面向兵棋推演的人机混合决策框架设计与验证

吴克宇[1], 黄魁华[1], 汪　玲[2], 许　诺[2], 李　健[2]

(1. 国防科技大学 系统工程学院,湖南 长沙 410073;2. 北京机械设备研究所,北京 100854)

摘　要:针对兵棋推演中态势高度动态、决策耦合紧密等复杂对抗特征,提出一种遵循"人在回路"原则的人机混合决策框架。框架引入基于任务时间紧迫性与决策复杂性的动态分工机制,将作战过程划分为战前规划、战中执行与战后评估3个阶段,明确指挥员与智能体的协同边界。基于灵弈平台集成武器目标分配、空战多目标打击、巡航弹弹迹规划与预警机协同战术规划4类智能体,构建可支撑复杂对抗推演的"数字参谋团"验证系统,并设置纯人工、纯智能体与人机协同3类条件开展对照实验。实验结果显示:人机协同模式取得4胜1负的最佳作战效能,显著优于其他两种模式;NASA-TLX负荷评估表明,所提框架有效降低了指挥员的认知负荷并提升了绩效表现。研究表明,所提框架在兼顾对抗效能与指挥负荷方面具有明显优势,可为智能化指挥系统的体系架构与交互机制设计提供参考。

关键词:兵棋推演; 人机混合; 人在回路; 动态分工; NASA-TLX

中图分类号:TP18;E91　　**文献标志码**:A　　**doi**:10.13705/j.issn.1671-6833.2026.04.022

随着"全域战""马赛克战"等新型作战概念的提出,现代军事对抗对指挥决策的速度、精准性与自适应性提出了前所未有的要求[1]。作为第0.999场战役,兵棋推演所面临的决策环境呈现出高维态势信息、多域力量耦合以及实时对抗博弈等特征[2],使指挥员需在强时间压力与高不确定性下完成复杂判断。在此背景下,人机混合智能成为提升推演效能的必然趋势。人工智能在大规模数据处理、多目标优化等方面具有突出优势,而人类指挥员在战略意图理解、价值权衡方面仍不可替代[3-4]。然而,当前兵棋推演系统的人机协同水平整体偏低,普遍采用静态、固化的协同模式,在高度动态的对抗环境中逐渐暴露出模式失配、认知过载和"人类离环"等问题,形成"1+1<2"的系统效能瓶颈[5-6]。

针对上述瓶颈,国内外学者从不同角度开展了人机协同决策研究[7],本文提出一种面向兵棋推演的人机混合决策框架,旨在实现"人为主导,智能赋能"的协同模式。本框架的核心创新体现在3个方面:①在理论层面,构建基于作战任务属性的角色动态分工机制,明确了从战略指挥到战术执行的人机协作方式;②在实践层面,基于灵弈智能推演系统(简称灵弈)平台集成4类功能互补的基线智能体,形成可支撑复杂对抗推演的"数字参谋团"验证系统;③在评估层面,设计兼顾作战效能指标与指挥员主观体验的综合验证体系,为框架有效性提供了客观与主观的双重依据。

1　人机混合决策框架

1.1　设计原则

本框架的设计遵循以下原则。

(1)人在回路原则。确立人类指挥员在决策中的核心地位,智能体作为辅助工具提供决策支持,确保人类智慧起主导作用[8]。

(2)渐进式智能演进原则。建立"训练-部署-反馈-再训练"的闭环优化机制,如图1所示,使智能体在过程中持续学习,形成良性人机协同进化模式。

收稿日期:2026-03-30;**修订日期**:2026-04-15

基金项目:国家自然科学基金资助项目(62273352)

作者简介:吴克宇(1990—),男,云南宣威人,国防科技大学副研究员,博士,主要从事智能博弈、优化决策、兵棋推演研究,E-mail:keyuwu@nudt.edu.cn。

通信作者:汪玲(1998—),女,湖南醴陵人,北京机械设备研究所中级工程师,主要从事博弈计算、效能评估研究,E-mail:776471577@qq.com。

图 1　智能体渐进式智能演进闭环优化机制

Figure 1　The closed-loop optimization mechanism for agent progressive intelligence evolution

(3)模块化与可扩展性原则。各智能体功能独立且接口统一,使系统根据作战需求灵活调整功能配置。

(4)态势感知透明性原则。通过可视化、可解释的交互设计,确保指挥员对智能体决策过程的充分理解,建立可靠的人机信任关系。

1.2　框架主体

本文提出的人机混合决策框架如图 2 所示,框架依托兵棋推演系统,以人为核心,智能体协同决策[7]。在作战 3 个阶段由不同角色主导。

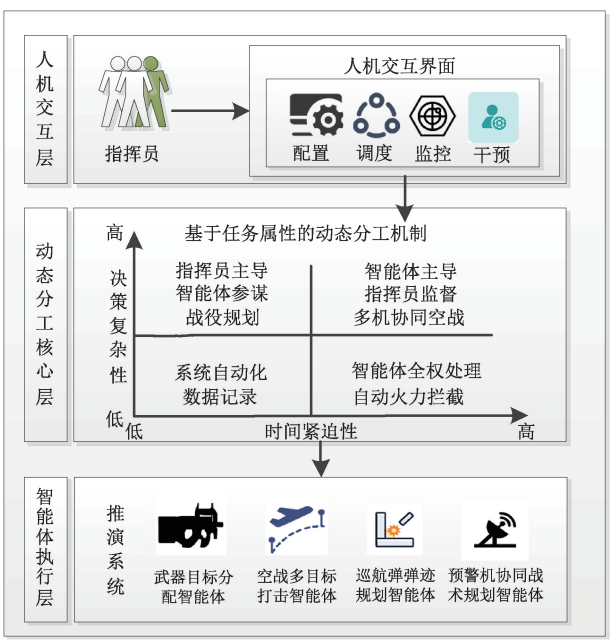

图 2　面向兵棋推演人机混合决策框架

Figure 2　Human-agent collaborative decision-making framework for wargaming

(1)战前规划阶段。指挥员在该阶段注入战略意图、作战理念与规则约束,负责制定总体作战构想与冲突解决方式。智能体则基于多源情报融合、态势评估及方案推演,为指挥员提供可行行动方案及其效能预估。通过“战略主导+智能增益”的方式,提升方案设计的质量与多样性。

(2)战中执行阶段。在战场态势高速变化的阶段,智能体承担高度实时性的感知、推理与执行决策,如威胁识别、火力协调与路径规划等,以“机器速度”压缩对手 OODA 环。指挥员则负责监控智能体行为,处理异常态势并对关键行动进行授权,以确保行动的安全性与可控性。

(3)战后评估阶段。智能体能够无偏差地回溯完整作战过程,生成时间线记录、效能评估指标与战术动作统计、克服人类回忆受限等问题。指挥员基于这些客观数据进行深度分析,提炼战法机理与经验规则,实现从对抗行为到战术知识的沉淀,形成闭环演化机制。

框架根据作战任务在时间紧迫性和决策复杂性两个维度上的属性,动态调整人与智能体的控制权。时间紧迫性衡量任务对响应速度的要求,如威胁预警、火力拦截等紧迫性高任务;决策复杂性衡量任务所需的信息处理深度、创造性推理程度,如战略规划、目标优先级调整等复杂性高任务。

基于以上二维属性空间,任务被划分为 4 种类型,对应不同的人机分工模式。通过闭环流程将战略意图逐步转化为具体行动。指挥员通过人机交互接口设定的总体作战目标,被系统性地分解为多个可计算的决策子问题,并根据上述分工原则调度对应的智能体模块进行求解。多个功能各异的作战智能体通过标准化接口对分配的实体单元进行控制,并将决策执行结果实时反馈给指挥员,形成完整的决策闭环。

各子系统通过统一的事件总线与状态共享机制进行协同。协同过程的数学模型可表述为一个多智能体马尔可夫决策过程的扩展形式,其中联合状态空间 $\boldsymbol{S}$ 包含所有智能体的局部观测与全局共享信息;联合动作空间 $\boldsymbol{A}_s$ 为各智能体动作的笛卡儿积;状态转移函数 P 和奖励函数 $\boldsymbol{R}_r$ 均由平台仿真引擎与协同规则共同定义。通过该框架,系统能够在分布式决策的基础上保持整体作战意图的一致性。

1.3　框架实例化

为验证所提出的人机混合决策框架可行性与适用性,以“要地防空”典型作战任务为例[7],对框架在实际对抗场景中的运行流程进行实例化说明。

在该任务中,红方需在有限时间、空间及防御资

源约束下,应对蓝方多类型空中威胁(如隐身战斗机、武装无人机等),实现对关键要地的多波次、高效率协同拦截。该任务场景高度动态并具有强对抗性,是验证框架有效性的典型代表。

首先,指挥员提出“确保核心要地安全”的作战意图后,系统将其分解为若干可计算的决策子问题,包括弹迹规划、武器目标分配、多目标空战协同等。随后,框架调度对应的智能体模块对各子问题分别进行求解,从而完成从战略目标到战术行为的结构化映射。

推演开始前,指挥员需完成对红方作战单元的初始设置,包括武器平台、目标信息与参考点等基础配置。随着智能体启动,各自所管控的作战单元将进入自主决策与执行流程。例如,武器目标分配智能体对地空导弹营的拦截任务进行动态优化;空战多目标打击智能体依据实时态势指挥战斗机编队执行空中拦截;巡航弹弹迹规划智能体则负责对蓝方重点目标实施精确打击并生成协同弹道。

在任务执行阶段,人机协同机制得到充分体现。指挥员通过统一的人机交互界面,可实时监控智能体运行状态与作战行为,并在需人工判断或战术突变时实施干预。例如,当智能体提出的目标分配方案需结合更高层次的意图判断时,指挥员可对智能体建议进行确认、调优或拒绝;当规划算法面临局部最优或路径受阻等问题时,指挥员可通过加入新的约束条件或调整任务优先级,引导智能体生成更符合战略目标的方案。

2 智能体与核心算法设计

本节详细阐述框架内的基线智能体系统,覆盖了从战略预警、战术空战到武器协同分配的多层次、多维度决策场景。由武器目标分配智能体、空战多目标智能体、巡航弹弹迹规划智能体和预警机协同战术规划智能体组成,分别对应武器目标分配问题、空战多目标打击分配问题、巡航弹弹迹规划问题和预警机协同战术规划问题。

针对兵棋推演中不同决策子问题的特性,依据“问题结构匹配算法优势”的原则选择算法。武器目标分配属于高维离散组合优化问题,MOPSO 擅长处理多目标、非线性的搜索空间;空战决策需要快速响应与规则可解释性,加权矩阵规则能满足实时性与透明性要求;巡航弹路径规划属于最短路径搜索,Dijkstra 算法能保证最优解且计算稳定;预警机协同涉及在不确定环境中序列决策与长期收益优化,PPO 作为策略梯度方法适合处理此类连续状态-动作空间的强化学习问题。这种差异化选择旨在使各智能体在其专长领域发挥最大效能,并通过上层框架实现协同互补。

2.1 武器目标分配智能体

本智能体的武器目标分配问题可形式化为一个多约束条件下的动态优化问题。传统推演平台的“满足条件即开火”规则易导致资源浪费与战略目标漏防[9]。本智能体旨在实现红方防御单元对来袭空中目标的全局最优分配,以最大化拦截价值并节约弹药资源。

基于上述问题将其建模为一个多目标优化问题,决策变量定义为一个二进制决策矩阵 $\boldsymbol{X}$,目标函数含两个相互冲突的目标。

$$\boldsymbol{X}=\begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \vdots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}。\tag{1}$$

最大化总拦截价值 f_1:

$$f_1(\boldsymbol{X})=\sum_{j=1}^{n} v_j\left(1-\prod_{i=1}^{m}(1-p_{ij})^{x_{ij}}\right)。\tag{2}$$

最小化武器使用次数 f_2:

$$f_2(\boldsymbol{X})=\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}。\tag{3}$$

式中:$x_{ij}=1$ 时表示武器 i 攻击目标 j;v_j 为目标 j 的价值;p_{ij} 为武器 i 对目标 j 的拦截概率。

采用多目标粒子群优化算法 MOPSO 进行求解,算法实现流程如图 3 所示。

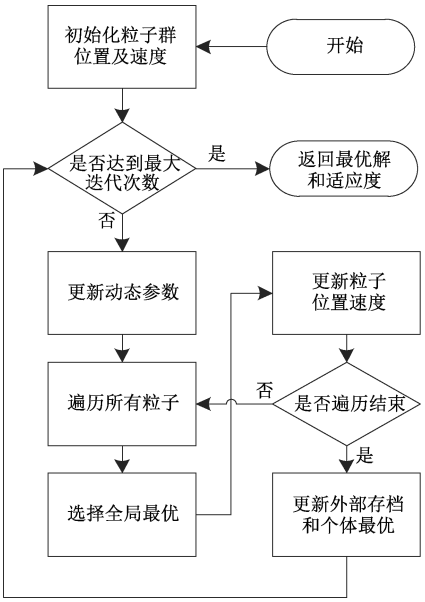

图 3 武器目标分配智能体算法流程

Figure 3 Algorithm flowchart of the weapon target assignment agent

MOPSO 算法的时间复杂度主要取决于种群规模 N、迭代次数 t 和目标函数计算开销。单次目标函数评估需计算所有武器-目标对的拦截价值与概

率,复杂度为 $O(MK)$,其中 M 为武器数量;K 为目标数量。总时间复杂度约为 $O(NTMK)$。

该智能体的交互逻辑通过“态势收集-算法求解-指令下发-效果评估”的闭环进行工作。态势信息从平台引擎获取,经 MOPSO 算法解算后,生成分配方案并通过平台接口下发打击指令。

2.2 空战多目标打击智能体

本智能体针对在空战中多目标打击问题,核心是动态多目标决策与资源分配问题。传统手操方式在应对多批次、多类型敌机来袭时,难以快速计算威胁优先级与武器命中率,易导致决策延迟和资源误配[10]。本智能体旨在通过实时计算目标威胁度与武器命中率,为红方空战单元(战斗机、地空导弹营)自动分配合适的打击目标,以提升击落率与生存率。

空战决策不仅基于静态加权矩阵,还引入了基于时序的博弈考量。在动态对抗中,智能体会依据敌我历史机动轨迹预测其下一步意图,并结合能量机动理论(energy-maneuverability theory)[11]评估红方飞机的态势优势。决策矩阵 $\boldsymbol{Q}$ 的生成不再仅是目标威胁度矩阵 $\boldsymbol{T}$ 与武器命中率矩阵 $\boldsymbol{R}$ 的静态加权,而是引入了时间窗内的态势变化率 $\Delta\boldsymbol{R}$ 与历史命中效能修正因子 $\boldsymbol{h}(t)$,公式扩展为

$$\boldsymbol{Q} = \alpha\boldsymbol{R} + \beta\boldsymbol{T} + \gamma\Delta\boldsymbol{R} + \delta\boldsymbol{h}(t)。\quad (4)$$

同时,阈值 Th 可根据对抗阶段动态调整(如初期保守、中期激进、末期保全)。此外,智能体嵌入了简单的机动策略库,当检测到敌机处于特定几何关系时,会建议或自动执行相应机动以争取优势。这些扩充使智能体能在一定复杂度上模拟空战的博弈性与时序性。

基于上述问题,其决策核心依赖于两个关键矩阵的加权计算。一是目标威胁度矩阵 $\boldsymbol{T}$,该矩阵量化了每个蓝方目标对红方每个单元的威胁程度,通过加权求和模型计算:

$$\boldsymbol{T} = \alpha_1\boldsymbol{C} + \alpha_2\boldsymbol{D} + \alpha_3\boldsymbol{A} + \alpha_4\boldsymbol{V} + \alpha_5\boldsymbol{H}。\quad (5)$$

式中:$\boldsymbol{C}$(目标价值威胁)、$\boldsymbol{D}$(距离威胁)、$\boldsymbol{A}$(攻击意图威胁)、$\boldsymbol{V}$(突防威胁)、$\boldsymbol{H}$(高度压制威胁)均为归一化后的威胁因子矩阵,权重 $\alpha_1 \sim \alpha_5$ 依据经验分别设定为 0.3,0.25,0.2,0.15,0.1。

二是武器命中率矩阵 $\boldsymbol{R}$:该矩阵量化了红方每个单元的武器对每个蓝方目标的命中概率,通过分层加权模型计算:

$$\boldsymbol{R} = \boldsymbol{R}_{\text{base}} f(\boldsymbol{d}) - g(\boldsymbol{v}) - h(\boldsymbol{m})。\quad (6)$$

式中:$\boldsymbol{R}_{\text{base}}$ 为基础命中率,反映武器固有精度;$f(\boldsymbol{d})$ 是距离衰减函数;$\boldsymbol{d}$ 为目标与武器发射距离,通常距离越远,命中率越低;$g(\boldsymbol{v})$ 为速度修正函数,与目标速度 v 紧密关联,目标速度越快武器命中率越低,速度修正越大;$h(\boldsymbol{m})$ 为目标机动修正函数,随机动参数 $\boldsymbol{m}$ 变化而变化,通过目标转向速率或敏捷度动态调整机动修正函数。最终的打击决策矩阵 $\boldsymbol{Q}$ 由威胁度矩阵与命中率矩阵加权得到,并通过与预设阈值 Th 比较,决定是否开火。

该规则算法的核心是矩阵运算。计算 $\boldsymbol{T}$ 和 $\boldsymbol{R}$ 矩阵需要对每个红方单元和每个蓝方目标进行 $O(1)$ 的标量运算,因此构建矩阵的复杂度为 $O(MK)$,后续的加权、阈值判断等操作也为 $O(MK)$。在典型规模下($M \leqslant 20, K \leqslant 20$),计算可在毫秒级完成,满足空战的高实时性要求。该智能体采用基于规则的快速决策算法,核心流程如图 4 所示。

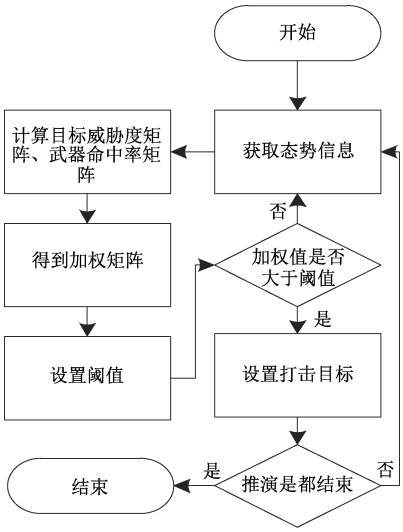

图 4 空战多目标打击智能体算法流程

Figure 4 Algorithm flowchart for multi-target air combat strike agent

从平台获取红方单元与蓝方目标的状态信息(位置、速度、高度、航向、类型等)后并行计算最终的打击决策矩阵,通过判断其与 Th 的大小关系来决定是否设置下一步的打击目标。

该智能体在推演开始后即持续运行,交互流程构成“感知-评估-决策-打击”的闭环。每步决策周期内,智能体从平台获取实时态势,通过核心算法计算打击方案,并通过平台接口为红方单元设置打击目标,同时监控燃油状态以保障任务连续性。

2.3 巡航弹弹迹规划智能体

本智能体针对巡航弹弹迹规划问题,核心是多约束条件下的实时路径规划与多弹协同问题。传统手操无法实现发射后的动态避障与弹道协调,导致导弹易被拦截且难以形成饱和打击效果[12]。本智能体旨在为已发射的巡航导弹在动态威胁环境中实时规划最优路径,并通过弹道延长时间实现多弹同

步抵达,以提升突防成功率与打击效能[13]。

该问题的求解分为两个核心阶段:路径规划和路径延长。路径规划将真实的经纬度坐标转换为二维栅格图,在此模型中,起点为红方单元位置,终点为蓝方目标位置,蓝方具备拦截能力的单元位置被建模为障碍点,栅格精度由参数 *ratio* 控制。路径延长是为实现多弹同步抵达(饱和打击)而采用的算法。该算法计算所有规划路径中的最大长度,对于长度不足的路径,在其路径中点插入一个在线段中垂线方向固定偏移距离的"之"字形绕行点,通过循环迭代使所有路径长度收敛至一致。

该智能体采用 Dijkstra 算法进行核心路径规划集成了路径延长模块。Dijkstra 算法在具有 V 个节点、E 条边的栅格图上时间复杂度为 $O((V+E)\log V)$。实验中的栅格图节点数 V 通常在 $10^2 \sim 10^3$ 量级。路径延长算法的复杂度为 $O(LN)$,其中 L 为平均路径长度,N_d 为导弹数量。因此,单次规划的总复杂度在毫秒到百毫秒级,重规划周期为 5 个仿真步长(通常大于 1 s),计算负荷在可接受范围内。实现流程如图 5 所示。

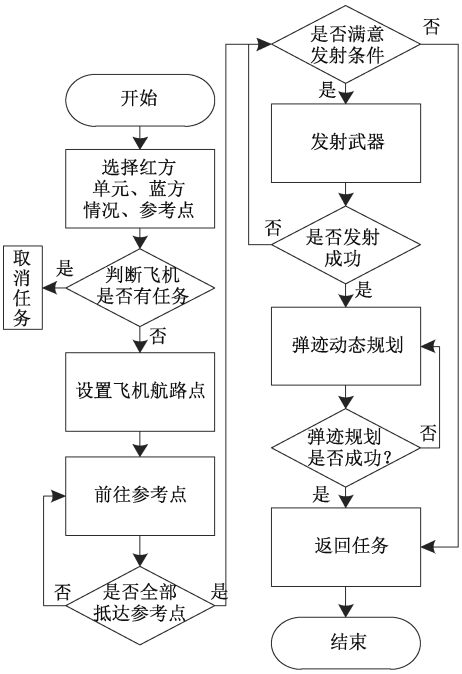

图 5 巡航弹弹迹规划智能体算法流程

Figure 5 Algorithm flowchart for cruise missile trajectory planning agent

(1)打击准备。控制载机飞往预设参考点,判断所有载机到达后,满足条件则发射巡航弹。

(2)初始规划。导弹发射后,立即调用 Dijkstra 算法,基于当前态势规划出从发射点到目标点的初始避障路径。

(3)动态重规划。在导弹飞行过程中,设定每 5 步(可配置)为重新规划步长,根据最新的战场威胁信息更新弹道。

(4)协同控制。对多枚导弹的规划路径调用路径延长算法,确保它们同时到达目标。

该智能体的交互流程是"感知-规划-控制-再规划"的闭环。智能体首先控制载机平台,完成导弹发射。导弹进入巡航阶段后,智能体持续监控战场威胁变化,周期性地调用路径规划算法为导弹生成新的航路点指令,并通过平台接口下发给导弹执行,直至其命中目标。

2.4 预警机协同战术规划智能体

本智能体针对预警机协同战术规划问题,旨在解决预警机-战斗机编队在动态威胁环境下的联合决策优化[14]。传统方法难以平衡雷达探测范围最大化与威胁规避的双重需求。本智能体通过强化学习训练预警机策略,实现情报探测效率与编队生存率的最优平衡。

基于上述问题建模,构建部分可观测的马尔可夫决策过程[15],状态空间内多维连续空间包含本机属性、传感器信息、友机信息、蓝方目标信息。动作空间是二维离散变量,包含 5 种机动动作(左切、右切、追击敌机、反向撤退、清除航路点)和 2 种雷达状态。奖励函数采用分层结构,包含基础存活奖励、区域覆盖奖励、目标探测奖励。

算法求解是采用 PPO(近端策略优化)强化学习算法,预警机作为学习智能体,通过 PPO 算法训练策略网络[16],战斗机作为规则智能体,执行以预警机为中心的护航任务[17],训练中使用专家经验引导探索,加速策略收敛[18]。

本智能体的交互逻辑是通过智能体在推演开始时启动,预警机根据训练好的策略网络实时决策机动动作和雷达开关,战斗机按规则维护护航阵型,形成协同探测与防御的闭环[19]。

3 指挥员的角色与人机交互设计

在构建人机混合框架时,确立指挥员的核心地位并设计高效的人机交互模式,是实现系统效能最大化的关键。本框架的核心理念在于构建"人在回路"的协同范式,旨在将人类指挥员的战略洞察与价值判断能力与智能体的优势深度融合,从而形成超越任何单一主体的决策优势[20]。

3.1 指挥员的核心职能与角色定位

在智能体协同作战框架中,指挥员通过人机交互接口实现对四大智能体的精准管控[21],核心职能体现在 3 个层面:一是作战编组管理,为各智能体分

派作战单元,完成力量编组;二是作战规则设定,通过参数配置贯彻战术意图;三是作战进程调控,基于战场态势动态调控智能体运行,确保了各智能体在恰当的时机发挥最大作战效能。

3.2 人机交互接口与协同机制

为有效支撑指挥员对前述智能体的精准驾驭,本框架设计了与之匹配的人机交互接口与协同机制,具体实现方式及对应的智能体反馈如表 1 所示。

表 1 人机协同机制的具体实现

Table 1 Implementation of human-agent collaboration mechanism

协同阶段	指挥员操作方式	智能体反馈与行为	实现示例
配置	通过客户端界面勾选单元、设定参数阈值	接收配置信息,更新内部状态和约束条件	指挥员为 WTA 智能体勾选地空导弹营,并设定决策间隔为 5 s
调度	点击“启动”按钮或通过任务触发器指派任务	激活核心算法,开始对分配的单元进行自主决策与控制	指挥员在界面上找到相应单元后点击“启动 WTA”,智能体开始解算目标分配方案
监控	观察界面上的态势图层、智能体状态图标和信息栏日志	持续推送决策结果、行为轨迹和关键事件通知	当 WTA 智能体拦截高价值目标后,界面左下方消息栏自动输出拦截和得分情况
干预	右键点击单元选择“接管”、调整参数滑块或点击“停止”	暂停自主控制权,等待下一轮指令或转为建议模式	指挥员基于态势考量,右键选中某单元并将其临时调离 WTA 智能体的管控列表

指挥员并非完全脱离回路,而是作为高阶决策者,通过“配置-调度-监控-干预”的流程,将智能体作为战术工具灵活运用于作战全过程。

(1)配置。指挥员在智能体调度前,设定运行的基本条件,包括单元选择、参数设定、规则阈值等,确保智能体行为符合当前战术意图。以武器目标分配智能体为例,指挥员可选择参与协同防空的地空导弹营。配置过程体现在灵弈平台客户端界面,手动勾选需纳入智能体管控的作战单元,如图 6 所示。

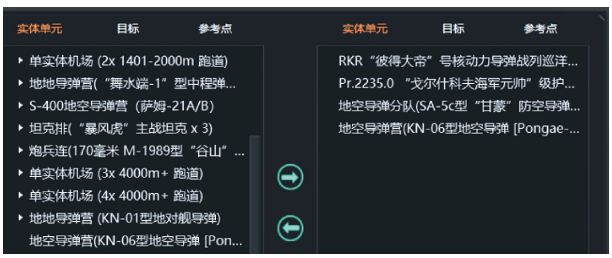

图 6 选择 WTA 单元

Figure 6 Selection of the WTA units

(2)调度。指挥员通过任务指派或任务触发方式启动智能体,赋予其特定任务权限,并在必要时指定目标或区域[22]。以武器目标分配智能体为例,设置相关单元后,在界面上找到相应单元后启动,点击如图 7 所示位置即可调度对应单元或智能体。

图 7 WTA 智能体调度设置

Figure 7 WTA Agent scheduling settings

(3)监控。智能体运行过程中,指挥员可通过客户端界面实时查看决策结果、行为轨迹与系统状态,确保按预期执行任务。以武器目标分配智能体为例,当该地面导弹营监测一架高价值蓝方侦察机进入开火范围时,基于智能体设计,会根据目标价值高低设计,优先打击高价值目标,打击成功后在界面左下方消息栏输出拦截和得分情况,如图 8 所示。

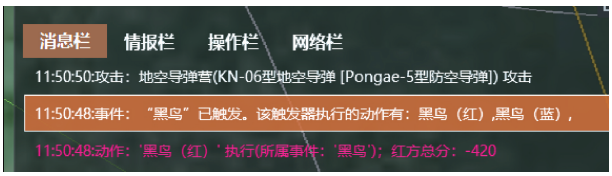

图 8 查看打击黑鸟侦察机得分情况

Figure 8 Check of the scoring results for the Blackbird reconnaissance aircraft attack

(4)干预。若智能体行为偏离预期或战场态势突变,指挥员可随时接管控制权,调整参数、重置任务或直接中止智能体运行。以武器目标分配智能体为例,当指挥员基于态势考量,需要增加某单元加入该智能体,则在界面选中该单元后右键进行调度或其他操作,如图 9 所示。

图 9 临机调整作战单元

Figure 9 Combat units of adjustment on the spot

表2列出了各智能体所支持的控制方式及其具体物理含义。

表2 智能体控制方式及物理含义

Table 2 Agent control mode and physical meaning

智能体名称	控制方式	物理含义
武器目标分配智能体	单元配置、参数配置	选择参与协同防空的单元;设定决策间隔以平衡实时性与系统负荷
空战多目标打击智能体	单元配置、参数配置	指定执行空战任务的单元;调整开火权重以控制拦截倾向性
巡航弹弹迹规划智能体	任务触发、目标指定	设置发射阵位并指定打击目标,触发智能体执行弹迹规划
预警机协同战术规划智能体	单元配置、区域预设	选择特定型号预警机;依赖预设关键区域引导航线规划行为

4 对抗场景验证及结果

4.1 实验平台与基础设置

为验证所提出的人机混合决策框架的有效性,本文基于灵弈平台构建了实验环境[23-24]。该平台具备高逼真度战场模拟、完善的裁决逻辑以及灵活的智能体集成接口,能够支持复杂对抗场景下的实时人机协同实验。在实验系统中集成了4类基线智能体。各智能体根据预设接口在推演过程中自主运行,并可接受指挥员的调控与干预,从而实现人机混合决策模式的验证。

4.1.1 硬件环境

实验计算机硬件配置统一为Intel Core Ultra 9 275HX处理器(主频3.6 GHz,24核)、NVIDIA GeForce RTX 5060 Ti显卡(16 GB显存)、32 GB DDR5 5600 MHz内存、1TB NVMe SSD固态硬盘,操作系统为Windows 11专业版(64位),实验硬件环境确保实验过程中无硬件性能瓶颈,保障推演的实时性与稳定性。

4.1.2 软件环境

基于灵弈平台构建了实验的软件环境。在实验系统中集成4类基线智能体:武器目标分配智能体、空战多目标打击智能体、巡航弹弹迹规划智能体以及预警机协同战术规划智能体。

4.2 实验设计

4.2.1 实验想定

实验采用第九届全国兵棋推演大赛智能博弈专项赛所使用的“要地冲突”场景作为研究案例[7]。在该想定中,以尤卡坦半岛真实地理环境构建红蓝双方的海空对抗局势。红方的作战目标在于维持“空运+重点防御”的补给体系,并通过主动打击削弱蓝方的持久作战能力;蓝方则旨在切断红方的陆上补给线。该场景涵盖了多域威胁、多类型作战平台及强对抗性行动,是检验人机协同框架在复杂战场环境中有效性的重要实验载体。

4.2.2 变量设置

实验设置3个组别。对照组A(纯人工)中红方由人类指挥员手动操作,不使用任何智能体辅助。对照组B(纯智能体)中红方由4个基线智能体自动决策运行,无人为干预。实验组(人机协同)中红方采用本文提出的动态人机协同框架,蓝方在3组实验中均采用完全相同的“人类离环”智能体策略,以确保自变量仅为红方的决策模式,每组进行5次完整推演对抗,记录每次对抗的胜负结果和详细数据,该数据来源于第九届全国兵棋推演大赛选手比赛数据。

4.2.3 评估指标

为全面评估不同人机协同模式的效能,本文采用多轮胜负记录和NASA-TLX负荷指数作为核心评估指标。这两个指标分别从系统作战效能和人的决策体验两个维度提供量化评估依据。

多轮胜负记录是通过在相同实验条件下进行多次重复对抗,统计各组别的获胜次数而形成的比分记录。在本实验中,每个组别进行5次完整的兵棋推演对抗,每次对抗根据推演结束时的双方得分判定胜负(得分高者胜),最终以"胜场数:负场数"的比分形式呈现。鉴于灵弈平台的推演胜负拥有较为完善的判断规则体系,故使用对抗比分来反映后续实验的作战效能情况。

NASA-TLX(task load index)是美国宇航局开发的多维负荷评估工具[25],通过6个维度评估任务负荷,分别是心智需求、时间需求、努力程度、绩效表现、挫折感、体力需求。每个维度采用0~100分评分,选手在对抗后进行打分,最终通过加权计算得出总体负荷分数。

4.3 实验结果与分析

表3展示了实验3个组别的选手在5次对抗中的胜负记录情况。

表3 组间对抗效能对比

Table 3 Comparison of inter-group effectiveness

组别	第1局	第2局	第3局	第4局	第5局	比分
对照组A	-565	-685	-2 300	2 260	-25	1:4
对照组B	-2 770	965	-965	4 925	-680	2:3
实验组	4 230	3 430	-30	3 180	195	4:1

表 4 展示了实验 3 个组别的 NASA-TLX 负荷评估结果情况。

表 4　NASA-TLX 负荷指数对比

Table 4　NASA-TLX load index comparison

组别	心智需求	时间需求	努力程度	绩效表现	挫折感	体力需求	总分
对照组 A	80±6	85±5	80±6	60±7	75±6	70±5	75.0
对照组 B	50±5	45±4	40±5	70±5	60±5	45±5	51.7
实验组	55±5	50±3	45±4	85±4	40±3	40±3	52.5

通过表 3 组间对抗效能情况可以看出实验组在胜负记录(4:1)和各项效能指标上均显著优于对照组。对照组 A(纯人工)表现最差,暴露出人类在复杂环境下的决策局限[26];对照组 B(纯智能体)虽优于纯人工,但在适应性方面存在不足[27]。

通过表 4 的 NASA-TLX 负荷指数对比情况不难看出,实验组在维持较高作战效能的同时,保持了合理的决策负荷水平。特别在绩效表现(85 分)和挫折感(40 分)维度表现优异;对照组 A 的高负荷状态(75.0 分)验证了"认知过载"问题的存在。

为更深入说明框架的有效性,对实验过程数据进行了细化分析,发现人机协同模式下,关键决策(如重要目标分配、关键时机打击)的响应时间比纯人工模式平均快 65%,决策方案的质量(以后续得分增益衡量)比纯智能体模式平均高 30%。在动态威胁处理上,人机协同模式成功应对了全部 5 次推演中蓝方发起的突发性饱和攻击,而纯智能体模式在其中 2 次应对失败,纯人工模式 3 次因决策延迟而处于劣势。这初步表明了本框架在提升决策速度、质量和动态适应性方面的综合优势。后续需通过更多实验批次和更细粒度的指标进行验证。

综合效能数据与负荷评估可见,人机混合决策框架在“作战效果-认知负荷”之间实现了优化平衡:其一,大幅提升作战效能;其二,有效降低操作疲劳与挫折感;其三,在动态威胁环境下保持较强的稳定性与泛化能力。以上结果充分验证了本文提出框架的可行性与显著优势。

5　结论

本文通过构建并验证一个面向兵棋推演的动态人机混合决策框架,系统评估了“人在回路”协同模式在复杂军事对抗场景中的效能表现。实验结果表明,该框架在作战效能与认知负荷之间实现了显著优于纯人工与纯智能体的平衡,验证了人类战略判断与智能体高速推理的互补优势。研究进一步指出,基于任务属性的动态分工机制及可解释性交互设计是提升协同质量的关键。未来需要在更加多样化的对抗环境和不同经验水平的指挥员群体中开展扩展验证,并深入研究长时推演中的信任动态与人机协同进化机制。此外,构建更加完善的智能决策技术框架也是后续研究的重要方向。

参考文献:

[1] Feng Weiqiang, Yan Zongrui. Development of US OSS construction and its inspiration[J]. Command Control & Simulation, 2018, 40(5): 137-140. [冯伟强, 严宗睿. 美军作战仿真系统建设发展及启示[J]. 指挥控制与仿真, 2018, 40(5): 137-140.]

[2] 欧阳歆. 战争工程论: 走向信息时代的战争方法学[M]. 修订版. 北京: 科学出版社, 2017.

[3] Tian Zhongliang, Liu Hao. Application of intelligent algorithm in military chess antagonism deduction[J]. Command Control & Simulation, 2021, 43(1): 40-47. [田忠良, 刘昊. 智能算法在兵棋对抗推演中的应用[J]. 指挥控制与仿真, 2021, 43(1): 40-47.]

[4] Wang Ze, Li Ni, Gong Guanghong, et al. An attention-based joint value estimation strategy for multi-agent coordination optimization[J]. Swarm and Evolutionary Computation, 2025, 99: 102132.

[5] Ding Yanyan, Feng Jianhang, Ye Ling, et al. Study on human-machine hybrid intelligent decision-making paradigm and its operational application[J]. Computer Science, 2024, 51(6): 272-281. [丁炎炎, 冯建航, 叶玲, 等. 人机混合智能决策范式及作战应用研究[J]. 计算机科学, 2024, 51(6): 272-281.]

[6] Liu Donghong. A new paradigm of human-machine hybrid intelligent mission planning based on computational mission tree[J]. Journal of Command and Control, 2023, 9(1): 85-92. [刘东红. 基于可计算任务树的人机混合智能任务规划新范式[J]. 指挥与控制学报, 2023, 9(1): 85-92.]

[7] Ma Yang, Zhang Chunxue Fang Zhongqi, et al. Research on human-machine collaborative strategy for campaign-level wargame based on "Lingyi"[C]//Proceedings of the 13th China Command and Control: CICC, 2025: 106-111. [马扬, 张春雪, 方仲琦, 等. 基于“灵奕”的战役级兵棋推演人机协同策略研究[C]//第十三届中国指挥控制大会. 北京: 中国指挥与控制学会, 2025: 106-111.]

[8] Gao Lei. Man-machine collaborative operation planning for shipborne aircraft support operations method study[D]. Zhengzhou: Zhengzhou University, 2023. [高磊. 舰载机保障作业人机协同作业规划方法研究[D]. 郑州: 郑州大学, 2023.]

[9] Su Jiongming, Cheng Lecong, Liu Hongfu. A review of

intelligent decision-making gaming methods for campaign and tactical wargaming[J]. Computer Simulation, 2026, 43(1): 21-31. [苏炯铭, 程乐聪, 刘鸿福. 战役战术兵棋推演智能决策博弈方法综述[J]. 计算机仿真, 2026, 43(1): 21-31.]

[10] Zhang Hongbing, Zhao Hong. Research on a wartime joint air operation task planning method[J]. Operations Research and Management Science, 2025, 34(9): 141-147. [张红兵, 赵红. 一种战时联合空中作战任务规划方法研究[J]. 运筹与管理, 2025, 34(9): 141-147.]

[11] Sun Yuxiang, Peng Yihui, Li Bin, et al. Overview of intelligent game: enlightenment of game AI to combat deduction[J]. Chinese Journal of Intelligent Science and Technology, 2022, 4(2): 157-173. [孙宇祥, 彭益辉, 李斌, 等. 智能博弈综述: 游戏 AI 对作战推演的启示[J]. 智能科学与技术学报, 2022, 4(2): 157-173.]

[12] Zhang Yijie, Pan Tao, Kong Zhe, et al. A collaborative midcourse guidance method based on multi-missile collaborative route planning[J]. Tactical Missile Technology, 2024(3): 114-122. [张一杰, 潘涛, 孔哲, 等. 一种基于多弹协同航迹规划的协同中制导方法[J]. 战术导弹技术, 2024(3): 114-122.]

[13] Gao Chunqing, Xiao Mingqing, Kong Qingchun, et al. Research on effectiveness optimization design of cruise missile cooperative penetration[J]. Computer Simulation, 2017, 34(1): 48-51. [高春庆, 肖明清, 孔庆春, 等. 巡航导弹协同突防的效能优化设计研究[J]. 计算机仿真, 2017, 34(1): 48-51.]

[14] Wang Xingyu, Yang Zhen, Huang Jichuan, et al. Collaborative strategy for hybrid actions of radar modes and maneuver decisions under observation errors[J]. Engineering Applications of Artificial Intelligence, 2025, 160: 111774.

[15] Zhang Yulong, Fan Changjun, Feng Yanghe, et al. Technical framework design and key issues analysis in task-level wargame intelligent decision making[J]. Journal of Command and Control, 2024, 10(1): 19-25. [张驭龙, 范长俊, 冯旸赫, 等. 任务级兵棋智能决策技术框架设计与关键问题分析[J]. 指挥与控制学报, 2024, 10(1): 19-25.]

[16] Li Chen, Huang Yanyan, Zhang Yongliang, et al. Multi-agent decision-making method based on Actor-Critic framework and its application in wargame[J]. Systems Engineering and Electronics, 2021, 43(3): 755-762. [李琛, 黄炎焱, 张永亮, 等. Actor-Critic 框架下的多智能体决策方法及其在兵棋上的应用[J]. 系统工程与电子技术, 2021, 43(3): 755-762.]

[17] Wang Junmin, Jiang Qingshan, Luo Zeming. Hierarchical decision models of formation cooperative air combat under command of AWACS[J]. Journal of Naval Aeronautical and Astronautical University, 2014, 29(5): 491-496. [王俊敏, 姜青山, 罗泽明. 预警机指挥编队协同空战分层决策模型[J]. 海军航空工程学院学报, 2014, 29(5): 491-496.]

[18] Zhang Jingke, Yang Kai, Li Chao, et al. Intelligent interference decision algorithm with prior knowledge embedded LSTM-PPO model[J]. Journal on Communications, 2024, 45(12): 227-239. [张静克, 杨凯, 李超, 等. 基于先验知识嵌入 LSTM-PPO 模型的智能干扰决策算法[J]. 通信学报, 2024, 45(12): 227-239.]

[19] Tao Wei. Application of game learning system for fighter guidance[J]. Chinese Journal of Ship Research, 2020, 15(S1): 166-172. [陶伟. 博弈学习系统在战斗机引导中的应用[J]. 中国舰船研究, 2020, 15(增刊 1): 166-172.]

[20] Yang Hongqi, Yu Qiuwei, Huang Mingjia, et al. Application of artificial intelligence in military decision making[J]. Information Technology and Network Security, 2025, 44(S1): 291-295. [杨红齐, 余秋伟, 黄铭佳, 等. 人工智能在军事决策中的应用[J]. 网络安全与数据治理, 2025, 44(增刊 1): 291-295.]

[21] Sun Yi, Zheng Yu, Huang Haiyan, et al. Multi-loop nested LLM-based multi-agent command and control processes[J]. Journal of Command and Control, 2024, 10(6): 732-739. [孙毅, 郑雨, 黄海燕, 等. 多循环嵌套的大语言模型多智能体指挥控制过程[J]. 指挥与控制学报, 2024, 10(6): 732-739.]

[22] Yan Zhenhua, Yan Zhenyu, Song Yafei, et al. Cascade threshold-regulated agent swarm firepower allocation model[J]. Aero Weaponry, 2025, 32(6): 51-60. [闫振华, 闫振宇, 宋亚飞, 等. 基于级联阈值调控的智能体群火力分配模型[J]. 航空兵器, 2025, 32(6): 51-60.]

[23] Cui Wenhua, Li Dong, Tang Yubo, et al. Framework of wargaming decision-making methods based on deep reinforcement learning[J]. National Defense Technology, 2020, 41(2): 113-121. [崔文华, 李东, 唐宇波, 等. 基于深度强化学习的兵棋推演决策方法框架[J]. 国防科技, 2020, 41(2): 113-121.]

[24] Cui Wenhua, Dong Zhihao, Liang Rongxiao, et al. Research on evaluation and analysis method for wargaming data[J]. Information Technology and Network Security, 2025, 44(S1): 302-306. [崔文华, 董志浩, 梁荣晓, 等. 兵棋推演数据评估分析方法研究[J]. 网络安全与数据治理, 2025, 44(增刊 1): 302-306.]

[25] Hart S G, Staveland, L E. Development of NASA-TLX

(task load index): results of empirical and theoretical research[J]. Advances in Psychology, 1988, 52(6): 139-183.

[26] He Xiaoyuan, Guo Shengming, Wu Lin, et al. Modeling research of cognition behavior for intelligent wargaming [J]. Journal of System Simulation, 2021, 33(9): 2037-2047. [贺筱媛, 郭圣明, 吴琳, 等. 面向智能化兵棋的认知行为建模方法研究[J]. 系统仿真学报, 2021, 33(9): 2037-2047.]

[27] Zhang Tao, Xiang Qi, Zheng Wanwen, et al. Application of path planning based on improved A* algorithm in war gaming of naval warfare[J]. Acta Armamentarii, 2022, 43(4): 960-968. [张韬, 项祺, 郑婉文, 等. 基于改进 A*算法的路径规划在海战兵棋推演中的应用[J]. 兵工学报, 2022, 43(4): 960-968.]

Design and Verification of Human-Agent Collaborative Decision-Making Framework for Wargaming

WU Keyu[1], HUANG Kuihua[1], WANG Ling[2*], XU Nuo[2], LI Jian[2]

(1. Systems Engineering College, National University of Defense Technology, Changsha 410073, China; 2. Beijing Institute of Mechanical Equipment, Beijing 100854, China)

Abstract: To address the highly dynamic and tightly coupled decision-making characteristics of wargaming confrontations, a human-agent collaborative decision-making framework grounded in the human-in-the-loop principle was proposed. The framework introduced a dynamic task-allocation mechanism based on task urgency and decision complexity, dividing the operational process into three stages, including pre-war planning, in-war execution, and post-war evaluation, to clarify the collaborative boundaries between commanders and agents. A verification system integrating four types of agents, namely weapon-target assignment, multi-target air combat strike, cruise missile trajectory planning, and airborne early warning collaborative tactical planning, was implemented on the LingYi platform to form a "digital staff group" capable of supporting complex adversarial wargaming. Comparative experiments were conducted under three conditions: human-only, fully autonomous agents, and human-agent collaboration. Results showed that the collaborative mode achieved the best operational performance with four wins and one loss, significantly outperforming the other two modes. NASA-TLX load evaluations further confirmed that the framework effectively reduced commanders' cognitive workload and enhanced performance. These findings demonstrated that the proposed framework achieved a favorable balance between operational effectiveness and command load, offering valuable insights for the design of the system architecture and interaction mechanism of the intelligent command system.

Keywords: wargaming; human-agent collaborative; human-in-the-loop; dynamic task allocation; NASA-TLX