

文章编号:1000-6893(2009)09-1727-13

基于群集智能的协同多目标攻击空战决策

刘波, 覃征, 邵利平, 高由兵, 王瑞

(西安交通大学 电子与信息工程学院, 陕西 西安 710049)

Air Combat Decision Making for Coordinated Multiple Target Attack Using Collective Intelligence

Liu Bo, Qin Zheng, Shao Liping, Gao Youbing, Wang Rui

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

摘要: 为使各武器平台利用数据链(TADIL)交换和共享信息,自主协同计算实现多目标攻击决策,本文基于 NASA/Ames 研究中心提出的群集智能(COIN)理论,构建以作战单元为智能体(Agent)的协同多目标攻击决策模型,并通过定义 Agent 的贡献度扩展群集智能架构,用于解决分布式异构 Agent 造成的系统难以准确收敛问题。为提高收敛速度,减少求解过程中的盲目搜索和提高已形成解的收敛,研究加入了构造性启发和改进性启发两类启发方法。实验表明所提算法同传统算法相比,具有更好的稳定性和扩展性,减少了计算量,收敛速度也有一定的提高。

关键词: 群集智能; 协同攻击; 多目标分配; 空战; 决策; 超视距

中图分类号: V247; E837 **文献标识码:** A

Abstract: In modern air combat, it is an important yet difficult task to coordinate multi-fighter to make multi-target attack decisions by using the tactical digital information links (TADILs) to share and exchange combat information. To address the problem, a novel distributed heterogeneous multi-target air combat decision-making model is proposed based on the collective intelligence (COIN) theory put forward by the NASA/Ames Research Center. In this model, each missile is defined as an agent and the conventional COIN framework is expanded by agent contribution rate to realize accurate convergence under a distributed heterogeneous agent environment. In order to improve the convergence rate, two classes of heuristic information is added to the algorithm, i. e., constructive heuristics and improvement heuristics, to reduce blind search in intermediate results and enhance convergence rate in ultimate results. Simulation results show that as compared with conventional methods, the proposed method is able to converge to the global optimum with more stability and higher expandability, decrease the amount of computation and improve convergence rate.

Key words: collective intelligence; coordinated attack; multi-target assignment; air combat; decision making; beyond visual range

超视距(BVR)空战环境下,参战各武器平台(飞机等)利用数据链,共享战场态势,进行协同多目标攻击是现代空战的重要形式。研究在超视距空战环境下融合多信息源,从而将信息优势转化为决策优势,进而转化为行动优势,实现各武器平台间的自主协同在现代战争中尤为重要。因空战中的协同多目标攻击决策是一个带约束条件的组合优化问题,现有的多目标空战决策方法多采用神经网络、遗传算法、蚁群、粒子群等仿生算法,找出其最优分配方案,进行攻击决策。例如文献[1]首先计算协同优先权,然后通过神经网络和自组

织特征映照(SOFM)网络训练形成分配方案;文献[2]采用标准的遗传算法得出决策方案;文献[3]通过确定飞机*i*攻击当前最大优势的敌机为“好”基因,然后在遗传中保留“好”基因改进遗传算法得出决策方案,并与标准的遗传算法作了对比;文献[4]将一种局部搜索方法引入自适应遗传算法,形成混合自适应遗传算法得出分配方案;文献[5]将协同多目标攻击限制在目标仅分配一枚导弹情况下,提出应用模拟退火遗传算法解决问题寻优;为了提高算法的收敛速度,文献[6]提出在粒子群算法中加入启发式信息改进局部搜索的启发式粒子群算法得出分配方案,文献[7]提出在蚁群算法中加入启发式信息进行局部改进的启发式蚁群算法得出分配方案。上述文献均使用仿生

算法解决协同多目标分配问题。此外,在无人机的协同多目标攻击决策上,文献[8]~文献[10]采用了拍卖算法对目标进行分配。上述这些方法关注算法本身对建模后组合优化问题的求解能力,而没有注重问题本身的分布式等特点,其主要形式是采用集中式的计算和通信模式,由中心节点获取全部战场态势信息进行全局计算,然后将结果传递给各作战单元(空空导弹等)执行命令。这种方式存在以下问题:

(1) 仿生算法本身或算法加入启发式信息进行局部改进,易导致“公共地悲剧^[11]”问题,从而陷入局部最优解。“公共地悲剧”问题是指,在资源受限情况下,个体追求自己利益最大化时伤害总体利益。在协同多目标攻击中,指出现了个体或部分集在确定攻击对象的过程中追求最大化自己目标的同时,总效用反而减少。例如,在文献[7]中,各导弹以对目标造成的威胁为主要分配依据,此时,易出现多个导弹攻击同一个威胁较大目标的情况,尽管导弹达到各自利益最大化,却损害了全局利益,从而导致“公共地悲剧”问题的出现。

(2) 没有考虑问题本身的分布式特点和应用背景,忽略了各武器平台本身的自主计算和协同能力。集中式计算带来的指挥层级过多,响应时间过长,组网形式单一等在现代信息化战争中已突显弊端^[12],不利于在各个武器平台间形成自主协同以加快“从传感器到射手”的决策和行动速度,且集中的计算和通信让各武器平台被动地等待协同攻击指令,易导致系统脆弱,在应用中存在很大局限^[12-13]。

针对以上问题,借鉴 NASA/Ames 研究中心在解决协同分布式计算问题时提出的群集智能理论^[14-16],实现了一种基于群集智能架构的协同多目标攻击决策算法。该算法利用数据链共享信息,基于各武器平台自身的计算能力形成自主协同,解决协同多目标攻击决策问题。在所提算法中,构建了以各作战单元(空空导弹等)为基本变量(智能体(Agent))的协同多目标攻击决策模型,通过定义 Agent 的贡献度,扩展了群集智能架构(原群集智能理论仅适用于同构的多 Agent 系统),解决了分布式异构 Agent 造成的系统难以准确收敛问题。为提高收敛速度,研究加入了构造性启发和改进性启发两类启发方法,分别用来减少求解过程中的盲目搜索和提高已生成解的收敛。

1 群集智能简介

群集智能是 NASA/Ames 研究中心群集智能项目提出用来解决大规模自主 Agent 的智能控制和协同分布式计算问题的一种架构及相关算法^[15]。群集智能定义为具有以下两个特征的大规模多 Agent 系统^[14, 17]:

① 系统含有一个或多个 Agent,这些 Agent 试图最大化其个体效用函数。

② 有一个明确的全局效用函数,可评价整个系统可能的行为。

群集智能的核心思想是构建系统内每一个 Agent 的个体效用函数,各 Agent 通过学习使该函数最大化,且当个体效用函数达到最大化时,系统全局效用函数也达到最大值。同传统解决协同分布式计算问题的方法相比^[14],群集智能不需对问题进行详细建模,即使是架构中使用的模型,也很少涉及动态性。

群集智能理论涉及 4 个基本概念,即:Agent,世界线,个体效用和全局效用。

群集智能中的 Agent,是指任何在逻辑上可区分的计算体。用 η 表示 Agent 实例,用 $\hat{\eta}$ 表示系统中除 η 外的其他所有 Agent 实例。

所谓世界线,是指保存整个系统的全部信息,包括环境、Agent 的外部行为、内部学习系统及参数的一个欧几里德向量,用 z 表示。

所谓个体效用,是指定义在世界线 z 上,可以评价 Agent 所有可能行为的函数,记 Agent η 的个体效用为 $g_{\eta}(z)$ 。在系统中, η 中学习算法的目标是最大化 $g_{\eta}(z)$ 。

所谓全局效用,是指对系统所有可能的历史,即对所有世界线进行评价的评价函数。记系统的全局效用为 $G(z)$ 。群集智能系统的最终目标是最大化 $G(z)$ 。

群集智能系统内的不确定性行为,用内置在世界线 z 中一个不可见部分的随机发生器产生,即不存在外部不可知的力量影响系统状态。

群集智能通过 1 个中心方程和 4 个显著特征(因子性(Factoredness)、易学性(Learnability)、智能性(Intelligence)和完美效用(Wonderful Life Utility))来实现其架构和算法。因子性、易学性、智能性定义在 Agent 的个体效用函数上,完美效用是同时具有因子性和易学性的一个实例。

效用函数 $g_{\eta}(z)$ 可评价 Agent η 的所有行动,对于 $g_{\eta}(z)$,若 η 任意导致 $g_{\eta}(z)$ 值增加的行动,

均导致全局效用 $G(\mathbf{z})$ 值增加,则称 $g_\eta(\mathbf{z})$ 具有因子性。定义如下:

定义 1(因子性) 对于任意两条世界线 \mathbf{z} 和 \mathbf{z}' , 有 $z_\eta = z'_\eta, z_\gamma \neq z'_\gamma$ 。即两条世界线仅在 η 处不同, 如 η 的个体效用函数满足

$$g_\eta(\mathbf{z}) \geq g_\eta(\mathbf{z}') \Leftrightarrow G(\mathbf{z}) \geq G(\mathbf{z}') \quad (1)$$

则称效用函数 $g_\eta(\mathbf{z})$ 满足因子性。

Agent η 的个体效用具有因子性意味着 η 任意导致自己获利的行动, 同样也会导致全局获利, 即 Agent 的行为不会以伤害集体来获利。

Agent η 通过学习提高个体效用 $g_\eta(\mathbf{z})$, 若 $g_\eta(\mathbf{z})$ 满足因子性, 则 η 可通过提高 $g_\eta(\mathbf{z})$ 来提高全局效用 $G(\mathbf{z})$ 。定义易学性来衡量 η 学习提高 $g_\eta(\mathbf{z})$ 的难易程度, 具有高易学性的 $g_\eta(\mathbf{z})$, Agent 越容易学习。定义如下:

定义 2(易学性) Agent η 关于效用函数 $g_\eta(\mathbf{z})$ 的易学性定义为

$$\lambda_{\eta, \varepsilon_\eta}(\mathbf{z}) \equiv \frac{\|\partial_{z_\eta} g_\eta(\mathbf{z})\|}{\|\partial_{z_\gamma} g_\eta(\mathbf{z})\|} \quad (2)$$

式中: $\partial_{z_\eta} g_\eta(\mathbf{z})$ 表示 η 的效用函数 $g_\eta(\mathbf{z})$ 对 z_η 求偏导; $\partial_{z_\gamma} g_\eta(\mathbf{z})$ 表示 η 的效用函数 $g_\eta(\mathbf{z})$ 对 z_γ 求偏导^[14]; $\|\cdot\|$ 表示求范数。

易学性定义为 $g_\eta(\mathbf{z})$ 中由 η 行动引起的变化与由其他 Agent 的行动引起变化的比, 反映了 Agent 自身的行动对其的影响, 也可说是 Agent η 的行动与其他 Agent 行动的独立度。如同信号处理中的“信噪比”, 分子表示信号强度, 分母表示噪声强度。效用函数具有高的易学性能够使 Agent 迅速感知到不同行动的结果。

Agent η 学习提高 $g_\eta(\mathbf{z})$, 需要辨别可行行为集中各行为之间的差别, 以选择最好的行动, 加速学习。定义 $g_\eta(\mathbf{z})$ 的智能性以评价每个行为的优劣。

定义 3(智能性) Agent η 关于效用函数 $g_\eta(\mathbf{z})$ 的智能性定义为

$$\varepsilon_{\eta, \varepsilon_\eta}(\mathbf{z}) \equiv \int d\mu(\mathbf{z}'_0) \{H[g_\eta(\mathbf{z}) - g_\eta(\mathbf{z}'_{\text{time} < 0} \cdot C(\mathbf{z}'_0))] \delta(\mathbf{z}'_{\eta, 0} - \mathbf{z}_{\eta, 0})\} \quad (3)$$

式中: $H(x)$ 为 Heaviside theta 函数; $\delta(x)$ 为 Dirac delta 函数; time 为关于 \mathbf{z} 的时间量; $C(\mathbf{z}'_0)$ 为系统在世界线 \mathbf{z}' 上, 时间 $\text{time} \geq 0$ 的动态性; $\mathbf{z}'_{\eta, 0}$ 为世界线 \mathbf{z}' 在分量 η 上, 时间 $\text{time} = 0$ 时的值, $\mathbf{z}_{\eta, 0}$ 定义与此相似; 设 $\int d\mu(\mathbf{z}'_{\text{time}}) = 1$ ^[14]。

智能性是关于每个行动评级的一个特征, 对

于 Agent η 的一个行动 a , 如果其值等于 0.7, 则代表在 Agent 所有可采用的行动中, 有 70% 的行动得到的效用值低于采用当前行动 a 得到的效用值。效用函数具有高的智能性可使 Agent 得到高的个体效用, 继而得到高的全局效用。

在协同系统中, Agent 间存在相互影响, 有些影响内在的 Agent 间关系比较清晰, 如 Agent 集合是领导者-跟随者模式^[15], 跟随者的行动由领导者确定。定义影响集以利于系统中 Agent 间相互影响的确定和 Agent 对个体效用的学习。

定义 4(影响集) Agent η 的影响集定义为

$$S_\eta^{\text{eff}}(\mathbf{z}) \equiv \{i \mid \partial_{z_i}(\mathbf{z}_i) \neq \mathbf{0}\} \quad (4)$$

式中: $\partial_{z_i}(\mathbf{z}_i)$ 表示对 \mathbf{z} 中对象 i 所关联的分量关于 z_i 求偏导。

影响集说明系统内 Agent 间行动的相互影响范围, Agent η 的影响集表示受 η 行动影响的 Agent 的集合。

在群集智能架构中, 对 Agent 个体效用的学习, 需要“虚拟”出 Agent 实际没有做的行动以进行计算和对比。定义这些“虚拟”行动为系统内的一种操作。

定义 5 $CL_\sigma^*(\mathbf{z})$ 定义为一种操作, 表示将 \mathbf{z} 中 σ 转化为指定的向量值 κ (固定值或函数值), 例如, $CL_\eta^0(\mathbf{z})$ 指将 \mathbf{z} 中的 Agent η 对应的向量转化为向量 $\mathbf{0}$ 。

$CL_\sigma^*(\mathbf{z})$ 操作可以使系统仿真各种实际上没有在系统内发生的情况, 计算出各种“虚拟”系统的值, 达到学习的目标。

定义 6(完美效用) 完美效用是同时具有因子性和易学性的一个实例, 定义为差分方程

$$WLU_\eta(\mathbf{z}) \equiv G(\mathbf{z}) - G(CL_{S_\eta^{\text{eff}}}^*(\mathbf{z})) \quad (5)$$

经 $CL_{S_\eta^{\text{eff}}}^*(\mathbf{z})$ 操作后, η 及其影响集被转化为固定的向量值, 使 $G(CL_{S_\eta^{\text{eff}}}^*(\mathbf{z}))$ 独立于 Agent η 。

定理 1 如果差分效用可以描述为

$$D_{g_\eta}(\mathbf{z}) \equiv G(\mathbf{z}) - \Gamma(f(\mathbf{z})) \quad (6)$$

式中: $\Gamma(\cdot)$ 为独立于 η 的一个函数, $f(\cdot)$ 为关于 \mathbf{z} 的函数, 则 $D_{g_\eta}(\mathbf{z})$ 关于 $G(\mathbf{z})$ 满足因子性(证明略)。

由定理 1, 可以得出 $WLU_\eta(\mathbf{z})$ 关于 $G(\mathbf{z})$ 满足因子性。

设 $CL_{S_\eta^{\text{eff}}}^*(\mathbf{z})$ 中向量 κ 为 $\mathbf{0}$, 即将 η 的影响集固定为 $\mathbf{0}$, 此时, $WLU_\eta(\mathbf{z})$ 仅为 η 的影响集对系统的作用, 由定义 2 可知, 此函数具有高的易学性。

群集智能中心方程的理论基础为贝叶斯理论。

$$P(G | s) = \int d\mathbf{e}_G \{P(G | \mathbf{e}_G, s) \cdot \int d\mathbf{e}_g P(\mathbf{e}_g | \mathbf{e}_G, s) P(\mathbf{e}_g | s)\} \quad (7)$$

式中: $\mathbf{e}_g \equiv [\epsilon_{\eta_1, g_{\eta_1}}(\mathbf{z}) \quad \epsilon_{\eta_2, g_{\eta_2}}(\mathbf{z}) \quad \dots]$ 为 Agent 关于各自个体效用函数智能性的向量; $\mathbf{e}_G \equiv [\epsilon_{\eta_1, G}(\mathbf{z}) \quad \epsilon_{\eta_2, G}(\mathbf{z}) \quad \dots]$ 为 Agent 关于全局效用函数智能性的向量; s 为 \mathbf{z} 中所有系统设计者可以自由改变的量。

定理 2 若对系统内所有 Agent η , 有关于个体效用函数智能性

$$\epsilon_{\eta, g_{\eta}}(\mathbf{z}) = 1 \quad (8)$$

则 \mathbf{z} 为所有 Agent 集的纳什均衡点。

证明 式(3)对所有 Agent 满足式(8), 则根据智能性的定义, 在 \mathbf{z} 点 Agent η 行动为最优行为, 任何其他行动都劣于当前行动, 即 Agent 选择行动完全理性。给定其他 Agent 的行为, η 的行动可以最大化个体效用 g_{η} 。依照纳什均衡的定义^[18], \mathbf{z} 为纳什均衡点。证毕。

推论 1 若对系统内所有 Agent η , 有关于全局效用函数智能性

$$\epsilon_{\eta, G}(\mathbf{z}) = 1 \quad (9)$$

则 \mathbf{z} 为 $G(\mathbf{z})$ 的一个局部最优点(证明略)。

由定理 2 和推论 1, 如果 $\epsilon_{\eta, g_{\eta}}(\mathbf{z})$ 和 $\epsilon_{\eta, G}(\mathbf{z})$ 等同, 则若 Agent 能最大化其个体效用函数, 就可使全局效用达到一个局部最优点。

综合上述定义和定理, 建立好全局效用函数和计算出个体效用函数后, 通过设计式(7)中 $P(\mathbf{e}_g | s)$ 的 s , 使得 Agent 关于个体效用的智能性接近 1(要求个体效用函数有较高的“信噪比”), 最大化个体效用, 进而使 $P(\mathbf{e}_G | \mathbf{e}_g, s)$ 中 Agent 关于总体效用的智能性接近 1(个体效用函数“因子化”于全局效用 G), 使总体效用达到局部最优, 最后通过 $P(G | \mathbf{e}_G, s)$ 使得 G 最大(避免陷入局部最优), 得到全局最优^[14]。

Agent 内部对个体效用函数进行学习的算法一般使用强化学习算法^[14], 本文中使用的 Q 学习算法^[15, 19]。

$$Q(s, a) = r + \max_{a'} Q(s', a') \quad (10)$$

式中: r 为当前行动的回报; s' 为在状态 s 下采取行动 a 达到的下一个状态; a' 为在状态 s' 下 Agent 可采取的行动; $\max_{a'} Q(s', a')$ 为状态 s' 下 Q 函数能达到的最大值。Agent 学习算法的探索策略采用 Boltzmann 函数, Agent 选取行动 a 的概率为

$$P(a) = \frac{e^{Q(s, a)/t}}{\sum_{a' \in A} e^{Q(s, a')/t}} \quad (11)$$

式中: A 为行为集; t 为温度参数, 用来调节随机性, 随着学习的进行, 温度值衰减, 衰减因子为 d , $d \in (0, 1]$, 从而保证系统收敛。

现有群集智能的研究和应用主要有网络路由, 卫星集群的控制通信, 异构计算网络的作业调度, NASA 火星探测器项目, 直升机飞行编队及美国军方的一些项目等^[15-17]。在这些项目中, 群集智能可以迅速构建一个健壮、易扩展、结构灵活的分布式解决方案, 使众多 Agent 有效地协同工作。

在上述研究和应用中, Agent 往往是同构的, 甚至协同的 Agent 集初始位置集中或初始位置无关^[14]。此时, 构建以 Agent 为变元的全局效用函数模型较为简单, 因 Agent 同构且集中, 全局函数可描述为关于 Agent 数量的函数的线性组合。其步骤为

① 设计系统的全局效用函数 $G(\mathbf{z})$, 此函数可以评价系统所有的可能情况。

② 确定 Agent 的影响集 S_{η}^{eff} 。

③ 由式(5)计算出每个 Agent 的完美效用函数 $WLU_{\eta}(\mathbf{z})$ 。

④ Agent 内部采用强化学习算法对个体效用进行学习, 得出各自的最优策略实现自主协同。

当系统对时间、计算能力或通信能力等敏感时, 系统需在受限条件下计算协同方案。此时, 可通过调整 Agent 内部的学习算法平衡行为的“探索与利用^[19]”以满足约束条件, 但可能会因对行为信息学习不充分而导致 Agent 行动不能达到完全“理性”, 从而系统不能达到全局最优解。当然, 当系统具有某些固定特征时, 可将其作为启发信息加快系统学习。

2 协同多目标空战决策问题与建模

设有红(决策方)、蓝(目标方)双方战机进行军事对抗。红方通过陆基雷达或空中预警机等发现某一空域中有 N 架蓝机来袭, 立刻进行兵力资源部署, 出动 M 架具有多目标攻击能力的战机协同对来袭机群进行超视距攻击拦截。为简单起见, 假设每架红机携带 L 枚中远程空空导弹, 可同时攻击空战的 L 个目标。记所有红机为机群 $R, R = \{R_i, i = 0, 1, \dots, M-1\}$, 其中 R_i 为第 i 架红机; 记所有蓝机为机群 $B, B = \{B_j, j = 0, 1, \dots, N-1\}$, 其中 B_j 为第 j 架蓝机。

多机协同多目标攻击决策的目的就是, 在超视距条件下将 N 架来袭的蓝机有效地分配给 M 架红机进行攻击拦截, 更进一步将 N 架蓝机分配

给 M 架红机所挂载的 $M \times L$ 枚导弹进行攻击, 以实现最佳的攻击拦截效能。

2.1 协同多目标攻击决策问题

威胁估计是超视距环境下多机协同多目标攻击决策的前提和基础^[20]。本文结合超视距空战的特点, 根据空战双方作战飞机性能参数及相对战术几何关系, 采用非参量法^[20] 计算威胁估计。为研究方便, 假设空战在平面内进行, 红、蓝双方战机之间的空战几何态势关系如图 1 所示。

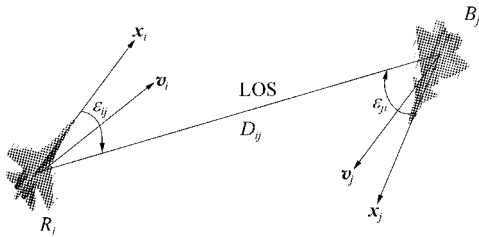


图 1 红、蓝双方战机之间的空战几何态势
Fig. 1 Air combat situation between R_i and B_j

图 1 中, LOS 为视线(目标线), 即红机 R_i 和蓝机 B_j 之间的连线; D_{ij} 为 R_i 相对于 B_j 的距离; x_i 与 v_i 分别为 R_i 的机体轴线和速度矢量; ϵ_{ij} 为 B_j 相对于 R_i 的离轴发射角; x_j, v_j 和 ϵ_{ji} 用类似方法定义^[7]。

定义 7 红机 B_i 对蓝机 B_j 的威胁定义为

$$th_{ij} = \omega_1 P_{ij}^D P_{ij}^\epsilon + \omega_2 P_{ij}^v + \omega_3 P_{ij}^E \quad (12)$$

式中: ω_1, ω_2 和 ω_3 为非负的权重系数, 其值由专家给出, 且满足 $\omega_1 + \omega_2 + \omega_3 = 1$; P_{ij}^D 为距离威胁, $P_{ij}^D \in [0, 1]$; P_{ij}^ϵ 为角度威胁, $P_{ij}^\epsilon \in [0.1, 1.0]$; P_{ij}^v 为速度威胁, $P_{ij}^v \in [0.1, 1.0]$; P_{ij}^E 为空战能力威胁, $P_{ij}^E \in [0, 1]$ ^[5, 7, 21]。由上述 4 个威胁的定义域可知 $th_{ij} \in [0, 1]$ 。假设在空战中某一时刻, 红机 R_i 击毁蓝机 B_j 的概率与此时 R_i 对 B_j 的威胁成正比, 则可用 R_i 对 B_j 的威胁 th_{ij} 表示 R_i 击毁 B_j 的概率。以同样的方式定义 B_j 对 R_i 的威胁 th_{ji} 。

基于红、蓝双方的相互威胁, 多目标攻击决策问题可转化为计算红方消除的总威胁, 或计算蓝方剩余的总威胁问题。

2.2 协同多目标攻击决策建模

为有效准确地实现信息化战争“从传感器到射手”, 形成自主协同架构, 协同的对象要由武器平台向具体的作战单元进行转变。用实际完成攻击的空空导弹建模比用战机进行建模描述空战效能更为直接和准确。协同空战决策问题可以描述

为: 寻找一个合适的导弹目标分配方案 π^* , 使协同攻击效能最佳, 即

$$\pi^* = G^{-1}(\text{optimal } G(\pi))_{\pi \in \Omega} \quad (13)$$

式中: π 为红方对蓝方的导弹目标分配方案; Ω 为所有可行导弹目标分配方案构成的解空间; $G(\pi)$ 为衡量协同攻击效能的效用函数。

根据拟定的多机协同多目标空战场景假设, 红方机群 R 的总导弹数为

$$Z = M \times L \quad (14)$$

目标数量(蓝机数量)

$$T = N \quad (15)$$

由于机群 R 的导弹数量及空战能力有限, 此处假设参与拦截攻击目标的导弹数量满足

$$T \leq Z \leq 2T \quad (16)$$

设机群 R 的 Z 枚导弹构成导弹集合 K , 且 $K = \{k, k=0, 1, \dots, Z-1\}$ 。这样, 可以定义 K 的第 k 枚导弹 M_k 与红机 R_i 的第 h 枚导弹的对应关系为

$$k = iL + h$$

$$\forall i \in \{0, 1, \dots, M-1\}, \quad \forall h \in \{0, 1, \dots, L-1\}$$

$$(17)$$

在协同空战决策中, 需要考虑每个目标可能分配的导弹数量及当前参与协同打击的导弹攻击数量。在本文中作如下约定: ①一个目标可分配 0~2 枚导弹; ②参与协同攻击的 Z 枚导弹均分配攻击目标。形式化为

$$\sum_{k=0}^{Z-1} X_{kj} \leq 2, \quad \forall j \in \{0, 1, \dots, N-1\} \quad (18)$$

$$\sum_{j=0}^{N-1} X_{kj} = 1, \quad \forall k \in \{0, 1, \dots, Z-1\} \quad (19)$$

式中: X_{kj} 为布尔值, $X_{kj} = 1$ 表示红方导弹 M_k 攻击蓝机 B_j , $X_{kj} = 0$ 表示红方导弹 M_k 不攻击蓝机 B_j 。

现令导弹目标分配方案 $\pi = [t_0 \quad t_1 \quad \dots \quad t_k \quad \dots \quad t_{Z-1}]$, 其中 t_k 为导弹 k 攻击的目标, $t_k \in B, k \in \{0, 1, \dots, Z-1\}$ 。求得最优导弹目标分配方案 π^* 后, 由式(17)确定导弹的所属战机, 最终确定协同空战决策方案。因不考虑同一载机上的各枚导弹之间的差别, 所以获得的 π^* 并不唯一, 但依据它们所确定的协同攻击决策达到的效能是唯一的。

衡量攻击效果有多种方式, 相应的, 对于问题的建模也有所区别。由于威胁事实上是由挂载在战机上可进行攻击的武器单元完成, 这些武器单元可共享战机作战效能和信息, 为简单起见, 假设红机 R_i 上挂载的导弹对蓝机 B_j 的威胁与红机的威胁成正比, 进一步, 假设等同于 th_{ij} 。即 R_i 挂

载的导弹 M_k 击毁 B_j 的概率等于 R_i 对 B_j 的威胁, 即 $th_{kj} = th_{ij}$, 则 $(1 - th_{kj})$ 为导弹 M_k 攻击 B_j 后 B_j 的存活率, $\prod_{k=0}^{Z-1} (1 - th_{kj})^{X_{kj}}$ 为在一次协同攻击后 B_j 的存活率, $1 - \prod_{k=0}^{Z-1} (1 - th_{kj})^{X_{kj}}$ 为一次协同攻击后 B_j 的被摧毁率。

红方计算时可从红、蓝机群两个角度考虑, 总效用函数分别设计为计算蓝方剩余总威胁 $G(\pi)$ 的式(20), 和红方消除总威胁 $G'(\pi)$ 的式(21), 从决策来讲, $G(\pi)$ 越小越好, $G'(\pi)$ 越大越好。

$$G(\pi) = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left\{ th_{ji} \left[\prod_{k=0}^{Z-1} (1 - th_{kj})^{X_{kj}} \right] \right\} \quad (\pi \in \Omega) \quad (20)$$

$$G'(\pi) = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left\{ th_{ji} \left[1 - \prod_{k=0}^{Z-1} (1 - th_{kj})^{X_{kj}} \right] \right\} \quad (\pi \in \Omega) \quad (21)$$

约束条件均为

$$\left. \begin{aligned} \sum_{k=0}^{Z-1} X_{kj} &\leq 2, \quad \forall j \in \{0, 1, \dots, N-1\} \\ \sum_{j=0}^{N-1} X_{kj} &= 1, \quad \forall k \in \{0, 1, \dots, Z-1\} \end{aligned} \right\} \quad (22)$$

易推导出

$$G'(\pi) = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} th_{ji} - G(\pi) \quad (\pi \in \Omega) \quad (23)$$

对特定的一个时间点, 双方态势信息是静态的, 因此 $\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} th_{ji}$ 为固定值, 故 $G'(\pi)$ 和 $G(\pi)$ 是等价的。

上述模型适用于集中式计算的协同多目标决策。此时, 无论是采用仿生算法还是拍卖算法, 均是在约束条件式(22)下生成决策方案, 然后进行改进, 直至形成最终决策方案。然而, 对于协同分布式计算的群集智能架构, 在进行学习时, 要根据式(4)确定 Agent 的影响集, 由定义 5 虚拟出对应的值, 以此计算 Agent 的个体效用函数。在此过程中, 往往需要虚拟出系统不存在若干 Agent 时系统的总效用值, 或求 Agent 对某目标集合进行等可能攻击时的期望值等, 直接使用上述模型会导致一些错误, 从而系统无法准确收敛。例如:

在一个决策方案中, 有导弹 M_{k_0} 和 M_{k_1} 联合攻击蓝机 B_j , 则 B_j 的被摧毁率为 $1 - (1 - th_{k_0j}) \cdot (1 - th_{k_1j})$ 。 k_0 和 k_1 消除的威胁为 $[1 - (1 - th_{k_0j})(1 - th_{k_1j})] \sum_{i=0}^{M-1} th_{ji}$ 。此时, 系统中其他情况不变, 计算没有 M_{k_0} 时 M_{k_1} 消除的威胁和没有

M_{k_1} 时 M_{k_0} 消除的威胁, 分别为 $th_{k_1j} \sum_{i=0}^{M-1} th_{ji}$ 和 $th_{k_0j} \sum_{i=0}^{M-1} th_{ji}$ 。此二者消除的威胁之和大于协同攻击消除的威胁, 即单个攻击的效果有可能被放大。在群集智能架构中, 这会导致 Agent 学习过程中对错误信息进行强化, 从而收敛到不可接受解。为解决此问题, 本文直接以导弹为基本变量, 定义了协同攻击时 Agent 的贡献度, 在此基础上构建全局效用函数。

设关于红、蓝双方及导弹的设定、约束条件不变。 $(1 - th_{kj})$ 为导弹 M_k 攻击蓝机 B_j 后 B_j 的存活率。 $1 - \prod_{k=0}^{Z-1} (1 - th_{kj})^{X_{kj}}$ 为一次协同攻击后, 蓝机 B_j 的被摧毁率, $[1 - \prod_{k=0}^{Z-1} (1 - th_{kj})^{X_{kj}}] \sum_{i=0}^{M-1} th_{ji}$ 为一次协同攻击后蓝机 B_j 被消除的总威胁。当各个导弹是分布式并行决策时, 为了衡量不同情况下导弹的攻击效果, 定义导弹对摧毁目标的贡献度如下。

定义 8 在一次协同攻击中, $\forall k, \exists j, X_{kj} = 1$, 以字母 d 建立索引, 记所有 $X_{kj} = 1$ 的导弹为 $k_0, k_1, \dots, k_d, \dots$, 且 $\forall d, th_{k_dj} > th_{k_{d+1}j}$ 。导弹 k_d 对其攻击目标被摧毁率的贡献度 $cont_{k_d}$ 为

$$\left. \begin{aligned} cont_{k_d} &= th_{k_dj} (1 - th_{index_{k_d}j}) \\ th_{index_{k_{d+1}}j} &= th_{index_{k_d}j} + cont_{k_d} \\ th_{index_{k_0}j} &= 0 \end{aligned} \right\} \quad (24)$$

贡献度定义明确了在一次协同攻击中, 单独攻击或协同攻击时 Agent 对攻击效果的贡献。式(24)定义的主要依据是空战中的优先攻击分配原则^[22]。对于无明确分配原则的 Agent 系统, 可定义 Agent 的贡献度为其在所有协同完成同一目标的 Agent 完成率之和中所占的比例。以本文场景为例, 导弹 M_k 对其攻击目标被摧毁率的贡献度为

$$cont_k = \frac{th_{kj}}{\sum_{\substack{p=0 \\ X_{pj} \neq 0}}^{Z-1} (X_{pj} th_{pj})} \quad (25)$$

式中: p 为用于 Agent 计数的下标; $X_{pj} \neq 0$ 保证了分母不为 0。 $cont_k \in (0, 1]$ 。

在式(24)基础上, 构建全局效用函数为

$$G''(\pi) = \sum_{k=0}^{Z-1} \left\{ \sum_{j=0}^{N-1} [X_{kj} \cdot cont_k \cdot \sum_{i=0}^{M-1} th_{ji}] \right\} \quad (\pi \in \Omega)$$

$$\text{s. t. } \begin{cases} \sum_{k=0}^{Z-1} X_{kj} \leq 2, \quad \forall j \in \{0, 1, \dots, N-1\} \\ \sum_{j=0}^{N-1} X_{kj} = 1, \quad \forall k \in \{0, 1, \dots, Z-1\} \end{cases} \quad (26)$$

易证明, $G''(\pi) = G'(\pi)$ 。同理可以构建基于式(25)的全局效用函数。

$$G'''(\pi) = \sum_{k=0}^{Z-1} \left\{ \sum_{j=0}^{N-1} \left\{ X_{kj} \cdot \frac{th_{kj}}{\sum_{\substack{p=0 \\ X_{pj} \neq 0}}^{Z-1}} (x_{pj} \cdot th_{pj}) \right. \right. \\ \left. \left. \left[1 - \prod_{q=0}^{Z-1} (1 - th_{qj})^{X_{qj}} \right] \sum_{i=0}^M th_{ji} \right\} \right\} \quad (\pi \in \Omega) \quad (27)$$

式中: q 为用于 Agent 计数的下标。约束条件同式(26)。两个全局效用函数模型的区别在于贡献度是根据优先攻击分配原则确定还是根据威胁大小的比例关系确定, 本文实验采用式(24)计算贡献度。

式(26)是在特定时间片内, 基于双方态势信息构建的全局效用函数。在任一时间片内, 设双方态势信息为静态数据, 则在约束条件下, 式(26)可唯一地确定一个全局效用函数值。满足群集智能中全局效用的定义。

Agent 的贡献度明确确定了 Agent 在协同攻击中的贡献, 故 Agent 对其个体效用的学习及行为选择时, 可不考虑其他 Agent 的行动, 即 Agent 间相互独立。此时, Agent 的影响集仅为自身, 基于式(26), 由式(5)可计算出每个 Agent 的个体效用的完美效用值。由定理 1, 完美效用值满足因子性, 由 Agent 间相互独立和定义 2, 计算出的完美效用值具有较高的易学性。

3 协同多目标攻击决策分析

现代战争已经演化为“军事网格”与“军事网格”的对抗。在信息化战争中, 网格内各作战单元决策的核心依据是最大化全局效用, 在此基础上, 形成自主协同攻击决策方案^[12]。图 2 是协同多目标攻击空战态势。

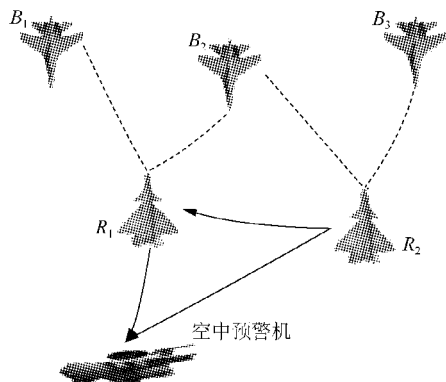


图 2 协同多目标攻击战术分析

Fig. 2 Analysis of coordinated multi-target attack tactics

图 2 中, 红方机群 R 包含 R_1, R_2 , 以及空中预警机, 通过数据链实现信息共享, R_1 和 R_2 各挂载 2 枚同型导弹; 蓝方机群 B 包含 B_1, B_2 和 B_3 。 R 采用协同战术与 B 空战。协同多目标攻击决策的目标是要找到一组导弹-目标分配对, 使得在所有导弹攻击后, 所消除的总威胁最大。

数据链的发展使得在战机间和战机与雷达、预警机间建立低延迟、高带宽、灵活的网络架构成为可能, 这可使战机间实时地交换信息和战术目标, 进行自主计算, 形成自主协同。图 2 中 R_1 和 R_2 协同攻击 B_1, B_2 和 B_3 , 以求达到全局效用最优。在决策过程中, 直接以全局效用最优作为 Agent 分配目标的依据, 采用群集智能架构, Agent 以个体效用函数为分配目标的依据, 因个体效用具有因子性和高的易学性和智能性, Agent 很容易学习提高个体效用, 从而提高总体效用。从理论上可避免“公共地悲剧”问题。

同时, 群集智能架构使用分布式计算, 无中心控制节点, Agent 通过信息共享进行自主计算, 各 Agent 的计算量及通信节点的计算量相比集中式有较大地减少, Agent 自主计算个体形成决策方案, 系统具有较好的扩展性。但对通信的带宽和速度有更高的要求。

无人机采用顺序拍卖方法形成目标分配方案^[8-9]和仿生算法中以武器对目标的威胁值计算分配值, 然后以分配值为启发信息形成初始攻击方案并进行局部优化, 最终确定攻击决策方案的方法^[7], 均易导致“公共地悲剧”问题。例如, 假设图 2 中, B_1, B_2 和 B_3 对机群 R 的总威胁分别为 (1.20, 1.50, 1.00), R_1 和 R_2 对 B_1, B_2 和 B_3 的威胁分别为 (0.90, 0.95, 0.50), (0.50, 0.85, 0.80)。以仿生算法为例, 仿生算法以分配值为启发信息产生初始方案, 简单说, 就是每架红机攻击其对蓝机机群中威胁最大的蓝机。此时易产生攻击方案: R_1 攻击 B_1 和 B_2, R_2 攻击 B_2 和 B_3 , 总效用为 $0.90 \times 1.20 + [1 - (1 - 0.95) \times (1 - 0.85)] \times 1.50 + 0.80 \times 1.00 = 3.36875$ 。但是, 假设 R_1 攻击 B_1 和 B_2 不变, R_2 连续攻击 B_3 , 则总效用为 $0.90 \times 1.20 + 0.95 \times 1.50 + [1 - (1 - 0.80) \times (1 - 0.80)] = 3.46500$, 后产生的方案全局效用大于按照仿生算法形成方案的全局效用, 仿生算法中 R_2 攻击其最有利攻击目标且目标对 R 总威胁最大, 但全局效用减少了。

产生“公共地悲剧”问题的原因是 Agent 间存在相互影响, 而 Agent 在决策时未将自己的个体目

标与全局目标严格一致化。此外,上述仿生算法和拍卖方法通信方式均为有中心的网络结构,仅能由中心节点计算并将结果发给各个作战单元,易产生延迟和不易扩展,对中心节点的计算要求高,且易受敌方攻击^[12],在应用中有很大的局限。

4 基于群集智能的协同多目标攻击决策算法

4.1 基于群集智能的算法设计

传统的群集智能架构主要解决同构多 Agent 协同问题,甚至有些系统中 Agent 初始位置集中或初始位置无关,相应全局效用函数形式可表示为

$$G(\boldsymbol{\pi}) = \sum_j f_j(\sum_k X_{kj}) \quad (28)$$

式中: $f_j(\cdot)$ 为关于分配给第 j 个目标的 Agent 数量的函数,全局效用为各关于 Agent 数量的函数的线性和。例如,从某基地发射导弹对目标区域的若干威胁目标进行打击,若各导弹相同且时间上具有一定的同步性,即在攻击决策时可视为同时攻击。此时,世界线 \boldsymbol{z} 可简单地定义为一个 $(N+1) \times M$ 矩阵, N 为目标数量, M 为导弹总数。若 Agent k 对目标 B_j 攻击, $X_{kj}=1$,否则 $X_{kj}=0$, \boldsymbol{z} 中 $j=0$ 表示导弹未对任何目标进行攻击。令 $x_j(\boldsymbol{z}) = \sum_{k=0}^{Z-1} X_{kj}$,表示所有攻击 B_j 的导弹数量。则全局效用函数可表示为

$$G(\boldsymbol{z}) = \sum_{j=0}^N x_j(\boldsymbol{z}) e^{\frac{x_j(\boldsymbol{z})}{c_j}} \quad (29)$$

式中: c_j 为攻击 B_j 所需的最优导弹数量; c_0 为基地中不需要进行攻击的最优导弹数量, $c_0 = Z - \sum_{j=1}^N c_j$ 。若攻击前目标信息已知,则 c_j 信息可计算得出,此时,全局效用函数 $G(\boldsymbol{\pi})$ 为关于攻击每个 B_j 的 Agent 数量 $x_j(\boldsymbol{z})$ 的函数的线性和。同时,此函数可以评价系统所有可能的情况,满足群集智能理论条件。

对于 Agent 集位置分布、异构,目标集位置分布的协同多 Agent 系统,全局效用函数形式可表示为

$$G(\boldsymbol{\pi}) = \sum_j h_j(g_i(\sum_l f_{li}(\sum_{l,i} X_{lk,j}))) \quad (i \in A) \quad (30)$$

式中: k_i 为类型(或角色)为 i 的 Agent; l 为 Agent 分布的位置; $X_{lk,j}=1$ 表示位于 l ,类型为 i

的 Agent 分配给目标 j , $X_{lk,j}=0$ 为未分配给 j ; $f_{li}(\cdot)$ 为关于位于地点 l 的 i 型 Agent 的函数; $g_i(\cdot)$ 为关于 Agent 不同类型的函数; $h_j(\cdot)$ 为分配给目标 j 的关于 Agent 不同类型函数的函数。以本文场景为例,挂载于不同战机上的导弹可攻击不同的目标,由于导弹对目标的威胁与战机效能等相关,故挂载于不同战机的导弹对同一目标的威胁是不同的,若以导弹为 Agent 对决策问题进行建模,则是异构分布式多 Agent 协同系统。与式(28)表述的示例相同,世界线 \boldsymbol{z} 可简单地定义为一个 $(N+1) \times M$ 矩阵, N 为目标数量, M 为导弹总数,若 Agent k 对 B_j 攻击, $X_{kj}=1$,否则 $X_{kj}=0$, \boldsymbol{z} 中 $j=0$ 表示导弹未对任何目标进行攻击。以式(21)为参考可相应地给出 $f_{li}(\cdot)$, $g_i(\cdot)$ 和 $h_j(\cdot)$,因在本文前述假定条件下,位于不同战机的导弹视为不同的 Agent,不区分不同的 Agent 类型,所以本例中的 $f_{li}(\cdot)$ 和 $g_i(\cdot)$ 可以简化为一个函数 $g'(\cdot) = 1 - \prod_{k=0}^{Z-1} (1 - \text{th}_{kj})^{X_{kj}}$, 而 $h_j(\cdot) = \sum_{i=0}^{M-1} \{\text{th}_{ji} g'(\cdot)\}$,其全局效用函数直观地可表示为式(21)。由 2.2 节分析可知,上述效用函数不适用于群集智能架构,其原因是 Agent 异构导致 Agent 学习过程中对错误信息进行强化,从而收敛到不可接受解。为使群集智能架构可以解决异构 Agent 问题,基于定义的 Agent 的贡献度,可转换全局效用函数为式(26)。

有了全局效用函数,Agent 的个体效用函数,可由式(5)导出。在 Agent 的内部,采用式(10)定义的 Q 学习算法对个体效用函数进行学习。Agent 学习个体效用函数的步骤:

① 由 Boltzmann 函数计算 Agent 行动空间集(Q学习的 Q 表)中每个行动被选取的概率,Agent 的行动空间为 Agent 可攻击的目标集。

② 由系统产生的随机数与 Boltzmann 函数计算出的概率比较,确定选取的行动。

③ 由迭代次数对 Boltzmann 函数参数进行调整以控制收敛速度。

④ Agent 将选取的行动通信给通信节点进行计算,通信节点计算各个 Agent 的个体效用并返回结果。

⑤ 根据获取的个体效用值对 Q 表进行更新。群集智能架构基本步骤:

- ① 选取某个通信节点。
- ② 收集 Agent 的行动。

- ③ 计算全局效用函数值。
- ④ 计算 Agent 的个体效用函数值。
- ⑤ 将 Agent 的个体效用值传送给 Agent。

本文中,通信系统假定所有参战的作战单元组成一个 Ad hoc 网络,通过相邻节点将信息传输给通信节点。通信节点的选取通过一个动态的列表确定,列表根据态势信息动态确定和更新,作战区域内的作战单元共享此列表。系统运行初期首选项设为空中预警机,其他项由具备计算和通信能力的武器平台组成。

4.2 启发式信息

空战中的多目标协同攻击是一个有约束的组合优化问题,属于非确定性多项式(NP)完全问题。随着问题输入规模的扩大,求解时间呈几何级数上升。为了加快求解速度,使解在一定准确度内快速收敛,本文引入两类启发信息:构造性启发和改进性启发。

定义 9 构造性启发是指利用问题的特定知识,如经验法则、有根据的猜测、常识等构造满足一定条件的解的方法。

定义 10 改进性启发是指利用历史信息、探索和试错、评价函数等对已经生成的解进行局部优化的方法。

基于定义的两类启发式信息,引入攻击范围(定义 11)和协同攻击比(定义 12)来进行启发式改进。

Agent 的攻击范围 AS_i 为红机 R_i 所有攻击胜率大于阈值 S_{th} ($S_{th} \in [0, 1]$) 的蓝机。即

定义 11 $AS_i = \{B_j \mid th_{ij} \geq S_{th}, B_j \in B, S_{th} \in [0, 1]\}$

挂载于同一架飞机的导弹定义为一个 Agent 组,Agent 组的协同攻击比定义为组内不能明确确定攻击目标,需要通过试探等方法确定攻击目标的 Agent 的比例。即

定义 12 Agent 组的主动协同攻击比例为 p_{co} ,即所有 Agent 组内成员有 p_{co} , $p_{co} \in [0, 1]$ 的成员主动试探与其他飞机协同攻击。而 $1 - p_{co}$ 的成员则按照优先攻击分配原则确定攻击目标。

直观地看,定义 11 的主要依据是战机在空战过程中往往聚焦在满足一定攻击条件的目标集上,例如在一定距离内或攻击角小于某角度时,而不会攻击空域内的所有敌机,这属于构造性启发,将缩小解的搜索空间,当 $S_{th} = 0$ 时,Agent i 的攻击范围为所有蓝方目标。当 $S_{th} = 1$ 时,Agent i

攻击范围为 0,即不参与攻击。定义 12 的主要依据是挂载在同一架飞机上的导弹在攻击时往往按照优先攻击分配原则选取目标,优先攻击相对于友机更有利于攻击的目标和对己方机群威胁较大的目标。主动协同比确定了有多少 Agent 主动试探与其他战机协同攻击目标,而不是按照优先攻击分配原则从攻击范围内直接选取目标,属于改进性启发,当 $p_{co} = 0$ 时,Agent 组内成员(同一架飞机上挂载的导弹)依次选取攻击其攻击范围内威胁最大的目标。当 $p_{co} = 1$ 时,Agent 所有组内成员将在攻击范围内试探与其他导弹协同攻击某目标。

在 Agent 的学习过程中,可利用计算攻击目标集中所有目标的期望来提高准确度和精确度。本文中,定义了基于计算期望的 WLUA 以提高准确度。

$$WLUA = G(z) - G\left(E\left(\sum_{a=0}^{|A|} \frac{1}{|A|} CL_{S_{th}}^{off}(z)\right)\right) \quad (31)$$

式中: $|A|$ 为所有可能的行动数; a 为具体行动。

带有启发信息的群集算法如下:

```

{
    初始化;
    连接 Agent;
    while(不满足迭代终止条件){
        攻击矩阵 = Agent 选择目标();
        计算全局效用函数(){
            for 每个 Agent(导弹){
                由式(24)计算 Agent 对攻击目标的贡献度;
            }
            计算全局效用函数;
            计算每个 Agent 的效用函数;
            将个体效用传递给 Agent;
        }
        输出结果;
    }
}

```

Agent 内部学习算法:

```

{
    初始化,设置相关参数  $t, d$ ;
    计算  $AS_i, p_{co}$ ;
    初始化 Agent(导弹)  $k$  的 Q 表;
    选择目标(){

```

根据式(11)计算行动空间(Q表)中每个行动被选取的概率;

产生随机数,确定选取行动 a ;

将行动 a 传递给通信节点;

根据迭代次数等同步信号更新参数;

}

根据回传的WLU或WLUA更新Q表;

}

5 仿真实验

双方的交战态势如文献[7],设红方由陆基雷达或空中预警机发现某空域有蓝方14架飞机来袭,出动4架具有多目标攻击能力的战机对其进行拦截。蓝方以双机一字和机翼形为基本单元组成机群编队,相邻两蓝机左右间距5 km,前后间距10 km。每架红机携带4枚中远程空空导弹,以双机一字形为基本单元组成飞行编队以增强攻击力,通过新型高速数据链进行机群间及与陆基雷达和预警机间的信息共享,相邻两红机左右间距20 km,前后间距5 km。设在某一时刻空战态势如图3所示,红、蓝机群均在同一高度飞行,红方所有目标均在红方机群的联合攻击区内,红、蓝各机编号如图中所示。红方机群首先对蓝方所有目标发动协同攻击。红方机群R的总导弹量为 $Z=16$, R_1, R_2, R_3 和 R_4 携带的导弹编号分别为第0~3,4~7,8~11和12~15。

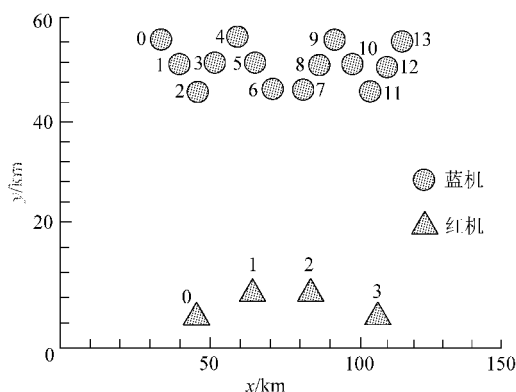


图3 红、蓝双方机群空战态势图

Fig. 3 Air combat situation between R and B

根据上述双方态势进行仿真实验,为消除实验过程中因随机原因出现的偶然情况,本文的所有实验数据均是各算法运行100次所取的平均。实验选取的 Boltzmann 温度参数为1,衰减率指每次迭代后,温度衰减为前此温度的比例。表1和表2分别列出了WLU和WLUA在标准、带有攻击范围、

带有攻击范围和主动协同比3种情况下的结果。

表1 WLU实验结果

Table 1 Results of WLU

衰减率	攻击范围	协同比	收敛到95%迭代次数	最终收敛值	最终收敛值的迭代次数
0.90	—	—	—	0.840	33
0.99	—	—	—	0.900	93
0.90	0.4	—	—	0.920	24
0.99	0.4	—	—	0.940	60
0.90	0.4	0.25	5	0.990	25
0.99	0.4	0.25	8	0.993	61
0.99	0.4	0.50	13	0.981	66

表2 WLUA实验结果

Table 2 Results of WLUA

衰减率	攻击范围	协同比	收敛到95%150次迭代迭代次数	150次迭代收敛值	500次迭代收敛值
0.90	—	—	50	0.960	0.965
0.99	—	—	97	0.970	0.975
0.90	0.4	—	21	0.987	0.991
0.99	0.4	—	48	0.984	0.994
0.90	0.4	0.25	5	0.992	0.992
0.99	0.4	0.25	6	0.992	0.992
0.99	0.4	0.50	15	0.995	0.997

表1和表2中攻击范围由战机的作战性能决定,选取值较大时,攻击范围小,选取值较小时,攻击范围大。本文中选取0.4,即红机将所有威胁超过0.4(红机对蓝机威胁)的蓝机列入攻击范围内。

表1是对式(5)确定的WLU的实验验证,3种情况均能很快收敛,在探索策略温度参数不变的情况下($t=1$),衰减率值越小,即每次温度衰减越快,则收敛越快,但收敛准确度越低。这是因为Agent对其行为集的探索学习次数少,达到“完全理性”的可能性较小。调高衰减率值,可增加Agent的学习探索次数,达到收敛的迭代次数增加了,但可提高收敛准确度。加入构造性启发式信息后,可明显提高收敛速度和准确度,而加入两类启发信息后,系统仅需5次就可收敛到全局最优值95%以上。稳定收敛后可达到全局最优值99.3%。提高协同攻击比增加了需主动协同的Agent数量,对并行决策在探索学习速度有影响,收敛速度有一定的降低,收敛到最优值95%的迭代次数由8增加到13。同时,准确度有一定的下降,这是因为在保证收敛速度的情况下,增加主动协同攻击Agent数量易使Agent关于个体效用

函数学习中智能性有所降低。

表 2 是对式(31)确定的 WLUA 的实验验证,3 种情况均能快速收敛到最优值 95%以上,随着实验次数的增加,3 种情况均能稳定地增加全局效用值。Boltzmann 温度衰减率的作用与 WLU 相同。加入构造性启发信息后,收敛速度和准确度有了较大的提高,加入两类启发信息后,仅需 5 次就可收敛到最优值 95%以上。提高主动协同攻击比,收敛速度有了一定的下降,但准确度却有了提高,这是因为 Agent 计算的是关于打击目标的期望。收敛速度的下降是因为增加了主动协同攻击 Agent 的数量,准确度的提高是因为低的协同攻击比在用计算期望进行学习时,易使协同攻击的 Agent 为防止有目标未分配导弹而强化对此类目标的学习,从而降低 Agent 关于个体效用的智能性,达不到完全“理性”,没有收敛到最优行为。而适当放宽主动协同比,则更易使 Agent 通过对个体效用的学习达到理性,从而最终收敛到最优值。

图 4(a)和(b)分别给出了基于 WLU 和 WLUA 的标准、带有攻击范围(前缀 A,如 AWLU,下同)、带有攻击范围和主动协同比(前缀 AS),Boltzmann 参数为 $t=1$,衰减率为 0.99 情况下的实验结果。

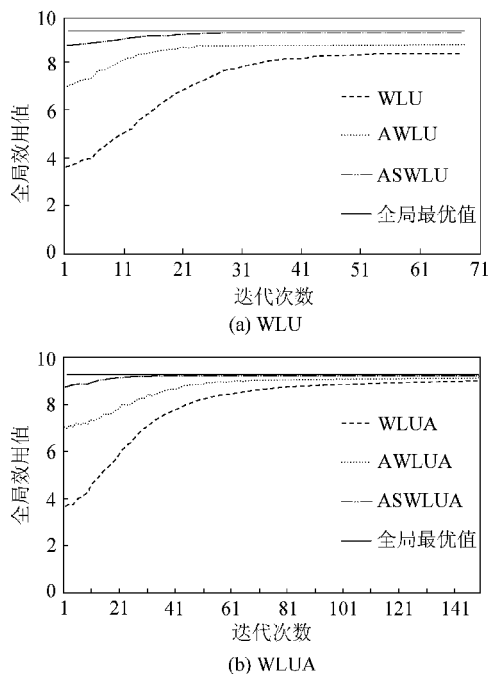


图 4 基于 WLU, WLUA 算法的实验结果

Fig. 4 Results of algorithms based on WLU and WLUA

由图 4 可见,对于基于 WLU 的个体效用函数,Agent 均能较快收敛,但收敛值不高。在加入

两类启发信息后,可快速收敛到 95%以上的性能。对于基于 WLUA 的个体效用函数,Agent 相对收敛较慢,但准确性有了较大提高,即使是标准的 WLUA,也能在 50 次后收敛最优值的 95%以上,且 3 种情况均稳定的向全局最优值逼近。加入启发式信息后,收敛速度有了很大的提高,带有攻击范围的 WLUA 可以在 21 次收敛到最优值的 95%以上,而同时加入两类启发信息的 WLUA 更是提高到了 5 次。

表 3 ASWLU 与其他算法的比较

Table 3 Comparison of ASWLU with other algorithms

算 法	参 数	收敛迭代 次数	收敛值	生成解 的个数
HPSO ^[6]	50 粒子	72	0.980	3 600
HGA ^[29]	50 基因	78	0.898	3 900
HACA ^[7]	30 蚂蚁	61	0.999	1 830
ASWLU	0.99	61	0.993	61

由表 3 可知,WLU 计算量总体上最小,通过与友机共享信息形成自主协同,减少了盲目搜索的次数和计算量。当学习次数较大时,采用 WLUA 可获得准确度非常高的攻击决策方案。在群集智能算法中,解的准确度取决于 Agent 的智能性,Agent 完全“理性”时,系统收敛到最优解,当 Agent 行动集较小时,可很快收敛到最优解,具有很好的稳定性,各 Agent 自主计算确定各自的攻击方案,系统具有良好的扩展性。

在实际应用中,群集智能架构解决协同多目标攻击问题时,为减少通信次数,可通过一次发送多组迭代值进行计算,从而迅速收敛。

6 结 论

为加快“从传感器到射手”的决策速度,增强系统的健壮性和扩展性,提出基于群集智能的算法来解决协同多目标攻击问题,构建的全局效用函数和基于 Agent 贡献度计算得出个体效用函数满足群集智能理论,从全局而言,可实现任意 Agent 任何提高个体效用函数值的行动均会导致全局效用函数值的提高,所以系统在理论上可避免“公共地悲剧”问题。同时,算法结合数据链,充分利用各武器平台本身的自主计算和通信能力,采用协同分布式计算方式,使各武器平台共享全局态势信息,并仅需聚焦在各自可攻击范围内进行独立决策,实现自主协同攻击。分布式的计算使得网络结构更加灵活,提高了系统的稳定性和健壮性,聚焦在自己攻击范围内并利用武器平台

自身的计算能力,减少了系统对强计算能力的要求和依赖,提高了系统的可扩展性。两类启发式信息减少了盲目搜索,提高了 Agent 对行为集的学习速度,加快了收敛,也降低了计算量。实验证实了上述结论。

下一步工作中,将重点研究以下两个方面:

(1) 动态不确定环境下建立动态模型,解决连续协同攻击决策问题。

(2) 在态势信息不完全情况下,即参战武器平台仅能与部分其他武器平台进行通信和信息共享,建立参战武器平台或作战单元协同多目标攻击决策模型,并研究解决方法。

参 考 文 献

- [1] 李林森, 佟明安. 协同多目标攻击空战决策及其神经网络实现[J]. 航空学报, 1999, 20(4): 309-312.
Li Linsen, Tong Ming'an. Air combat decision of cooperative multi-target attack and its neural net realization[J]. Acta Aeronautica et Astronautica Sinica, 1999, 20(4): 309-312. (in Chinese)
- [2] 高文正, 丁全心. 多机协同攻击中的目标分配研究[J]. 电光与控制, 2003, 10(4): 10-14.
Gao Wenzheng, Ding Quanxin. Target allocation in multi-fighter cooperative attacking[J]. Electronics & Control, 2003, 10(4): 10-14. (in Chinese)
- [3] 程红斌, 张凤鸣, 张晓丰. 多机协同空战目标分配算法[J]. 空军工程大学学报: 自然科学版, 2005, 6(2): 7-10.
Cheng Hongbin, Zhang Fengming, Zhang Xiaofeng. Multi-aircraft coordination target assignment[J]. Journal of Air Force Engineering University, Natural Science Edition, 2005, 6(2): 7-10. (in Chinese)
- [4] 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.
- [5] 罗德林, 王彪, 龚华军, 等. 基于 SAGA 的协同多目标攻击决策[J]. 哈尔滨工业大学学报, 2007, 39(7): 1154-1158.
Luo Delin, Wang Biao, Gong Huajun, et al. Air combat decision-making for cooperative multiple target attack based on SAGA[J]. Journal of Harbin Institute of Technology, 2007, 39(7): 1154-1158. (in Chinese)
- [6] Luo D L, Yang Z, Duan H B, et al. Heuristic particle swarm optimization algorithm for air combat decision-making on CM-TA[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2006, 23(1): 20-26.
- [7] 罗德林, 段海滨, 吴顺祥, 等. 基于启发式蚁群算法的协同多目标攻击空战决策研究[J]. 航空学报, 2006, 27(6): 1166-1170.
Luo Delin, Duan Haibin, Wu Shunxiang, et al. Research on air combat decision-making for cooperative multiple target attack using heuristic ant colony algorithm[J]. Acta Aeronautica et Astronautica Sinica, 2006, 27(6): 1166-1170. (in Chinese)
- [8] Pongpunwattana A, Rysdyk R. Real-time planning for multiple autonomous vehicles in dynamic uncertain environments[J]. Journal of Aerospace Computing, Information, and Communication, 2004, 1(12): 580-604.
- [9] 廖沫, 陈宗基. 基于多 Agent 分布协同拍卖的动态目标分配算法[J]. 北京航空航天大学学报, 2007, 33(2): 180-183.
Liao Mo, Chen Zongji. Dynamic target assignment method based on multi-agent decentralized cooperative auction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(2): 180-183. (in Chinese)
- [10] 廖沫, 陈宗基, 周锐. 基于 MAS 的多 UAV 协同任务分配设计与仿真[J]. 系统仿真学报, 2007, 19(10): 2313-2317.
Liao Mo, Chen Zongji, Zhou Rui. Design and simulation of multi-UAV coordinated task allocation based on MAS[J]. Journal of System Simulation, 2007, 19(10): 2313-2317. (in Chinese)
- [11] Hardin G. The tragedy of the commons[J]. Science, 1968, 162(3859): 1243-1248.
- [12] United States Department of Defense. Joint vision 2020 [EB/OL]. [2008-06-08]http://www.iwar.org.uk/military/resources/aspc/text/strategy/jv2020.htm.
- [13] Rockwell collins TTNT successfully demonstrated at JEFX 2006[EB/OL]. [2008-06-08]http://www.rockwell-collins.com/news/page7868.html.
- [14] Wolpert D H, Tumer K. An introduction to collective intelligence[R]. NASA-ARC-IC-99-63, 1999.
- [15] Agogino A K. Design and control of large collections of learning agents[D]. Austin: The University of Texas at Austin, 2003.
- [16] Agogino A K, Tumer K. Handling communication restrictions and team formation in congestion games[J]. Autonomous Agents and Multi-Agent Systems, 2006, 13(1): 97-115.
- [17] Fogel D B, Robinson C J. Computational intelligence: the experts speak [M]. Piscataway: IEEE Press, 2003.
- [18] 李光久. 博弈论基础教程[M]. 北京: 化学工业出版社, 2005.
Li Guangjiu. Game theory[M]. Beijing: Chemical Industry Press, 2005. (in Chinese)
- [19] 高阳, 陈世福, 陆鑫. 强化学习研究综述[J]. 自动化学报, 2004, 30(1): 86-100.
Gao Yang, Chen Shifu, Lu Xin. Research on reinforcement learning technology: a review[J]. Acta Automatica Sinica, 2004, 30(1): 86-100. (in Chinese)
- [20] 高永, 向锦武. 一种新的超视距空战威胁估计非参量法模型[J]. 系统仿真学报, 2006, 18(9): 2570-2572, 2592.
Gao Yong, Xiang Jinwu. New threat assessment non-parameter model in beyond-visual-range air combat [J].

- Journal of System Simulation, 2006, 18(9): 2570-2572, 2592. (in Chinese)
- [21] 高永, 向锦武. 超视距多机协同空战目标分配算法[J]. 北京航空航天大学学报, 2007, 33(3): 286-289.
Gao Yong, Xiang Jinwu. Target assignment in BVR air combat[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(3): 286-289. (in Chinese)
- [22] Lazarus E. The application of value-driven decision-making in air combat simulation[C]// 1997 IEEE International Conference on Systems, Man, and Cybernetics. 1997, 3: 2302-2307.
- [23] Luo D L, Shen C L, Wang B, et al. Air combat decision-making for cooperative multiple target attack using heuristic adaptive genetic algorithm[C]// Proceedings of 2005 International Conference on Machine Learning and Cybernetics. 2005, 1: 473-478.
- 作者简介:**
刘波(1979—)男, 博士研究生。主要研究方向: 移动计算, 群集智能。
Tel: 029-82663979
E-mail: dongdu@gmail.com
- 覃征(1956—)男, 博士, 教授, 博士生导师。主要研究方向: 移动计算, 数据融合等。
Tel: 029-82663979

(编辑: 张利平, 孔琪颖)