将机器学习与人类经验有效结合,可借鉴专家经验的作用。基于GAN的博弈对抗决策生成方法框架如图3所示。判别器 $D$ 的损失函数可表示为:

$$L_w = E_{\tau}(\ln D_w) - E_{\pi_{\theta}}(\ln(1 - D_w)) \quad (2)$$

其中, $E_{\tau}$ 为专家策略; $D_w$ 为判别器输出; $E_{\pi_{\theta}}$ 为生成器输出。生成器 $G$ 的损失函数为:

$$L_{\theta} = E_{\pi_{\theta}}(\ln(1 - D_w)) \quad (3)$$

该方法采用近端策略优化(PPO)策略梯度网络与仿真环境(Gym Env)结合构成生成器网络 $\pi_{\theta}$ ,两者交互生成数据与已有的专家数据 $\tau$ 共同作为判别



器 $D_w$ 的输入,判别器 $D_w$ 对专家数据 $\tau$ 的采样与生成器生成数据对比得到相应的损失输出 $L_w$ ,根据 $L_w$ 对生成器网络 $\pi_\theta$ 进行更新,反复迭代至学习到专家数据隐含的专家策略,生成最终模型。对抗时,仅需将实施的态势数据输入生成的模型,即可生成战术规划决策,再将该决策输入建立的无人机智能体模型即可执行相应的机动动作,反复迭代至完成作战任务。

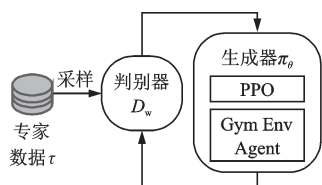


图3 基于GAN的博弈对抗决策生成方法框架

## 2.4 仿真试验

### 2.4.1 数据重建试验

试验采用无人机9轴传感器采集的时序数据用于训练模型,对无人机飞行姿态数据中偏航角、俯仰角和滚转角缺失的数据进行修复重建。

为了令采集的姿态信息更全面,采集数据过程中,无人机在飞行中会切换多种姿态。例如,在无人

机正常飞行过程中,偏航角变化较小,但有时会出现无人机环绕一周探查周围环境的情况,为兼顾到该类特殊情况,在采集数据集过程中,控制无人机在一定区域内原地旋转式飞行,偏航角会迅速变化。采集数据由多个连续采样点构成,采样频率100 Hz,共采集15万个样本点数据,选取前80%作为训练样本,后20%作为测试样本,并截取其中7000多个采样点用于展示修复结果。

数据修复结果如图4所示。对于真实数据,随机选取部分数据进行缺失,用绿色虚线标记出缺失数据的起始和结束位置,修复结果在上方小窗口中放大显示。图中,左侧小窗口对应修复片段1缺失数据的修复结果,右侧小窗口对应修复片段2缺失数据的修复结果;蓝色曲线表明真实数据,红色曲线表明修复生成的数据。修复试验中随机选取2段各50个采样点的数据进行缺失处理。由图4可见,本文方法能够较好修复出缺失数据的整体趋势,并使误差保持在较小范围内,具有较高修复精度,修复片段1对应于俯仰角、滚转角和偏航角3个角度的均方误差(MSE)分别为:1.590、0.298和1.149;修复片段2对应3个角度的MSE分别为:1.432、0.510和3.126。

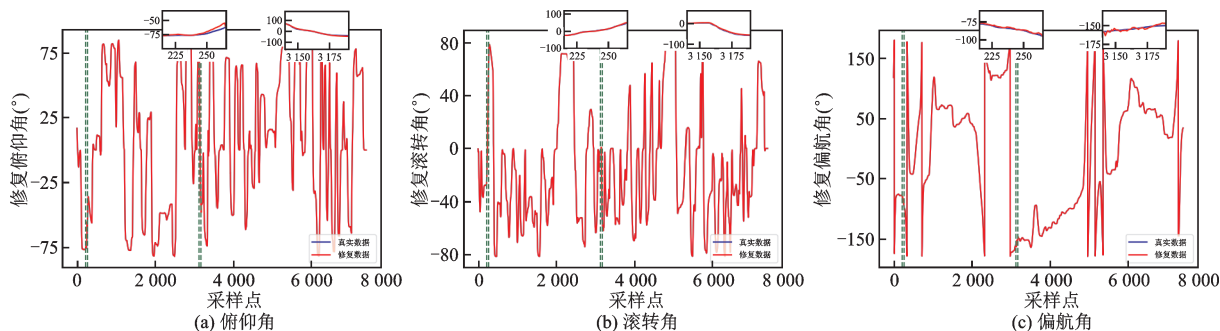


图4 数据修复结果

### 2.4.2 博弈决策生成试验

仿真场景设置为横向3 km、纵向2.5 km的场景,采用本文算法,在1v1无人机拦截对抗场景中进行仿真。敌方无人机采用基于马尔可夫决策过程(MDP)的控制方法,我方无人机采用本文方法,以修复重建后的态势数据训练模型,根据生成的模型实时生成战术规划决策,控制无人机完成对抗任务。

拦截对抗结果如图5所示。图中,红色五角星为敌方无人机攻击目标(我方军事基地),灰色圆形与方块为障碍物。由图可见,敌方无人机随机出现在地图左侧随机位置,按照不同攻击路径攻击我方军事基地,我方无人机均能按照较优的拦截路径完

成对敌机的拦截任务,验证了本文方法的有效性。

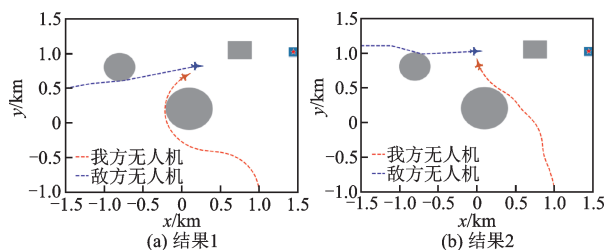


图5 拦截对抗结果

## 3 结束语

本文针对高智能强动态的无人机对抗环境,确

立了无人机智能体六自由度模型,提出使用LSTM-GAN对缺失战场态势信息进行修复重建,并结合GAIL方法学习已有专家的博弈对抗经验,实现博弈对抗决策生成,最后在仿真试验中验证了本文方法的有效性。本文方法主要有以下3个方面优点:1)针对复杂多变的战场环境,设计了简化的无人机智能体动作生成框架,能够快速准确地将战术规划层转换为底层的无人机飞行控制指令;2)针对战场博弈过程中存在的战场态势数据缺失问题,融合GAN与LSTM,能够自动完成对态势缺失数据的重建,为后续博弈对抗决策提供数据基础;3)将GAIL方法引入空战博弈对抗决策,使无人机智能体自主学习人类专家经验,生成博弈对抗策略,发挥了专家对抗经验在博弈决策生成中的借鉴作用。本文方法尚未考虑战场中其他因素(如风向和地理环境等),后续需进一步改进完善。

#### 参考文献(References):

- [1] 钟赞,张杰勇,邓长来.有人/无人机协同作战问题[J].指挥信息系统与技术,2017,8(4):19-25.
- [2] 高杨,李东生,程泽新.无人机分布式集群态势感知模型研究[J].电子与信息学报,2018,40(6):1271-1278.
- [3] MCGREW J S, HOW J P, WILLIAMS B, et al. Air-combat strategy using approximate dynamic programming[J]. Journal of guidance, control, and dynamics, 2010, 33(5): 1641-1654.
- [4] PEREZ-LEON H, ACEVEDO J J, MAZA I, et al. Integration of a 4D-trajectory follower to improve multi-UAV conflict management within the U-space context [J]. Journal of Intelligent & Robotic Systems, 2021, 102(3):62.
- [5] UWE G. UAV Swarm tactics: an agent-based simulation and Markov process analysis [EB/OL]. (2013-06-22)[2021-07-13]. <https://core.ac.uk/reader/36727281>.
- [6] NOWAK D J. Exploitation of self organization in UAV swarms for optimization in combat environments[EB/OL]. (2008-03-03)[2021-07-13]. <https://core.ac.uk/reader/288295184>.
- [7] 陈侠,赵明明,徐光延.基于模糊动态博弈的多无人机空战策略研究[J].电光与控制,2014,21(6):19-23.
- [8] LUO D L, SHEN C L, WANG B, et al. Air combat decision-making for cooperative multiple target attack: an approach of hybrid adaptive genetic algorithm [J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2006, 23(3):382-389.
- [9] 杨啟明,徐建城,田海宝,等.基于IMM的无人机在线路径规划决策建模[J].西北工业大学学报,2018,36(2):321-331.
- [10] MAO Y, CHEN Z J, YANG Y, et al. A novel adaptive heuristic dynamic programming-based algorithm for aircraft confrontation games [EB/OL]. (2021-08-25) [2021-08-30]. <https://doi.org/10.1016/j.fmre.2021.08.004>.
- [11] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal: ACM, 2014: 2672-2680.
- [12] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8):1735-1780.
- [13] HO J, ERMON S. Generative adversarial imitation learning [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona: ACM, 2016:4572-4580.
- [14] ANTOTSIU D, CILIBERTO C, KIM T-K. Adversarial imitation learning with trajectorial augmentation and correction [EB/OL]. (2021-03-26) [2021-07-11]. <https://arxiv.org/abs/2103.13887>.
- [15] CHOI S, KIM J, YEO H. Trajgall: generating urban vehicle trajectories using generative adversarial imitation learning [J]. Transportation Research Part C: Emerging Technologies, 2021, 128(6):1-22..
- [16] FU J, LUO K, LEVINE S. Learning robust rewards with adversarial inverse reinforcement learning [C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver: OpenReview, 2018:1-15.
- [17] WANG P, LIU D P, CHEN J Y, et al. Decision making for autonomous driving via augmented adversarial inverse reinforcement learning [EB/OL]. (2021-03-26) [2021-07-11]. <https://arxiv.org/pdf/1911.08044v3.pdf>.
- [18] 邢恩旭,吴小勇,李雅娴.基于迁移学习的双层生成式对抗网络[J].计算机工程与应用,2019,55(15):38-46.

#### 作者简介:

刘小峰,男(1974—),教授,研究方向为认知与机器人、人机交互和强化学习。

王欢,男(1986—),博士研究生,研究方向为多智能体博弈和强化学习。

王子洋,男(1999—),硕士研究生,研究方向为深度学习和人工神经网络。

(本文编辑:李素华)