基于预知博弈树的多无人机群智协同空战算法

周文卿¹, 朱纪洪^{2*}, 匡敏驰², 史恒²

1. 清华大学计算机科学与技术系, 北京 100084;

2. 清华大学精密仪器系, 北京 100084

* E-mail: jhzhu@tsinghua.edu.cn

收稿日期: 2021-07-15; 接受日期: 2021-10-14; 网络版发表日期: 2022-08-09

摘要 无人机广义上为不需要驾驶员登机驾驶的各式遥控飞行器. 在现代空战中, 无人机占据了越来越重要的地位. 然而目前关于无人机的空战算法的研究大都是在高度简化的低精度简单场景中进行的, 使用的方法也大都受限于已有的专家知识, 无法充分发挥智能空战算法的优势. 因此, 本文对基于预知博弈树的多无人机群智协同空战算法进行了研究. 首先使用Unity3D搭建了一套贴近于真实空战场景的仿真环境; 然后根据现有的空战知识封装了一套战术机动作, 实现了脚本化的飞机编队, 并设计了一套空战态势评估函数; 以此为基础提出了基于预知博弈树的空战AI算法框架, 通过预知博弈树算法完成了角色分配和机动作决策的任务, 使用XGBoost将其转化为一个在线的实时算法; 以状态机算法为baseline, 在高逼真度仿真平台上通过空战AI对抗实验, 验证了本文提出的空战算法的有效性.

关键词 多无人机群智协同, 智能空战, 预知博弈树算法, 高逼真度仿真平台, 自主控制

1 引言

无人机(unmanned aerial vehicle, UAV)广义上为不需要驾驶员登机驾驶的各式遥控飞行器. 无人机通常使用遥控、导引或自动驾驶的方法来进行控制, 在用途上则通常用于军事、科学研究、场地勘探、休闲娱乐等方面. 在军事方面, 无人机最早用作靶机或炸药的载具, 随后慢慢发展出了无人侦察机、诱饵机和具备对地攻击能力的作战飞机等. 随着近些年人工智能、5G通信和微电子技术的发展, 无人机在自主导航、精准定位、可靠通信、实时避障等方面都有了长足的进步. 而无人机由于其小型化、低成本、安全性和数量

规模优势, 在现代空战中占据了越来越重要的地位, 并逐渐发展出了蜂群战术和有人机/无人机群智协同等多种多样的作战方式^[1].

近年来, 国内外众多学者在无人机的机器视觉、自主飞行、集群算法、有人/无人协同和空战算法等方面做出了许多研究^[2,3]. 美国国防部高级研究计划局(Defense Advanced Research Projects Agency, DARPA)宣布其已经分别在2020年1月和9月完成了对于进攻性蜂群使能战术(offensive swam-enabled tactics, OFF-SET)的第三和第四次野外实验, 分别演示了异构蜂群无人机对于城区突袭任务和模拟目标保护的能力^[4]. 而针对无人机的自主控制, Alfeo等人^[5]对无人机在杂

引用格式: 周文卿, 朱纪洪, 匡敏驰, 等. 基于预知博弈树的多无人机群智协同空战算法. 中国科学: 技术科学, 2023, 53: 187–199

Zhou W Q, Zhu J H, Kuang M C, et al. Multi-UAV cooperative swarm algorithm in air combat based on predictive game tree (in Chinese). Sci Sin Tech, 2023, 53: 187–199, doi: 10.1360/SST-2021-0294

乱室外区域的目标搜索问题, 提出了一种利用生物激励机制协调无人机与不完善的传感器的方法. Varela 等人^[6]则提出, 利用空中平台监测无人机群, 探测环境条件以评价每个个体的状态, 并基于群体智能对无人机群进行实时协调.

国内方面, 孙佳琛等人^[7]建立了面向无人机集群的群体智能协同通信模型, 并阐述了无人机群的立体协同感知技术. 北京航空航天大学段海滨团队^[8,9]基于对鸽群的模拟和学习, 分别对仿生鸽归巢的导航技术和有人/无人机集群的一致性控制进行了研究. 胡多修等人^[10]则提出了一种改进的近端策略优化(proximal policy optimization, PPO)算法, 对无人机的自主引导跟踪与避障进行了研究. 在2020年10月, 中国电子科学研究院进行了陆空协同固定翼无人机蜂群系统的验证试飞, 标志着中国首个实用化无人机蜂群的诞生^[11].

而在无人机的空战算法方面, 2016年辛那提大学的Ernest等人^[12]研发了基于遗传模糊树的空战智能体AlphaAI, 其操纵F-15击败了人类退役空军上校驾驶的F-22. 而在DARPA于2020年举办的AlphaDogfight近距离格斗比赛中, 苍鹭公司使用长短期记忆网络(long short-term memory, LSTM)的算法实现了一种空战智能体, 并最终击败了其他一众AI以及人类飞行员^[13]. 国内方面, 有研究者使用深度强化学习的方法, 在二

维兵棋推演平台上对多机协同空战进行了研究^[14]. 此外, 极大极小搜索的算法^[15]和状态机的方法^[16]也被使用在空战机动决策的研究中.

然而上述对于无人机空战算法的研究, 大都是在经过大量简化的低精度简单场景中进行的, 与实际的应用场景相距较大^[12~15], 使用的方法也较为简单或者受限于已有的专家知识^[15,16]. 因而根据上述需求, 我们针对更为复杂真实的仿真环境, 提出了一种基于预知博弈树的多无人机群智协同空战算法, 并对其进行了仿真验证.

2 空战对抗模型

从引言中的介绍可以发现, 现有的对无人机空战算法的研究中, 很多工作是在低精度甚至二维的环境中进行的, 与实际的场景相去甚远. 为了能够研究出在应用方面具有实际意义的算法, 我们首先使用游戏开发引擎Unity3D搭建了一套贴近于真实空战场景的仿真环境. 高逼真度仿真平台中包括了8个部分——即空气动力学、红外、雷达、诱饵弹等空战引擎内核部分和视界与UI、VR、天气与地景、多功能电子触摸屏等人机交互部分.

基于高逼真度仿真平台, 我们设置了图1所示的空

图1 空战实验场景

Figure 1 Scene of air combat experiment.

战实验场景. 实验环境的地景为海中央的小岛, 最高海拔小于4 km, 作战区域则为地景上方的空域, 大小为150 km×150 km×20 km. 如果飞机或导弹在对抗过程中与地景相撞则会立即被销毁. 仿真平台中对战的双方为数量相同的F22战机, 考虑到仿真占用的内存和算法搜索空间限制, 战机规模控制在1架对1架到9架对9架之间. 每架飞机分别挂载两枚AIM-120中距弹、六枚AIM-9X近距弹和八枚诱饵弹, 导弹有一定概率被诱饵弹引走丢失目标.

高逼真度仿真平台中, 最核心的是飞行器的六自由度气动力学建模. 将飞机分为机身和机翼, 机翼分为主翼、副翼、襟翼、前缘襟翼和尾翼. 飞机在飞行时受到的空气动力主要来自于机翼, 因此对所有翼面进行划分和有限元分析模拟, 得到其升阻力系数. 综合操纵杆和节流阀的输入, 得到各个翼面产生的升阻力、机体阻力、发动机推力和力矩, 可以计算出飞机所受的合力和力矩的方向、大小. 升力的计算公式^[17]为

$$L = qS_{\text{ref}}C_L, \quad (1)$$

式中 L 为升力, q 为动压, S_{ref} 为参考面积, 一般指机翼面, C_L 为升力系数. 阻力的计算公式^[17]为

$$D = qS_{\text{ref}}C_d, \quad (2)$$

式中 D 为阻力, q 为动压, C_d 为阻力系数. 力矩的计算公式^[17]为

$$M = \frac{1}{2}\rho v^2 S_{\text{ref}} C_m l_{\text{ref}}, \quad (3)$$

式中 M 为力矩, ρ 为空气密度, v 为空速, C_m 为力矩系数, l_{ref} 为参考长度, 一般指弦长. 类似地, 可以计算得到导弹的升阻力、推力和力矩, 根据上面求得的这些参数, Unity3D中的物理引擎RigidBody可以方便地对飞行器的位置、速度、加速度以及姿态等信息进行实时更新.

系统中还包括了红外系统、雷达系统和诱饵弹等设置. 其中, 雷达系统对目前先进战斗机采用的有源相控阵雷达进行了模拟. 而雷达系统中的目标探测跟踪子模块会将所有可被探测、跟踪到的目标的位置和速度信息传递给AI模块. 若敌方目标不在探测范围内, 或雷达系统受到电磁干扰、诱饵等的影响, 则不能准确获得目标的信息. 相较于文献[12,13]等完备信息的环

境, 基于本文方法所研制的仿真平台能够使得使用系统的人类飞行员和空战算法都能在完全模拟实战的环境下进行战斗.

此外, 如图2所示, 视景与UI、VR、天气与地景、多功能电子触摸屏等子系统能够为人机交互提供便利, 方便AI编写时的调试和以后人机对抗的测试.

3 动作集封装

对于空战算法而言, 决策空间相当巨大, 而其中的有效解非常的稀疏, 根据基本的空战技战术可以发现可选的机动动作大部分都没有正面收益甚至会造成负面的影响. 因此, 为了压缩决策空间、降低算法的复杂度, 同时避免决策时的波动, 我们结合现有的空战知识设计了一套战术动作. DCS World是俄罗斯军用飞行模拟团队开发的一套较为逼真的空战仿真游戏平台, 目前还承担着俄罗斯军方空战AI的开发测试任务. 我们与国内的顶级DCS World玩家团队合作, 获取了大量顶级玩家的对战数据. 通过对数据的分析, 我们提取空战中的人类经验, 将复杂的机动分解成几种简单的动作, 并设计了如表1所示的一套机动动作集.

其中, 前3种动作为飞机向正前方飞行, 同时, 受目标俯仰角的影响, 其又分为直飞、拉升和俯冲. 盘旋是指飞机滚转后拉升机头执行转弯操作, 参数包括目标滚转角和操纵杆的俯仰输入. 追踪是指飞机对目标进行跟踪, 目标可能是飞行中的敌机, 也可能是给定的某一个预测点(如跟丢目标时, 可根据丢失视野前敌机的速度和加速度, 对目标可能的位置进行预测). 筋斗则是指固定操纵杆俯仰输入, 使飞机的俯仰角剧烈地变化. 通过4和6两个动作的组合, 可以使得飞机完成

图2 VR看透座舱功能

Figure 2 See through the cockpit with VR.

表 1 机动动作集

Table 1 Maneuver set

编号	机动动作	参数1		参数2	
		符号	含义	符号	含义
1	直飞	Θ_T	目标俯仰角	—	—
2	拉升	Θ_T	目标俯仰角	—	—
3	俯冲	Θ_T	目标俯仰角	—	—
4	盘旋	Θ_T	操纵杆俯仰输入	ϕ_T	目标滚转角
5	追踪	T	追踪目标	—	—
6	筋斗	Θ_T	操纵杆俯仰输入	—	—
7	发射导弹	—	—	—	—
8	发射诱饵弹	—	—	—	—

滚转180°再拉俯仰调转方向至水平的动作,实现对导弹的规避机动。最后两个动作则是发射导弹攻击敌机和发射诱饵弹,诱饵弹有一定的概率将敌方导弹引走。根据操纵杆的输入,我们采用PID控制器进行控制,使得飞机能够达到机动动作集中设置的目标滚转角或目标俯仰角。而操纵杆俯仰输入则是控制操纵杆保持一个固定的输入。此外,通过节流阀还可以控制飞机的加速度,使其加速、减速或匀速运动。

与经典的NASA提出的7种基本机动动作^[18]和细化扩展得到的25种典型战术动作^[19]相比,表1中的动作更加的简洁、基础。虽然数量不多,但是通过参数的变化、节流阀的变化和动作的组合,能够灵活地生成文献[18,19]中的复杂机动动作和文献中没有的其他动作。通过这样的方式,我们不仅可以在降低搜索空间的同时满足空战AI机动动作的丰富程度,还可以减少人为的限制,使其能够通过搜索得到更加有效的动作组合。

4 基于预知博弈树的空战算法

在空战中,飞机首先需要从起飞地点以编队的形式抵近敌机(或其他任务空域)。在接近敌方目标后,每架飞机需要根据分配的角色和战场情况决定所要执行的战术动作和动作强度。在此过程中,人类飞行员驾驶的有人机还可以担任空中指挥官的角色,与无人机群形成人机协同。有人机能够在空战中汇总所有无人机探测到的信息,完成对战场全局的把控和无人机决策的审核任务,而较少地甚至不承担单机作战任务。如

图3所示,我们采用脚本化的方式对无人机群进行编队,而对于战术协同层和单机决策层则综合了蒙特卡罗搜索树算法和对空战场未来态势的预测和评估,提出了一种预知博弈树的算法(图4)。

对于编队模式和各编队数量的决策,由于没有很多的选择且目前已有较为简单明确的战术规则,我们根据经典的编队模式,使用脚本化的方法直接给出机动策略。战机以不同的高低、左右、前后多维度地展开以迎敌。每个小队一般由两架飞机构成,互相之间的距离保持在6 km以内。不同小队之间的间距则视作战任务和战场态势不同,保持在6~28 km之间,保证不会使得多个小队同时受到敌方武器的威胁,又能在短时间之内快速支援。编队中的飞机数量和编队的任务与敌机的分布有关——执行主要任务的编队规模较大,而承担诱敌、掩护等任务的编队规模则相对较小。编队中的飞机应在局部空战场中形成数量优势,完成自己的任务后再与其他编队会合,以持续保持局部的优势。

飞机以脚本化编队的方式抵达作战空域后,我们提出了一种预知博弈树的算法来对飞机的任务角色和具体的机动动作进行预测、搜索(图5)和决策。预知博弈树算法的本质,是通过未来的推演对当前战场态势做出评估,因此这里的预测搜索过程和前面仿真平台中的仿真模拟是不完全相同的。预知博弈树算法中的初始战场状态取决于仿真平台中空战AI能获取到的信息,是不完备的。但是在后续的模拟、搜索过程中,应该认为初始场景中可探测到的敌方目标对于AI都是可知的,而初始场景中探测范围外的目标则不会

图 3 多机展开编队模式

Figure 3 Multi UAV formation mode.

图 4 基于预知博弈树的空战算法

Figure 4 Air combat algorithm based on the predictive game tree.

被考虑。这样才能保证对当前战场态势做出最准确的评估。同样的, 当停止搜索, 返回评估函数时, 此时的评估函数可获取到的信息也应该包括预知博弈树初始状态下可探测到的所有敌方目标的信息(此外还有模拟过程中新生成的目标)。

根据当前战场状态, 基于预知博弈树的多无人机

群智协同空战算法可以对未来进行预测。算法控制的每一架飞机, 在一步模拟中(模拟步长为0.02 s)分为战术协同层和单机决策层两部分。战术协同层中, 包含了空战中飞机的不同任务角色(探测、突击、回避、掩护等)对应的子节点; 而单机决策层中则根据分配的任务角色, 包含了每一架飞机的战术动作和动作强度对

应的子节点. 因此, 若己方有 n 架飞机, 则在一步模拟中就产生了一棵 $2n$ 层的搜索树, 而从根节点到叶子节点的一条通路对应着在这一步中所有己方飞机的一组决策. 另一方面, 敌机使用同样的方法对搜索树进行扩展, 与我方飞机做出的决策组一同进行仿真计算, 即可得到该决策组下 0.02 s后的战场态势. 重复这样的过程, 我们就可以实现对未来的预测, 并在不同的决策组中选择最优的一组.

预知博弈树的底层是蒙特卡罗搜索树, 其每一轮搜索包括选择、扩展、模拟和回溯四个步骤. 首先如公式(4), 综合此前访问的胜率 and 访问的频率, 根据信心上界树(upper confidence bound for tree, UCT)算法选择当前最值得访问的节点. 式中 s_i 为第 i 个节点的UCT值, w_i 为第 i 个节点当前胜利的场数, n_i 为第 i 个节点当前的对战场数, N 为当前总的对战场数, C 为设定的常数, C 越大博弈树越倾向于对之前较少访问的节点进行搜索, C 越小博弈树则越倾向于访问之前胜率高的节点.

$$S_i = \frac{w_i}{n_i} + C \sqrt{\frac{\ln N}{n_i}}. \quad (4)$$

重复上述的选择过程, 直到到达一个可扩展节点. 一个节点是可扩展的, 当且仅当其所对应的状态是非停止状态, 且拥有未被访问过的子状态. 如图5中的步骤(a), 最终选择到左下角 $3/3$ 的节点. 对该节点进行扩展, 如步骤(b)向 $3/3$ 节点下面添加一个子节点以拓展搜索树. 然后是模拟, 我们使用文献[16]中的状态机算法代替随机决策, 快速地向后搜索直到一方战败或者满足其他终止条件(如达到一定时间)为止. 由于空战的

持续时间可能较长, 甚至会双方都打空导弹或都丢失目标且无法再次捕获目标而形成和局, 可能无法达到一方战败的状态, 因此需要加上时间限制. 已知状态机AI之间的对战一般持续十几秒就会结束, 所以我们规定, 如果向后模拟了750步也就是15 s之后, 仍然没有一方被完全击毁, 则会停止搜索, 并将后面的评估函数作为结果返回. 最后一步是回溯, 根据图5(c)的模拟我们得到了胜负的结果或是对终止状态下态势的评估, 从扩展的节点将该结果回溯地向上更新, 直到根节点的下一层.

通过上面的四步, 我们更新了从第一层到拓展节点这条路径上的对战场数和胜利的场数(终止状态下的评估函数值经过正则化后也可以认为是胜利的场数). 重复图5中的步骤, 可以估计出第一层所有节点获胜的概率, 即可选择获胜概率最大的作为最终的决策.

由于搜索空间非常大, 所以我们还加入了一些剪枝来提升搜索的效率:

- (1) 防撞地剪枝: 当飞机面临与地面、山体等相撞的风险时, 将不进行预知博弈树的搜索, 直接执行拉升的机动动作;
- (2) 脱离战场剪枝: 当飞机挂载的导弹全部发射, 且当前没有敌方导弹锁定本机时, 也无需进行预知博弈树的搜索, 沿固定的方向直飞脱离战场;
- (3) 索敌剪枝: 当视野中没有敌机且没有敌方导弹锁定时, 若在5 s内曾经观测到过敌机(即丢失了敌机的视野), 则根据观测的位置、速度、加速度预测其现在的位置, 并对预测出的坐标点执行追踪机动. 否则执行

图5 预知博弈树的搜索步骤

Figure 5 Search steps of the predictive game tree.

盘旋机动搜索敌机;

(4) 发射导弹剪枝: 对于双方飞机不符合发射距离或角度限制, 或者距离上次发射时间间隔小于 2 s 的情况, 在搜索中不考虑发射导弹的动作。

除了上述剪枝以外, 在对未来的预测中, 如果导弹丢失了目标也不再对其进行仿真。这样也能够在一定程度上降低算法的算力损耗, 提高预知博弈树算法的效率。

4.1 空战态势评估函数

空战态势评估的核心是对于全向攻击导弹攻击区的判定。我们模拟了人类飞行员使用武器时的思维模式, 用距离和角度等物理量来进行评估而不去计算导弹的包线等复杂信息, 这样能在保证评估效果的同时压缩所需的计算量。

评估函数分为两部分, 一个是我方飞机相对敌机的位置优势评估。其又可分为距离和角度两部分, 距离的优势如图6所示, 使用一个分段函数进行模拟。我们认为 500 m 是导弹的最小有效攻击射程; 20 km 为不可逃逸射程, 此处距离优势函数返回 0.5; 20 km 外距离优势函数呈反比例函数递减。距离优势函数的公式如下:

$$L = \begin{cases} \frac{D}{500}, & 0 < D \leq 500, \\ \frac{39500-D}{39000}, & 500 < D \leq 20000, \\ \frac{10000}{D}, & D \geq 20000, \end{cases} \quad (5)$$

式中, D 表示目标到我方飞机的距离, L 表示我方飞机相对目标的距离优势。角度的优势则如图7所示。

角度的优势使用正态分布的概率密度函数进行模拟。我们通过一些关键点对参数进行调整——角度为 0° 时角度优势函数应该返回 1, 角度为 $\pm 180^\circ$ 时返回值应近似为 0。而我们设定角度为 $\pm 30^\circ$ 时, 角度优势函数返回值为 0.5。由此可以得出角度优势函数的公式为

$$A = e^{-\frac{\theta^2}{1300}}, \quad -180^\circ \leq \theta \leq 180^\circ, \quad (6)$$

式中, θ 表示目标到我方飞机的连线与我方飞机轴线方向之间的夹角, A 表示我方飞机相对目标的角度优势。经过计算可以验证, 当角度为 $\pm 180^\circ$ 时, 公式(6)中计算的结果为 1.5×10^{-11} , 符合近似为 0 的要求。

评估函数的另一部分则是敌方导弹对于我方飞机的威胁评估。为了鼓励 AI 最大限度地利用好每一枚导弹, 避免短时间内发射多枚造成攻击能力的浪费, 我们仅考虑飞机当前面临的威胁最高的导弹。图8中的导弹威胁函数大致是将图6中的距离优势函数沿坐标轴翻转得到的, 但是要注意的是, 对于这个威胁评估, 我们不再需要考虑最小有效攻击射程。敌方导弹离我

图6 距离优势函数 L

Figure 6 Distance advantage function L .

图7 角度优势函数 A

Figure 7 Angle advantage function A .

图8 导弹威胁函数 T

Figure 8 Missile threat function T .

方飞机越近, 威胁就越大. 导弹威胁函数的公式为

$$T = \begin{cases} \frac{D}{40000} - 1, & 0 \leq D \leq 20000, \\ -\frac{10000}{D}, & D > 20000, \end{cases} \quad (7)$$

式中, D 表示我方飞机到敌方导弹的距离, T 表示敌方导弹对我方飞机的威胁程度. 综合上述的两部分, 可以得到完整的评估函数

$$\begin{cases} V = \sum_{j=1}^m \left(\lambda T_j + \eta \sum_{k=1}^n L_{jk} \cdot \theta_{jk} \right), \\ \text{Value}(R) = V(R) - V(B), \\ \text{Value}(B) = V(B) - V(R), \end{cases} \quad (8)$$

式中, m 和 n 分别表示我方和敌方的飞机数量, $\text{Value}(R)$ 和 $\text{Value}(B)$ 表示红方和蓝方最终的态势评估函数值. λ 是勇敢值参数, 可以用于调整AI是激进进攻的类型还是保守的类型. η 则表示当前我方飞机是否还有导弹剩余, 如果无法发射导弹击落敌机, 则我方飞机相对敌机就不存在位置优势. 其中, 我方和敌方的飞机, 都是预知博弈树初始场景中可探测到的; 导弹则是初始场景中可探测到的, 或是在模拟过程中发射的. 初始不可见的目标在整个模拟和评估的过程中都不会被考虑.

根据上面的方法, 我们可以使用预知博弈树算法, 在每一种空战的态势下进行预测和模拟, 估计出在当前态势下选择每一种机动动作的情况下对应的胜率, 并选择其中胜率最高的任务角色和机动动作作为下一步的决策.

表 2 表征空战态势的状态量

Table 2 State representing air combat situation

类型编号	参数	参数解释
1	姿态角	本机的姿态角
2	敌机距离	本机与目标之间的距离
3	敌机方位	目标位置相对本机速度方向的夹角
4	本机速度	将本机速度沿本机与目标的连线进行分解, 得到相对连线的径向速度和法向速度
5	高度	本机的海拔高度
6	距撞地时间	延长本机速度方向直到与地面地景相交, 连线长度除以速度得到距离撞地的时间
7	敌方导弹距离	对于锁定本机的敌方导弹, 计算其中威胁度最高的导弹的距离
8	敌方导弹角度	威胁度最高的导弹相对本机速度方向的夹角
9	敌方导弹速度	威胁度最高的导弹相对与本机连线的径向速度和法向速度
10	剩余导弹	本机剩余的导弹的数量

4.2 实时算法

尽管采取了一些剪枝的措施, 但是前面的算法消耗的时间依然很长, 每一步的搜索和决策过程需要 5 min的时间. 我们的仿真环境是使用Unity3D搭建的, 其物理引擎的仿真步长为0.02 s, 而我们设置空战AI的决策步长为0.5 s. 可以发现决策所需的时间远长于仿真步长和决策步长, 甚至长于整场空战的时间, 无法做到实时决策, 前面所述的搜索树只能进行离线的决策. 为此, 我们通过多轮的运行, 利用离线的方法生成了一个关于态势及对应的决策的数据集. 然后使用XGBoost的方法, 对离线方法得到的数据集进行学习. 也就是通过监督学习, 对离线预知博弈树算法进行了一个近似, 从而将预知博弈树的算法由离线算法变为一个在线的实时算法.

我们随机初始化飞机的高度、位置、朝向和使用的AI类型(状态机AI或离线预知博弈树AI), 生成了三千多场不同初始场景的空战数据. 挂载离线预知博弈树AI的飞机, 每隔0.5 s进行一次决策并在下一次决策前都保持同样的机动动作(0.5 s指仿真环境中的时间而不是现实中的时间). 将每次决策时, 表2中的各个状态的值和离线预知博弈树最终的决策结果保存下来, 得到了一个二十多万条数据的数据集.

生成数据集后, 我们使用XGBoost对其进行监督学习. XGBoost是一个开源机器学习项目, 其对梯度提升决策树(gradient boosting decision tree, GBDT)在工程实现上做了大量的优化. 其中Boosting是一种集成学习的方法, 它串行地训练多个不同的基分类器, 每

一轮训练结束之后, 对该轮中基分类器分错的样本, 给予更高的权重. 测试的时候则根据各轮训练的分类器的结果加权得到最后的结果.

同样的, XGBoost在训练的过程中, 不断添加新的回归树. 每添加一个新的回归树, 就是学习了一个新的函数来拟合此前预测的残差, 有

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i), \quad (9)$$

其中 $\hat{y}_i^{(t)}$ 为第 t 次迭代后样本 i 的预测结果, $f_t(x_i)$ 为第 t 棵树对样本 i 的预测结果. XGBoost的目标函数为

$$f^{(t)} = \sum_{i=1}^n l(\hat{y}_i^{(t)}, y_i) + \sum_{j=1}^t \Omega(f_j), \quad (10)$$

其中 n 为样本的数量, t 为回归树的数量, C 为常数, $l(\hat{y}_i, y_i)$ 为损失函数, $\Omega(f)$ 为抑制模型复杂度的正则项, 防止模型过拟合:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2, \quad (11)$$

其中 T 为第 t 棵树中叶子节点的个数, ω_j^2 为其中第 j 个叶子节点的得分的L2正则化, γ 为叶子节点数量控制系数, λ 是叶子节点分数控制系数.

由于XGBoost是一种Boosting的方法, 每一轮训练分步向前, 因此可以将第 t 轮的模型单独出来:

$$f^{(t)} = \sum_{i=1}^n l(\hat{y}_i^{(t-1)} + f_t(x_i), y_i) + \Omega(f_t) + \sum_{j=1}^{t-1} \Omega(f_j), \quad (12)$$

公式(12)中只有一个变量 $f_t(x_i)$, 其余的都是可以计算得到的常数项. 忽略公式(12)中的常数项, 并进一步对目标函数进行泰勒展开(损失函数以平方损失函数为例), 可得到近似的目标函数:

$$\begin{cases} f^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t), \\ g_i = \frac{\partial l(\hat{y}_i^{(t-1)}, y_i)}{\partial \hat{y}_i^{(t-1)}} = -2(y_i - \hat{y}_i^{(t-1)}), \\ h_i = \frac{\partial^2 l(\hat{y}_i^{(t-1)}, y_i)}{\partial (\hat{y}_i^{(t-1)})^2} = 2, \end{cases} \quad (13)$$

其中, g_i 和 h_i 分别是损失函数的一阶导数和二阶导数. 根据公式(13)中的目标函数和贪心算法不断生成新的树, 最终将每一轮生成的回归树的预测结果加权平均. 使用XGBoost对于上面生成的数据集进行学习后, 准确率可以达到99%以上.

综合脚本化的编队方法, 对战术协同层和单机决策层使用上面的实时预知博弈树算法进行搜索和决策, 我们实现了完整的基于预知博弈树的多无人机群智协同空战算法.

5 仿真验证

图9中展示的是本文仿真验证的总体框架. 在前述高逼真度仿真平台的基础上, 我们根据飞行控制知识和已有的一些空战技战术封装了一套动作集, 然后以

图 9 研究内容总体架构

Figure 9 Framework of our research.

其为基础设计了我们的空战AI. 以文献[16]中的状态机算法作为baseline, 在仿真平台中进行了验证和对抗实验.

如图10所示, 我们选取不同的空战规模(如第2节中所述, 双方各自操纵1架到9架F22)、初始位置、高度、偏航角配置了一千多种不同的空战实验场景, 进行了大批次的对战以测试我们的算法的效果. 对于空战AI对抗实验, 我们规定双方分别使用本文提出的基于预知博弈树的空战算法和文献[16]中提出的基于状态机的空战算法. 双方使用雷达来扫描、锁定敌方飞机, 未被雷达扫描到的敌方单位不会被飞机感知到, 己方飞机之间通过数据链可以共享探测到的信息. 对抗实验的终止条件为一方的飞机被全部击落, 或者5 min内双方都仍未决出胜负(此时判定为平局).

图11中为一场预知博弈树AI与baseline状态机AI的一对一对抗实验. 上面为战场整体的态势图, 下面为红方飞机的第一视角. 其中红色曲线为状态机AI的轨迹线, 蓝色曲线为预知博弈树AI的轨迹线, 绿色曲线和黄色曲线分别为其各自发射的导弹的轨迹线. 可以看到, 预知博弈树AI以较小的机动躲避状态机AI的导弹后持续追击, 最终将其击落.

图12中则是一场三对三的空战AI对抗实验, 对战双方都是预知博弈树AI. 红蓝双方通过脚本化的方式形成编队进入战场, 随后通过预知博弈树进行搜索, 完成任务分配和机动动作的决策, 形成两块小型的局部战场, 在空中缠斗.

图13中展示了预知博弈树算法与baseline对战时一些主要参数的变化情况. 算法耗时方面, 本文中方法的耗时大部分在2~3 ms之间, 虽然比baseline的耗时稍长, 但相比0.5 s的决策间隔和0.02 s的仿真步长短很多, 可以验证算法的实时性. 高度和速度方面, 本文中方法的高度和速度更加的稳定, 相比于人类飞行员采用三九线机动或其他大幅度机动躲避导弹, 基于预知博弈树的AI倾向于通过对未来的预测, 采用小幅度的机动躲避导弹. 这样的机动避免了过度躲避, 相比人类飞行员的操作更加精确, 能减少高度优势的损失. 过载方面, 本文中方法的过载变化幅度稍大, 但整体仍在合理的范围内. 通过灵活的机动, 能够让飞机处于更具有优势的位置.

在共计1152场博弈树AI与状态机AI之间的对抗实验中, 基于预知博弈树的算法获得了其中700场的胜

图10 九对九空战实验配置

Figure 10 9v9 air combat experiment configuration.

图11 一对一空战AI对抗实验

Figure 11 1v1 air combat AI confrontation experiment.

利, 基于状态机的算法获得了其中451场的胜利, 还有1场双方打成了平局, 本文提出的方法相对于baseline, 胜率达到了近61%.

6 结论

本文针对多无人机群智协同空战的算法进行了相关研究. 搭建了一套高逼真度仿真平台, 并根据现有的空战技战术封装了一套动作集; 提出了基于预知博弈树的空战AI算法框架, 实现了脚本化的飞机编队, 设计了一套空战态势评估函数; 据此完成了角色分配和机动决策的任务, 实现了基于预知博弈树的多无人机群智协同空战算法, 并使用XGBoost将其转化为一个

图 12 多对多空战AI对抗实验

Figure 12 Many-to-many air combat AI confrontation experiment.

图 13 空战中主要参数分析

Figure 13 Analysis of main parameters in air combat.

在线的实时算法; 在高逼真度仿真平台上通过空战AI对抗实验将上述算法与状态机AI进行对比, 验证了本文提出的空战算法的有效性.

另一方面, 由于算力的限制, 目前的方法提供的可

选动作依然较少, 因此对于作为baseline的状态机AI还没有形成压倒性的优势. 在未来的研究中, 需要进一步地扩充AI可选的行为模式, 更加准确地还原空战的真实情况, 提升AI在空战中的能力.

参考文献

- 1 Duan H B, Shen Y K, Zhao Y J, et al. Review of technological hotspots of unmanned aerial vehicle in 2020 (in Chinese). *Sci Tech Rev*, 2021, 39: 233–247 [段海滨, 申燕凯, 赵彦杰, 等. 2020年无人机热点回眸. *科技导报*, 2021, 39: 233–247]
- 2 Xin X, Xie C. Application and development of artificial intelligence in UAV system (in Chinese). *Aerodyn Missile J*, 2021, 5: 48–51 [辛昕, 谢川. 人工智能在无人机系统中的应用与发展研究. *飞航导弹*, 2021, 5: 48–51]
- 3 Duan H B, Qiu H X, Chen L, et al. Prospects on unmanned aerial vehicle autonomous swarm technology (in Chinese). *Sci Tech Rev*, 2018, 36: 90–98 [段海滨, 邱华鑫, 陈琳, 等. 无人机自主集群技术研究展望. *科技导报*, 2018, 36: 90–98]
- 4 Unmanned Systems Technology. DARPA conducts ourth OFFSET autonomous swarm experiment. (2020-01-03). <https://www.unmannedsystems-technology.com/2020/09/darpa-conducts-fourth-offset-autonomous-swarm-experiment>
- 5 Alfeo A L, Cimino M G C A, De Francesco N, et al. Swarm coordination of mini-UAVs for target search using imperfect sensors. *Intell Decision Technol*, 2018, 12: 149–162
- 6 Varela G, Caamamo P, Orjales F, et al. Swarm intelligence based approach for real time UAV team coordination in search operations. In: *Proceedings of the 3rd World Congress on Nature and Biologically Inspired Computing*. Salamanca, 2011. 365–370
- 7 Sun J C, Wang J L, Chen J, et al. Cooperative communication based on swarm intelligence: Vision, model, and key technology (in Chinese). *Sci Sin Inf*, 2020, 50: 307–317 [孙佳琛, 王金龙, 陈瑾, 等. 群体智能协同通信: 愿景、模型和关键技术. *中国科学: 信息科学*, 2020, 50: 307–317]
- 8 Duan H B, Xin L, Deng Y M. Technology research review on homing pigeon-inspired navigation (in Chinese). *CAAI Trans Intell Syst*, 2021, 16: 1–10 [段海滨, 辛龙, 邓亦敏. 仿信鸽归巢行为的导航技术研究进展. *智能系统学报*, 2021, 16: 1–10]
- 9 Zhao J X, Duan H B, Zhao Y J, et al. Consensus control of manned-unmanned aerial vehicle swarm based on hierarchy interaction of pigeons (in Chinese). *J Shanghai Jiao Tong Univ*, 2020, 54: 973–980 [赵建霞, 段海滨, 赵彦杰, 等. 基于鸽群层级交互的有人/无人机集群一致性控制. *上海交通大学学报*, 2020, 54: 973–980]
- 10 Hu D X, Dong W H, Xie W J. Proximal policy optimization for UAV autonomous guidance, tracking and obstacle avoidance (in Chinese). *J Beijing Univ Aeronaut Astron*, 2021, doi: 10.13700/j.bh.1001-5965.2021.0182 [胡多修, 董文瀚, 解武杰. 无人机自主引导跟踪与避障的近端策略优化. *北京航空航天大学学报*, 2021]
- 11 NBC news. US airstrike kills top Iran general, Qassem Soleimani. (2020-01-03). <https://www.nbcnews.com/news/world/airstrike-kills-top-irangeneral-qassim-suleimani-baghdad-airport-iraqi-n110-9821>
- 12 Ernest N, Cohen K, Kivelevitch E, et al. Genetic fuzzy trees and their application towards autonomous training and control of a squadron of unmanned combat aerial vehicles. *Unman Syst*, 2015, 03: 185–204
- 13 Sirota S. DARPA announces AlphaDogfight Trials with simulated, AI-controlled UAVs. *Inside Pentagon*, 2019, 35: 7
- 14 Shi W, Feng Y H, Cheng G Q, et al. Research on multi-aircraft cooperative air combat method based on deep reinforcement learning (in Chinese). *Acta Automat Sin*, 2021, 47: 1610–1623 [施伟, 冯旸赫, 程光权, 等. 基于深度强化学习的多机协同空战方法研究. *自动化学报*, 2021, 47: 1610–1623]
- 15 He X, Jing X N, Feng C. Air combat maneuver decision based on MCTS method (in Chinese). *J Air Force Eng Univ (Nat Sci)*, 2017, 18: 36–41 [何旭, 景小宁, 冯超. 基于蒙特卡洛树搜索方法的空战机动决策. *空军工程大学学报(自然科学版)*, 2017, 18: 36–41]
- 16 Zhou W Q, Zhu J H, Kuang M C. An unmanned air combat system based on swarm intelligence (in Chinese). *Sci Sin Inf*, 2020, 50: 363–374 [周文卿, 朱纪洪, 匡敏驰. 一种基于群体智能的无人空战系统. *中国科学: 信息科学*, 2020, 50: 363–374]
- 17 Anderson J D. *Fundamentals of Aerodynamics*. New York: McGraw-Hill, 2005
- 18 Austin F, Carbone G, Falco M, et al. Automated maneuvering decisions for air-to-air combat. *AIAA J*, 1987, doi: 10.2514/6.1987-2393
- 19 Hu B K. One to one air combat simulation system of fighter and its application (in Chinese). *Aviat Syst Eng*, 1992, 5: 35–43 [胡秉科. 歼击机一对一空战模拟系统及其应用. *航空系统工程*, 1992, 5: 35–43]

Multi-UAV cooperative swarm algorithm in air combat based on predictive game tree

ZHOU WenQing¹, ZHU JiHong², KUANG MinChi² & SHI Heng²

¹ Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;

² Department of Precision Instrument, Tsinghua University, Beijing 100084, China

An unmanned aerial vehicle (UAV) broadly refers to all kinds of remotely controlled aerial vehicles that do not require pilots to board and drive. Due to its small size, low cost, large quantity, and safety, a UAV is widely used in modern air combat. However, current studies on UAV air combat algorithms are mostly conducted in highly simplified scenarios with low precision. In addition, most of the methods used in these studies are limited by existing expert knowledge and cannot fully exploit the advantages of intelligent air combat algorithms. Therefore, this study investigates a multi-UAV cooperative swarm algorithm in air combat based on the predictive game tree. First, Unity3D is used to build a simulation environment close to a real air combat scene. In addition, a human-machine interaction environment, including UI, VR, weather system, and multifunctional screen, is realized. Then, a set of tactical maneuvers is encapsulated on the basis of existing air combat knowledge. Scripted aircraft formations are realized, and a set of air combat evaluation functions is designed. On this basis, an air combat artificial intelligence framework based on the predictive game tree is proposed. This algorithm completes the task of role assignment and maneuver decision-making, and XGBoost is used to transform it into an online real-time algorithm. Using the state machine algorithm as the baseline, the effectiveness of our air combat algorithm is verified through air combat confrontation experiments on a high-precision simulation platform.

multi-UAV cooperation based on warm intelligence, intelligent air combat, predictive game tree algorithm, high-fidelity simulation platform, autonomous control

doi: [10.1360/SST-2021-0294](https://doi.org/10.1360/SST-2021-0294)