

基于精英改选机制的粒子群算法的空战纳什均衡策略逼近

王 昱^{1,2†}, 章卫国¹, 傅 莉², 黄得刚¹, 李 勇³

(1. 西北工业大学 自动化学院, 陕西 西安 710129; 2. 沈阳航空航天大学 自动化学院, 辽宁 沈阳 110136;

3. 沈阳工业大学 电气工程学院, 辽宁 沈阳 110870)

摘要: 针对无人机协同攻击的动态多策略性, 应用纳什均衡概念, 考虑联合生存概率和武器消耗等因素, 融合双方价值函数计算和双矩阵对策纳什均衡点的求解方法, 建立一种多战斗步空战动态目标分配优化模型. 提出基于精英改选机制的粒子群(elite re-election particle swarm optimization, ERPSO)算法, 在群体极值引导能力不足时, 通过对其克隆、变异和重新初始化等操作增加个体的多样性, 保留传统粒子群算法结构简单、快速收敛等优点, 改善算法易于陷入局部极小的问题. 将ERPSO算法应用于目标分配模型求解纳什均衡点, 获取更为精确的双方混合策略, 确保实时性和准确性, 验证了模型和方法的有效性.

关键词: 纳什均衡; 动态目标分配; 优化; 精英改选粒子群; 多样性; 混合策略

中图分类号: V247.1; TP301.6 **文献标识码:** A

Nash equilibrium strategies approach for aerial combat based on elite re-election particle swarm optimization

WANG Yu^{1,2†}, ZHANG Wei-guo¹, FU Li², HUANG De-gang¹, LI Yong³

(1. School of Automation, Northwestern Polytechnical University, Xi'an Shaanxi 710129, China;

2. School of Automation, Shenyang Aerospace University, Shenyang Liaoning 110136, China;

3. School of Electrical Engineering, Shenyang University of Technology, Shenyang Liaoning 110870, China)

Abstract: A multi-combat step dynamic target assignment optimization model is built based on Nash equilibrium concept by considering the joint survival probability and weapons consumption factors for dynamic multi-strategy UAV cooperative attack. In building the model the value function calculation for the two belligerent parties and the solving method of bimatrix game Nash equilibrium point are applied. Then, an elite re-election particle swarm optimization (ERPSO) is proposed based on the elite re-election mechanism; therefore, the diversity of individuals can be increased when leading capacity of group extreme value is deficient through cloning, mutation and re-initialization operation. The advantages of traditional particle swarm optimization (PSO) algorithm such as simple structure, fast convergence are retained, but the shortcoming of falling into local minimum is corrected. Finally, the ERPSO algorithm is applied to the dynamic target assignment model to solve the Nash equilibrium point for obtaining accurate hybrid strategies for both sides. It is confirmed that the real-time demands and accuracy are satisfied, and the validity of the proposed model and method is established.

Key words: Nash equilibrium; dynamic target assignment; optimization; ERPSO; diversity; mixed strategies

1 引言(Introduction)

多无人作战飞机(unmanned combating air vehicle,UCAV)在实现信息共享的前提下, 协同空战能够完成各种复杂的攻势作战任务, 引起各国的广泛关注, 这也使得多平台多武器火力分配(weapon-target assignment, WTA)问题成为航空领域研究的热点. 目前有关WTA问题的研究方法有随机算法(random strategy)、单元贪婪策略(unit greedy strategy)、团队优化

(team optimal strategy)等^[1], 它们各有弊端, 随机算法中每个单元的策略选择是随机的, 很难保证其收敛性; 贪婪算法是基于每个单元对自身的性能优化, 不能满足作战整体寻优, 也使得“协同”失去意义; 而团队优化只从本队性能优化的角度出发忽略了其他队伍的策略选择, 属于间接式对抗, 这不符合空战本质.

无人机协同空战是作战双方策略的对抗、竞争, 或面对一种局势时的对策选择, 决策的最终结果不但取

收稿日期: 2015-01-05; 录用日期: 2015-03-26.

†通信作者. E-mail: yuki1120@sina.com; Tel.: +86 13514207196.

国家自然科学基金项目(61374032), 辽宁省教育厅科学研究一般项目(L2015412), 沈阳市科技创新团队项目(src201204)资助.

Supported by National Natural Science Foundation of China (61374032), General Science and Technology Research Project of Department of Education of Liaoning (L2015412) and Science and Technology Innovation Team of Shenyang City (src201204).

决于本方,也取决于敌方决策,这非常符合博弈论所研究的范畴。WTA问题主要来源于指挥与控制(C2)自动化的要求^[2],作为现代无人机空战指挥决策中的关键问题,解决如何合理地部署我方兵力来迎击敌方兵力,以期达到最佳的作战效果。WTA问题按其战斗步骤的考量主要分为静态-WTA(SWTA)和动态-WTA(DWTA)^[3]。随着对无人机空战复杂性研究的深入,以多战斗步空战的DWTA逐渐取代了只考虑一步空战的SWTA,每一步的决策都是建立在对前一战斗步战况的评估上,这更符合现代空战实际。

为找寻DWTA问题的满意解,应用优化算法对DWTA问题进行求解是合适的。其中粒子群算法具有算法机制简单、性能好、容易实现等优点^[4],其收敛速度快的优点更是受到DWTA研究领域的青睐。然而,粒子群算法(particle swarm optimization, PSO)的缺点是全局搜索能力不强,目前,针对这方面的改进层出不穷,如调节控制参数平衡局部搜索能力和全局搜索能力^[5-7]、设计不同的邻域拓扑取代全局拓扑^[8-9]的改进算法,以及将辅助搜索策略引入传统PSO算法的混合算法^[10-13]。但这些算法在改变种群多样性的同时,却牺牲了收敛速度并使算法结构更加复杂^[14]。

为更好的描述无人机空战的动态对抗性,获取更为合理的对抗双方的目标分配方案,本文建立了新的多战斗步DWTA模型,提出了更为有效的优化求解方法,研究主要内容有:1)从博弈的思想出发,以联合生存概率及武器消耗为指标提出动态目标分配价值函数的新模型,在评估作战双方攻击策略的基础上,应用纳什均衡概念建立博弈目标函数。2)受到自然界和人类社会“领导能力考核、竞争、换届”等思想的启发,在保留传统粒子群算法的快速收敛性的基础上,引入“精英改选机制”,提出精英改选机制的粒子群(PSO based on elite re-election, ERPSO)算法,以增加个体多样性、加强引导能力,改善传统PSO算法易于陷入局部极小的问题;3)将ERPSO算法应用于所建立的目标分配模型的纳什均衡点的求解中,获取作战双方各自更为精确的概率意义上的混合协同攻击策略,为后续作战选择提供依据。

2 基于纳什均衡的动态目标分配模型(Dynamic target assignment model based on Nash equilibrium)

2.1 动态目标分配价值函数模型(Dynamic target assignment value function model)

将无人机空战双方设为蓝方、红方,从博弈的意义上看,交战双方是在不知道对方选择何种策略的情况下,选择各自的攻击策略,因此无论两方决策的时间是否真正同时,都可以视为是同时做出决策。在此前提下建立蓝、红两方单架无人机的存活概率计算函数

(假设战斗分 K 步进行,且同一作战飞机使用同一类型武器):

$$L_i^B(k) = L_i^B(k-1) \cdot \prod_{j=1}^{N_R} [1 - p_{ji}^R(k)]^{u_{ji}^R(k)}, \quad (1)$$

$$L_j^R(k) = L_j^R(k-1) \cdot \prod_{i=1}^{N_B} [1 - p_{ij}^B(k)]^{u_{ij}^B(k)}, \quad (2)$$

$$p_{ji}^R(k) = \beta_w \cdot p_{ji}^R \cdot L_j^R(k-1), \quad (3)$$

$$p_{ij}^B(k) = \beta_w \cdot p_{ij}^B \cdot L_i^B(k-1), \quad (4)$$

其中: N_B, N_R 为蓝、红两方战机架数; $u_{ij}^B(k), p_{ij}^B(k)$ 分别表示第 k 战斗步蓝方第 i 架战飞机攻击红方第 j 架战飞机使用的武器数目以及单一武器的毁伤概率, $u_{ji}^R(k), p_{ji}^R(k)$ 分别表示第 k 战斗步红方第 j 架战飞机攻击蓝方第 i 架战飞机使用的武器数目以及单一武器的毁伤概率; p_{ji}^B, p_{ji}^R 表示理想环境下(如天气等)武器的毁伤概率; β_w 为环境因素影响因子($0 \leq \beta_w \leq 1$)。 $L_i^B(k)$ 和 $L_j^R(k)$ 分别为第 k 个战斗步后,蓝方第 i 架无人机、红方第 j 架无人机的存活概率($L_i^B(0) = 1, L_j^R(0) = 1$ 表示双方初始生存概率均为1), $k = 1, 2, \dots, K$ 。

由此可见,本方此刻的生存概率大小依赖己方和对手在前一时刻的生存概率,这符合实际。因此,可建立蓝、红两方联合生存概率函数分别为

$$\sum_{i=1}^{N_B} L_i^B(K), \quad \sum_{j=1}^{N_R} L_j^R(K).$$

文献[15]以我方联合生存价值最大、敌方联合生存价值最小,结合目标价值参数建立了动态目标分配模型,未考虑本方武器消耗较少等损耗因素,本文综合武器消耗,加入规范化因子,建立蓝、红双方作战总收益函数:

$$J^B = \frac{\mu_L^B \sum_{i=1}^{N_B} c_i^B L_i^B(K)}{\mu_u^B \sum_{k=1}^K \sum_{j=1}^{N_R} \sum_{i=1}^{N_B} u_{ij}^B(k)} - \frac{\mu_L^R \sum_{j=1}^{N_R} c_j^R L_j^R(K)}{\mu_u^R \sum_{k=1}^K \sum_{i=1}^{N_B} \sum_{j=1}^{N_R} u_{ji}^R(k)}, \quad (5)$$

$$J^R = \frac{\mu_L^R \sum_{j=1}^{N_R} c_j^R L_j^R(K)}{\mu_u^R \sum_{k=1}^K \sum_{i=1}^{N_B} \sum_{j=1}^{N_R} u_{ji}^R(k)} - \frac{\mu_L^B \sum_{i=1}^{N_B} c_i^B L_i^B(K)}{\mu_u^B \sum_{k=1}^K \sum_{j=1}^{N_R} \sum_{i=1}^{N_B} u_{ij}^B(k)}, \quad (6)$$

其中: J^B, J^R 为蓝、红两方的作战总体收益; c_i^B, c_j^R 分别代表蓝方 i 机和红方 j 机相对于蓝方的价值参数; c_i^R, c_j^B 分别代表蓝方 i 机和红方 j 机相对于红方的价值参数; $\mu_L^B, \mu_L^R, \mu_u^B, \mu_u^R$ 分别为蓝方、红方生存概率,导弹消耗的规范化系数,采用式(7)计算:

$$\mu_p^q = \frac{p^q - \min p^q}{\max p^q - \min p^q}, \quad (7)$$

其中: $p = \{\sum_{i,j,k} L, \sum_{i,j,k} u\}, q = \{B, R\}$ 。

根据作战双方对武器使用情况的控制, 确定模型约束条件为

$$\sum_{k=1}^K \sum_{i=1}^{N_B} u_{ij}^B(k) \leq W_j^B, \forall j = 1, 2, \dots, N_R, \quad (8)$$

$$\sum_{k=1}^K \sum_{j=1}^{N_R} u_{ji}^R(k) \leq W_i^R, \forall i = 1, 2, \dots, N_B, \quad (9)$$

$$\sum_{j=1}^{N_R} u_{ij}^B(k) \leq w_i^B(k), \forall i = 1, 2, \dots, N_B, \quad (10)$$

$$\sum_{i=1}^{N_B} u_{ji}^R(k) \leq w_j^R(k), \forall j = 1, 2, \dots, N_R, \quad (11)$$

式中: $u_{ij}^B(k) \geq 0, u_{ji}^R(k) \geq 0, k = 1, 2, \dots, K$.

式(8)和式(9)描述了空战中蓝/红方最多分配给红方 j 机/蓝方 i 机的武器数目约束, 分别用 W_j^B 和 W_i^R 表示; 式(10)–(11)描述了第 k 个战斗步中蓝方 i 机/红方 j 机最多发射的武器数目约束, 分别以 $w_i^B(k)$ 和 $w_j^R(k)$ 表示.

从模型的建立方法可以看出, 作战双方策略的选取取决于战斗步数 K 和作战双方对武器使用情况的控制(约束条件(8)–(11))等. 而战斗资源有限, 这些条件均满足有限性, 双方的空战策略集为有限集合, 满足文献[16]纳什定理1阐明的纳什均衡解存在性条件—每一个有限次博弈都至少有一个混合策略纳什均衡. 故本文所建立的空战对抗博弈模型纳什均衡点存在, 可应用相应的方法予以求解.

2.2 目标分配问题的混合纳什均衡点求解模型 (Mixed Nash equilibrium point solving model for target assignment problem)

在无人机空战中, 作战双方属于完全敌对状态, 一方的得就是另一方的失, 任何一方都想寻求相对于另一方策略、策略组合的最佳对策, 这非常符合博弈论中的非合作双矩阵对策的概念.

设非合作对策问题记为 $G = (S_I, S_{II}, (A, B))$, 其中局中人 I 和 II 的有限策略集分别为

$$S_I = \{\alpha_1, \alpha_2, \dots, \alpha_n\}, S_{II} = \{\beta_1, \beta_2, \dots, \beta_m\},$$

$A = [a_{ij}]_{n \times m}, B = [b_{ij}]_{n \times m}$ 分别为双方的收益矩阵, 即 a_{ij}, b_{ij} 分别为局中人 I 采取策略 α_i , 局中人 II 采取策略 β_j 时双方的收益值. 若局中人 I 和 II 分别以概率分布 $X = (x_1, x_2, \dots, x_n)$ 和 $Y = (y_1, y_2, \dots, y_m)$ 随机在可选策略中进行选择, 则可得局中人 I 和 II 的混合策略^[16]:

$$S_I^* = \{X = (x_1, x_2, \dots, x_n) \mid \sum_{i=1}^n x_i = 1, x_i \geq 0\},$$

$$S_{II}^* = \{Y = (y_1, y_2, \dots, y_m) \mid \sum_{j=1}^m y_j = 1, y_j \geq 0\}.$$

定义 1(纳什均衡解) 在双矩阵 $G = (S_I, S_{II}, (A, B))$ 中, 若存在一对混合策略组合 (X^*, Y^*) , 对于

任何其他混合策略 $X \in S_I^*, Y \in S_{II}^*$, 均有

$$XAY^{*T} \leq X^*AY^{*T}, X^*BY^T \leq X^*BY^{*T},$$

那么策略组合 (X^*, Y^*) 为 G 的纳什均衡解.

由定义1易知, 给定博弈双方的策略后, 纳什均衡策略组合中各博弈方的策略都是相对于其他博弈方策略组合的最佳对策, 因此任何一方若单独改变策略, 收益上都不会得到好处.

定理 1 一对策略 (X^*, Y^*) 是双矩阵博弈 $G = (S_I, S_{II}, (A, B))$ 的纳什均衡解的充要条件^[17]为

$$A_i \cdot Y^{*T} \leq X^*AY^{*T}, X^*B_{\cdot j} \leq X^*BY^{*T},$$

其中: A_i 为矩阵 A 的第 i 行, $B_{\cdot j}$ 为矩阵 B 的第 j 列, $i = 1, 2, \dots, n, j = 1, 2, \dots, m$.

由定理1, 可建立非合作双矩阵博弈纳什均衡点求解的适应度函数^[18]:

$$f(x, y) = \max\left\{\max_{1 \leq i \leq n} (A_i \cdot Y^T - XAY^T), 0\right\} + \max\left\{\max_{1 \leq j \leq m} (XB_{\cdot j} - XBY^T), 0\right\}. \quad (12)$$

易知, 非合作双矩阵博弈的混合策略组合空间内只有纳什均衡点的适应度最小, 为0, 即使函数 $f(x, y)$ 为最小值0的混合策略 (X^*, Y^*) 即为所求的纳什均衡点, 因此, 这是一个极值已知的优化问题.

2.3 综合模型(Integrated model)

鉴于博弈论思想, 动态目标分配是建立在作战双方掌握双方作战策略但未知敌方的策略选择的情况下的. 可根据约束条件确定作战双方的策略集, 针对每组策略组合由式(5)–(6)计算双方的作战收益, 建立收益矩阵 A, B . 代入式(12)中, 求解问题的纳什均衡点, 即作战双方相对于对方的最佳策略组合. 值得注意的是, 由于这是一个理想极值已知(为0)的最小值优化问题, 所采用的算法的优良性可通过极值与0的误差大小予以衡量.

3 ERPSO算法求解动态目标分配问题的纳什均衡点 (Nash equilibrium point solving based on ERPSO algorithm for dynamic target assignment problem)

快速的收敛性和全局搜索能力是衡量群体智能优化算法性能的两大关键因素. 其中, 粒子群(PSO)算法以其快速的收敛性, 被广泛应用于任务分配及其相关问题中, 算法中粒子的特征用位置、速度和适应度值3项指标表征, 通过跟踪个体极值 P_{best} (个体粒子本身目前所找到的最优解)和群体极值 G_{best} (整个种群目前找到的最优解)更新粒子的速度(移动的方向和距离), 进而更新个体位置. 假设在一个 D 维搜索空间中, 由 n 个粒子组成的种群 $X = (X_1, X_2, \dots, X_n)$, 其中 $X_i = (X_{i1}, X_{i2}, \dots, X_{iD})^T$ 代表第 i 个粒子在 D 维搜

索空间中的位置,也即问题的一个潜在解.根据目标函数可计算出每个粒子位置 X_i 对应的适应度值.设第 i 个粒子的速度为 $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})^T$,其个体极值和种群的群体极值为 $P_{best_i} = (P_{i1}, P_{i2}, \dots, P_{iD})^T$, $G_{best} = (g_{i1}, g_{i2}, \dots, g_{iD})^T$.粒子群算法首先在可行解空间中初始化 n 个粒子,每个粒子通过个体极值和群体极值在进化过程中更新自身的速度和位置^[19],即

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (g_{id}^k - X_{id}^k), \quad (13)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1}, \quad (14)$$

其中: ω 为惯性权重,随进化代数增加从最大权值 ω_{max} 线性递减为最小权值 ω_{min} ,即

$$\omega = \omega_{max} - \frac{k}{k_{max}} \cdot (\omega_{max} - \omega_{min}), \quad (15)$$

其中: $d = 1, 2, \dots, D; i = 1, 2, \dots, n; k$ 为当前迭代次数, $k = 1, 2, \dots, k_{max}; c_1, c_2$ 为加速度因子,是非负常数; r_1, r_2 是分布于 $[0, 1]$ 区间的随机数.

从PSO算法的迭代公式可以看出,所有粒子的位置和速度更新都由 G_{best} 所引导,因而使算法具有快速收敛性能.然而,也因此产生问题—若 G_{best} 陷入局部最优将导致整个种群出现“早熟”现象,陷入局部极小,无法找到问题的全局最优解.为改善PSO算法的性能,引入“精英改选”机制,提出ERPSO(PSO based on elite re-election)算法,在 G_{best} 引导能力不足时,引入若干有力竞争者,增加种群的多样性,提高算法全局搜索能力.算法的流程如图1.

下面对精英改选条件做以详细说明:设 $G_{best}(k)$, $P_{best_i}(k)$ 为种群进化到第 k 代的群体极值和第 i 个个体极值,则根据算法原理有:

$$\sum_{i=1}^{N_{pop}} \delta_{P_{best_i}}(k) = \sum_{i=1}^{N_{pop}} f(P_{best_i}(k)) - \sum_{i=1}^{N_{pop}} f(P_{best_i}(k-1)) \leq 0, \quad (16)$$

$$\delta_{G_{best}}(k) = f(G_{best}(k)) - f(G_{best}(k-1)) \leq 0, \quad (17)$$

其中: $\sum_{i=1}^{N_{pop}} \delta_{P_{best_i}}(k)$, $\delta_{G_{best}}(k)$ 分别表征在当前种群极值的引导下个体极值联合优化和群体极值优化程度,其幅值越大,说明 $G_{best}(k)$ 的引导作用越强,问题解朝着好的方向发展;反之 $G_{best}(k)$ 的引导作用减弱,问题或接近全局最优或陷入局部最优点.因此需引入有竞争力的新精英个体,提高引导作用,以增加个体多样性,改善算法的“早熟”现象.具体方法:设定个体误差阈值error,若 $|\sum_{i=1}^{N_{pop}} \delta_{P_{best_i}}(k)| < N_{pop} \cdot error$,且 $|\delta_{G_{best}}(k)| < error$,则判别精英改选条件成立.精英改选策略的流程如图2所示.

精英改选策略的流程如图2所示.

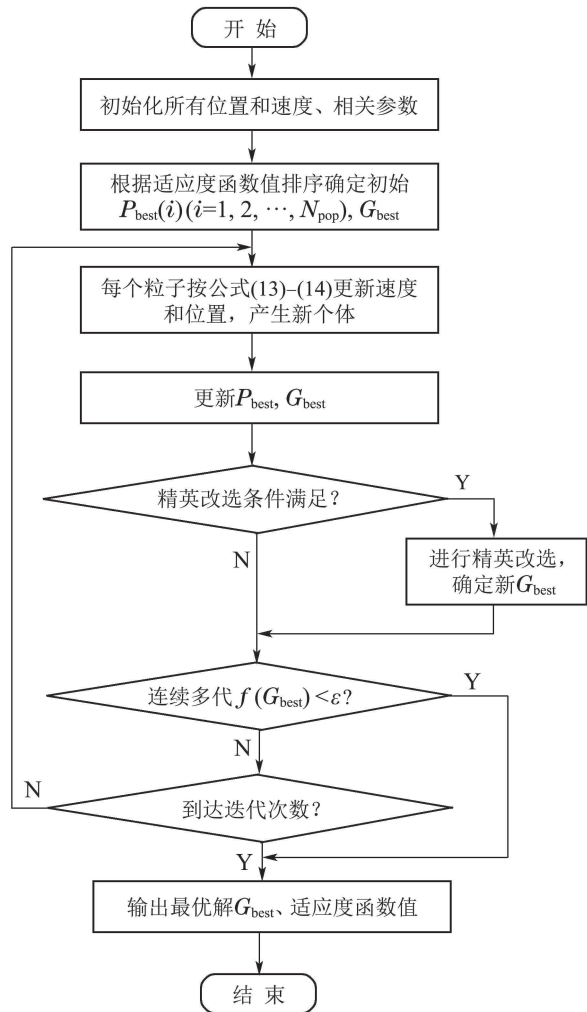


图1 ERPSO算法总流程图

Fig. 1 Flow chart of ERPSO algorithm

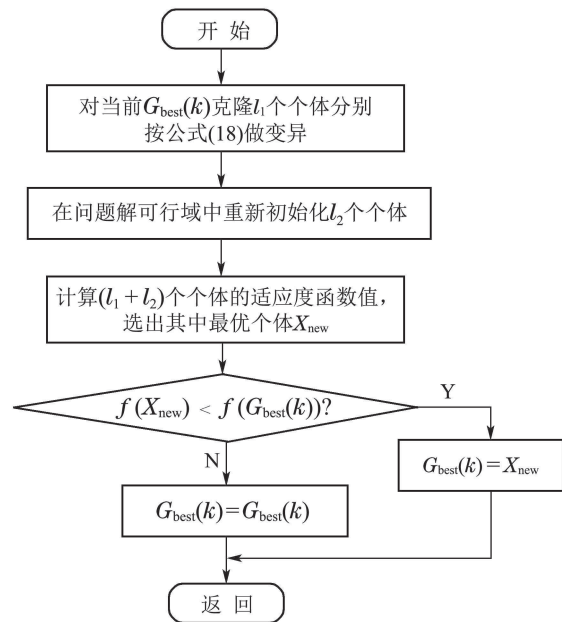


图2 精英改选策略流程图

Fig. 2 Flow chart of elite re-election strategy

这里,采用文献[20]的多项式变异策略:克隆后的 l_1 个 $G_{\text{best}}(k)$ 作为父代个体 p ,对每个 p ,按照以下方式生成子代个体 c :

$$c_k = p_k + (x_k^u - x_k^l)\delta_k, \quad (18)$$

式中: c_k, p_k 为 c, p 的第 k 个分量; x_k^u 为决策变量第 k 个分量的上限,这里为1; x_k^l 为决策变量第 k 个分量的下限,这里为0. 计算 δ_k 的多项式方程为

$$\delta_k = (2r_k)^{\frac{1}{\eta_m+1}} - 1, \quad 0 \leq r_k < 0.5, \quad (19)$$

$$\delta_k = 1 - [2(1 - r_k)]^{\frac{1}{\eta_m+1}}, \quad 0.5 \leq r_k \leq 1, \quad (20)$$

式中: r_k 为均匀分布于区间 $[0, 1]$ 上的随机数; η_m 为变异参数,控制变异程度.

下面应用本文提出的ERPSO算法对基于非合作博弈纳什均衡理论的无人机空战目标分配优化步骤作以说明:

第1步 确定武器约束、空战步数、毁伤概率门限等条件,对空战策略进行筛选,确定双方空战策略(设蓝方 n 条,红方 m 条).

第2步 按照式(5)–(6)对每组策略组合计算双方收益,形成蓝、红两方的收益矩阵 $A_{n \times m}, B_{n \times m}$.

第3步 以作战双方混合策略概率组合为决策变量,联合各分量速度,构造粒子.在约束范围内随机生成规模为 N_{pop} 的初始种群Pop;令初始进化的代数 $\text{gen} = 1$,确定终止迭代次数 gen_stop .每个粒子 X_i 的编码格式为

$$[x_1, \dots, x_n, y_1, \dots, y_m, v_1^x, \dots, v_n^x, v_1^y, \dots, v_m^y],$$

其中: x_1, \dots, x_n 为蓝方混合策略,满足 $\sum_{i=1}^n x_i = 1, \forall x_i \geq 0; y_1, \dots, y_m$ 为红方混合策略,满足 $\sum_{j=1}^m y_j = 1, \forall y_j \geq 0; v_1^x, \dots, v_n^x, v_1^y, \dots, v_m^y$ 为 X, Y 各分量的速度.

第4步 根据每个粒子适应度函数值,找到粒子的初始个体极值 $P_{\text{best}_i}(1)(i = 1, 2, \dots, N_{\text{pop}})$,群体极值 $G_{\text{best}}(1)$.

第5步 按照式(15)计算惯性权重,按式(13)–(14)更新粒子的速度和位置,确定新的个体极值 $P_{\text{best}_i}(k), G_{\text{best}}(k), k$ 为进化代数 $k = \text{gen}(\text{gen} = \text{gen} + 1)$.

第6步 按照式(16)–(17)来计算 $\sum_{i=1}^{N_{\text{pop}}} \delta_{P_{\text{best}_i}}(k)$ 和 $\delta_{G_{\text{best}}}(k)$.根据设定的误差限 error ,判别是否满足精英改选条件,若满足,转第9步,不满足继续.

第7步 判别终止条件是否满足,若满足继续,不满足转第5步.

终止条件为若满足 $G_{\text{best}}(k)$ 的适应度函数值连续多代(仿真实例中取5)小于指定的误差限 ε (仿真实例中取0.003),若满足则直接跳出迭代,否则考察迭代次数是否满足.

第8步 输出适应度函数值 f 和最优个体 G_{best} .

第9步 按流程图2进入精英改选模块,输出改选后的 $G_{\text{best}}(k)$,转第7步.

4 仿真算例(Simulation example)

简单起见,不失一般性,本文作以下场景假设:

设 $N_B = N_R = 2, K = 2, W_j^B = W_i^R = 2, w_i^B(k) = w_j^R(k) = 1(i, j, k = 1, 2)$;这里假定双方相对于对方的价值参数均相等,即

$$c_i^B = c_j^B = c_i^R = c_j^R = 1;$$

武器理想毁伤概率为

$$p_{11}^B = 0.5, p_{12}^B = 0.2, p_{21}^B = 0.3, p_{22}^B = 0.9,$$

$$p_{11}^R = 0.9, p_{12}^R = 0.7, p_{21}^R = 0.55, p_{22}^R = 0.75,$$

$$\beta_w = 1.$$

若双方毁伤概率门限均为0.9(对生存率 < 0.1 的敌机不再分配武器).进行双方策略分析:按照上述假设,在两回合作战中,每回合每架战机最多发射1枚武器,整个空战每架战机最多被分配2枚武器,为战斗充分,假设第1回合双方均充分发射武器,第2回合考虑降低消耗,可减少武器发射量;且注意到毁伤概率 $p_{11}^R = 0.9, p_{22}^B = 0.9$ 均已达到毁伤概率门限,所以制定策略时若第1步已安排红1攻打蓝1,则不需再攻打蓝1;若第1步已安排蓝2攻打红2,则不需再攻打红2,如此策略数得以缩减.由此,双方各可采取11条策略,按式(5)–(6)计算可得到双方价值收益矩阵 A 和 B ,由于 $A = -B$,此处只列出蓝方收益矩阵 A ,如表1所示.表1中第1行、第1列分别为红、蓝双方在两个战斗步中的策略分配,所列数字为攻击敌方机号,表中其余参数为双方采取相应策略时蓝方的收益.例如 $\langle R : \{1, 2; 2, 0\}; B : \{2, 1; 0, 2\} \rangle$ 表示空战中红方采取策略为第1步红1攻击蓝1,红2攻击蓝2,第2步红1攻击蓝2;蓝方采取策略第1步蓝1攻击红2,蓝2攻击红1,第2步蓝2攻击红2. -0.7283 为双方采取此策略组合蓝方收益.

针对收益矩阵 A, B 分别采用GA, PSO以及本文提出的ERPSO算法进行仿真.为更有说服力,比较在同一组初始值的情况下进行,且对每种算法各运行3次,取最好的一次结果,3种算法的运行结果如图3所示.算法参数如下:

初始种群个数100;终止迭代次数100;算法加速度因子 $c_1 = c_2 = 1.3$;精英改选策略中克隆变异的个体数 $l_1 = 50$,重新初始化个体数 $l_2 = 50$;变异算子 $\eta_m = 5$ (GA中交叉、变异算子均取20);精英改选条件误差阈值 $\text{error} = 0.05$;算法提前终止精度 $\varepsilon = 0.003$,连续满足精度代数 $\text{count_end} = 5$;算法权值范围 $\omega_{\text{max}} = 0.9, \omega_{\text{min}} = 0.35$.

表1 蓝方博弈收益
Table 1 Game income of blue side

蓝方策略	红方策略										
	{1, 2; 2, 0}	{2, 1; 0, 2}	{2, 2; 1, 1}	{1, 2; 0, 2}	{2, 1; 2, 1}	{2, 2; 0, 1}	{2, 2; 1, 0}	{2, 1; 0, 1}	{2, 1; 1, 2}	{2, 1; 1, 0}	{2, 1; 2, 0}
{2, 1; 0, 2}	-0.7283	-0.4760	-0.5349	-0.7467	-0.3720	-0.5407	-0.6673	-0.4880	-0.4510	-0.5450	-0.4540
{1, 1; 2, 2}	-0.5953	-0.4137	-0.3564	-0.6584	-0.2926	-0.4682	-0.3507	-0.4250	-0.3155	-0.3721	-0.3380
{1, 2; 1, 0}	-0.2083	0.1600	0.2215	-0.1625	0.1697	0.4467	0.1833	0.1585	0.1063	0.0400	0.1050
{2, 1; 2, 1}	-0.8406	-0.6250	-0.5210	-0.8543	-0.4800	-0.5653	-0.6603	-0.6340	-0.5393	-0.6768	-0.6085
{1, 1; 2, 0}	-0.6942	-0.4900	-0.3195	-0.7783	-0.3440	-0.4167	-0.2600	-0.5050	-0.3745	-0.4345	-0.3890
{2, 1; 0, 1}	-0.8133	-0.5780	-0.5540	-0.8317	-0.4485	-0.5662	-0.6928	-0.5900	-0.5275	-0.6470	-0.5560
{2, 1; 1, 2}	-0.7429	-0.4660	-0.4069	-0.7567	-0.3608	-0.4132	-0.5082	-0.4750	-0.4200	-0.5177	-0.4495
{1, 2; 0, 1}	-0.2000	0.1150	0.1021	-0.1542	0.1360	0.2875	0.0242	0.1135	0.0725	-0.0050	0.0600
{1, 1; 0, 2}	-0.5575	-0.3700	-0.3857	-0.6417	-0.2540	-0.5050	-0.3483	-0.3850	-0.2845	-0.3145	-0.2690
{2, 1; 1, 0}	-0.8250	-0.5150	-0.3869	-0.8433	-0.4013	-0.3433	-0.4700	-0.5270	-0.4803	-0.5840	-0.4930
{2, 1; 2, 0}	-0.8377	-0.5720	-0.4819	-0.8560	-0.4440	-0.4700	-0.5967	-0.5840	-0.5230	-0.6410	-0.5500

由图3可见, GA曲线在100代处仍有下降趋势, 并未收敛, 快速性不佳; PSO曲线虽在100代内收敛, 但准确性有欠缺; 而ERPSO算法于30代内提前收敛于终止精度, 快速性和准确性均得到保证。

3种算法得到的解值(混合策略)及适应度函数值比较见表2。

由表2可以看出, 蓝方更加地倾向于作战策略3{1, 2; 1, 0}和8{1, 2; 0, 1}, 红方更倾向于作战策略1{1, 2; 2, 0}和4{1, 2; 0, 2}, 武器数均为3, 这验证了模型考虑导弹消耗的效果。同时由3种算法中策略的概率分布显见, ERPSO算法策略结果倾向性明显更加明确。

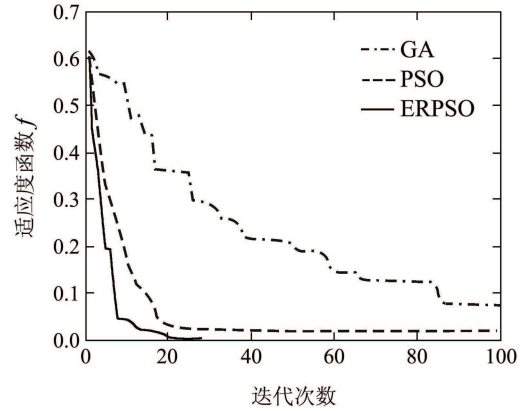


图3 GA, PSO, ERPSO运行结果比较

Fig. 3 Operation results' comparison of GA, PSO and ERPSO

表2 3种算法得到的纳什均衡近似解比较

Table 2 Comparison of approximate Nash equilibrium solution obtained by three algorithms

比较项目		可选策略序号										
		1	2	3	4	5	6	7	8	9	10	11
蓝方混合策略X*	GA	0.0026	0.0005	0.9522	0	0	0	0	0.0188	0.0259	0	0
	PSO	0	0	0.4981	0	0	0	0	0.5019	0	0	0
	ERPSO	0	0	0.2448	0	0	0	0	0.7552	0	0	0
红方混合策略Y*	GA	0.6623	0.0003	0.0030	0.1784	0.0391	0	0.0600	0	0	0.0568	0
	PSO	0.7225	0	0	0.2775	0	0	0	0	0	0	0
	ERPSO	1	0	0	0	0	0	0	0	0	0	0
适应度函数f值(GA/ PSO/ ERPSO)		0.07240 / 0.01687 / 0.002040										

上述仿真给出了红、蓝两方纳什均衡意义下的策略组合。在双方正确评估对方作战策略的情况下, 获取的策略(概率)组合对双方来说均是最佳策略, 任何一方若单独改变自身的策略组合, 都可能降低战斗价值。为验证此结论, 下面对ERPSO仿真结果

进行进一步研究。

1) 纳什均衡条件判别。

根据纳什均衡解的定义, 采用表2中的混合策略(X*, Y*), 计算X*AY*^T = -0.2020, X*BY*^T = 0.2020。随机生成100组混合策略组合([X]_{100×11},

$[Y]_{100 \times 11}$), 分别求取蓝方单独改变策略组合时随机收益 $[XAY^{*T}]_{100 \times 1}$, 红方单独改变策略组合时随机收益 $[X^*BY^T]_{100 \times 1}$, 所得结果如图4所示。

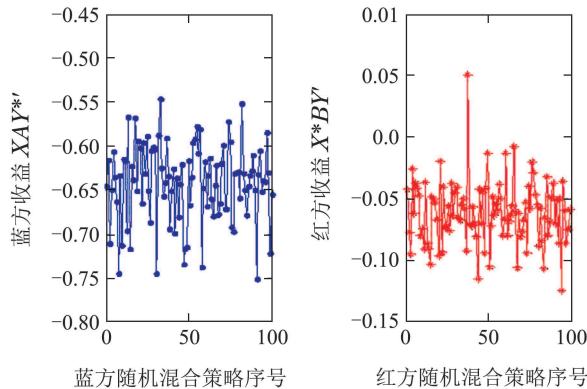


图 4 单方改变混合策略时的100组随机收益比较
Fig. 4 100 groups of random income comparison when mixed strategy is unilaterally changed

显而易见, 蓝方单独改变策略组合而红方保持纳什均衡策略时的蓝方获取收益均小于 $-0.2020 (X^*AY^{*T})$; 与此相似, 红方单独改变策略组合而蓝方保持纳什均衡策略时红方获取收益均小于 $0.2020 (X^*BY^T)$, 验证了纳什均衡条件

$$XAY^{*T} \leq X^*AY^{*T}, X^*BY^T \leq X^*BY^{*T}.$$

同时, 也说明了纳什均衡双方单方改变策略无法提高收益。

2) 近似纯策略比较。

由表2, 蓝方策略8概率为0.7552(接近为1), 红方策略1概率为1. 因此蓝、红两方可近似采用纯策略组合(蓝8VS红1), 将此策略组合与其它策略组合合作收益比较(蓝1, 2, ..., 11VS红1)、(蓝8VS红1, 2, ..., 11), 如图5所示。

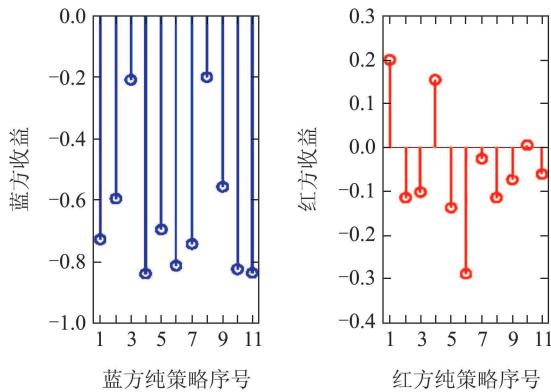


图 5 单方改变纯策略时收益比较
Fig. 5 Income comparison when pure strategy is unilaterally changed

如图5所示, 当红方采用纳什纯策略1时, 蓝方采取纯策略8收益最大; 而当蓝方采用纯策略8时, 红

方采取策略1收益最大, 从而验证了最佳策略组合(蓝8VS红1), 同样说明了单方改变策略并不能得到好处的结论。

鉴于本文仿真实例相对简单, 得到的纳什均衡策略指向性明确. 若复杂化模型参数, 可能求得多重纳什均衡解, 针对此种情况结合文献[16]等给出的多重纳什均衡解的选择方法本文提出如下思路: 将策略风险较低作为首要条件而策略收益较大作为次要条件。

设共解得 L 组蓝、红双方纳什均衡解 $(X_1^*, Y_1^*), (X_2^*, Y_2^*), \dots, (X_l^*, Y_l^*), \dots, (X_L^*, Y_L^*)$, 针对具体情况设定双方收益期望下限 α_A, α_B (低于此收益视为风险), 随机生成 n 组策略概率分布 $X_1, X_2, \dots, X_i, \dots, X_n, Y_1, Y_2, \dots, Y_j, \dots, Y_n$ (理论上 n 越大越好), 针对 L 组纳什均衡解, 分别计算蓝(红)方坚持己方纳什均衡混合策略而红(蓝)方随机改变混合策略时, 双方的收益 $X_l^*AY_j, X_iBY_l^* (l = 1, 2, \dots, L, i, j = 1, 2, \dots, n)$, 最优策略选择分两步进行: 1) 首先笔者统计满足 $X_l^*AY_j > \alpha_A$ 的 Y_j 个数 λ_A 和满足 $X_iBY_l^* > \alpha_B$ 的 X_i 个数 λ_B , 以该个数与 n 的比例超过一定阈值(设为 v , 根据具体情况而定), 即 $\lambda_A/n > v, \lambda_B/n > v$ 的 (X_l^*, Y_l^*) 作为备选策略, 即倾向双方策略风险率均较小的混合策略组合; 2) 在此基础上取策略收益和最大者的纳什均衡策略解为最优策略, 即

$$(X^*, Y^*) = \{(X_l^*, Y_l^*) | \max_l (\sum_{j=1}^n X_l^*AY_j + \sum_{i=1}^n X_iBY_l^*)\}, l = 1, 2, \dots, L. \tag{21}$$

5 算法性能分析(Performance analysis of the algorithm)

为考查本文提出的ERPSO算法的性能, 将传统PSO和ERPSO分别应用于上述实例, 在相同初始个体的情况下各运行100次, 得到比较如表3所示。

表 3 算法性能比较

算法	平均迭代	平均运行/s	\bar{f}	提前收敛	0值
PSO	94	3.8023	0.0171	8	1
ERPSO	78.36	5.0368	0.0061	33	5

由表3可以看出, ERPSO算法大幅降低了目标极值, 使其更接近理想值0, 同时虽然ERPSO算法增加了克隆、变异等操作, 理论上大幅度增加了收敛时间(100代运行时间6.6 s左右), 但ERPSO算法在确保极值误差更小的情况下, 通过迭代次数的减少, 使

算法的平均运行时间减少1.6 s, 在决策时间允许的情况下将提供更为精确的信息。

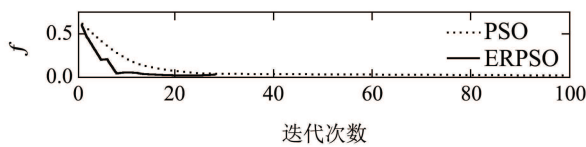
由于本文提出的ERPSO算法是改善传统PSO算法容易陷入局部极小的问题, 在原 G_{best} 显示出引导能力不足时, 改善个体多样性, 扩大搜索范围是本算法的主要手段。因此, 有必要在算法结果比较时加入粒子位置和速度多样性的考查。这里的多样性采用文献[21]提出的方法:

$$\begin{cases} \bar{X}_j = \frac{1}{N_{\text{pop}}} \sum_{i=1}^{N_{\text{pop}}} X_{ij}, \\ D_j^X = \frac{1}{N_{\text{pop}}} \sum_{i=1}^{N_{\text{pop}}} |X_{ij} - \bar{X}_j|, \\ D_X = \frac{1}{n+m} \sum_{j=1}^{n+m} D_j^X, \end{cases} \quad (22)$$

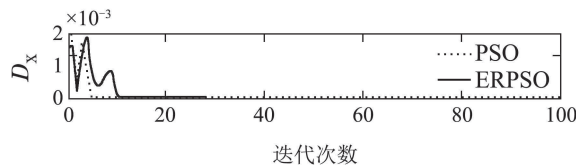
$$\begin{cases} \bar{V}_j = \frac{1}{N_{\text{pop}}} \sum_{i=1}^{N_{\text{pop}}} V_{ij}, \\ D_j^V = \frac{1}{N_{\text{pop}}} \sum_{i=1}^{N_{\text{pop}}} |V_{ij} - \bar{V}_j|, \\ D_V = \frac{1}{n+m} \sum_{j=1}^{n+m} D_j^V, \end{cases} \quad (23)$$

其中: X 代表粒子的位置变量, V 为粒子的速度变量, N_{pop} 为每代粒子的个数, $n+m$ 为位置的维度。

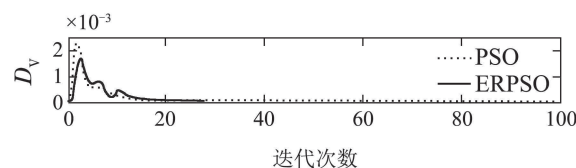
为说明ERPSO算法如何通过增加个体的多样性改善优化效果, 下面通过曲线图6对图3中的结果进行ERPSO算法与传统PSO算法的位置、速度多样性比较。



(a) 适应度函数值



(b) 位置多样性



(c) 速度多样性

图6 PSO与ERPSO算法多样性比较

Fig. 6 Diversity comparison of PSO and ERPSO

由图6的适应度函数曲线可以看出, 传统的PSO

算法曲线平滑, 在20代左右收敛到0.01687, 与期望值0有较大差距, 未达到提前收敛条件, 故需运行100代, 显然, 此次运行结果陷入局部极小; 而本文提出的ERPSO算法曲线在20代内有两处明显转折, 使得曲线急剧下降, 在约23代处进入误差限0.003, 到第28代(经过5代)跳出循环, 输出结果, 既达到满意效果又降低了运行时间。

由多样性曲线比较可以看出, 传统的PSO算法在5代以内显示出较强大的多样性, 但之后几乎没有变化, 这也直接导致适应度函数值陷入局部极小; 而ERPSO算法10代内有较大改变, 且有两处明显转折(对应适应度值的转折), 这归功于10代内的两次 G_{best} 的大范围改选, 引导个体重新调整, 趋向理想结果。

6 结论(Conclusion)

本文将动态目标分配问题(DWTA)与博弈论中的纳什均衡概念相结合, 并考虑生存概率和武器消耗的双重目标, 创建了无人机协同空战的多战斗步DWTA博弈模型, 赋予了目标分配问题的动态对抗性。

为精确寻求模型的纳什均衡点, 获取作战双方的概率意义下的混合策略, 提出了基于精英改选机制的粒子群(ERPSO)算法, 引入竞争机制, 改选精英个体, 增加个体多样性, 引导粒子向更优的方向行进, 在保证求解的快速性的基础上, 提高了全局搜索能力, 改善传统PSO算法易于“早熟”现象。

将提出的ERPSO算法应用于所建立的DWTA博弈模型, 成功获取了纳什均衡意义下的双方最佳策略组合, 且运行时间和解的准确性得以保证, 验证了所提出的模型和算法的正确性和有效性, 对信息化无人机空战决策起到一定的指导作用。

参考文献(References):

- [1] GALATI D G, SIMAAN M A. Effectiveness of the nash strategies in competitive multi-team target assignment problems [J]. *Transactions of Aerospace and Electronic Systems*, 2007, 43(1): 126 - 134.
- [2] ATHANS M. Command and control (C2) theory: a challenge to control science [J]. *IEEE Transactions on Automatic Control*, 1987, 32(4): 286 - 293.
- [3] XIN B, CHEN J, PENG Z, et al. An efficient rule-based constructive heuristic to solve dynamic weapon-target assignment problem [J]. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 2011, 41(3): 598 - 606.
- [4] 王光辉, 陈杰, 蔡涛, 等. 多目标分解随机粒子群优化算法及其在直线电机优化设计中的应用 [J]. *控制理论与应用*, 2013, 30(6): 693 - 701.
- [5] (WANG Guanghui, CHEN Jie, CAI Tao, et al. A multi-objective decomposition-based stochastic particle swarm optimization algorithm and its application to optimal design for linear motor [J]. *Control Theory & Applications*, 2013, 30(6): 693 - 701.)

- [5] SHI Y, EBERHART R C. Fuzzy adaptive particle swarm optimization [C] // *Proceedings of the Congress on Evolutionary Computation*. Honolulu, HI: IEEE, 2001, 1: 101 – 106.
- [6] CLERC M, KENNEDY J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space [J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(1): 58 – 73.
- [7] RATNAWEERA A, HALGAMUGE S K. Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients [J]. *IEEE Transactions on Evolutionary Computation*, 2004, 8(3): 240 – 255.
- [8] MENEDS R, KENNEDY J, NEVES J. The fully informed particle swarm: simpler, maybe better [J]. *IEEE Transactions on Evolutionary Computation*, 2004, 8(3): 204 – 210.
- [9] LIANG J J, QIN A K, SUGANTHAN P N, et al. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions [J]. *IEEE Transactions on Evolutionary Computation*, 2006, 10(3): 281 – 295.
- [10] BRITS R, ENGELBRECHT A P, VAN DEN BERGH F. Locating multiple optima using particle swarm optimization [J]. *Applied Mathematics and Computation*, 2007, 189(2): 1859 – 1883.
- [11] LI J P, WOOD A S. Species-conserving particle swarm optimization for multimodal functions [J]. *International Journal of Modelling, Identification and Control*, 2009, 8(4): 290 – 300.
- [12] ZHAN Z H, ZHANG J, LI Y, et al. Adaptive particle swarm optimization [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2009, 39(6): 1362 – 1381.
- [13] 金敏, 鲁华祥. 一种遗传算法与粒子群优化的多子群分层混合算法 [J]. *控制理论与应用*, 2013, 30(10): 1231 – 1238. (JIN Min, LU Huaxiang. A multi-subgroup hierarchical hybrid of genetic algorithm and particle swarm optimization [J]. *Control Theory & Applications*, 2013, 30(10): 1231 – 1238.)
- [14] CHEN W N, ZHANG J, LIN Y, et al. Particle swarm optimization with an aging leader and challengers [J]. *IEEE Transactions on Evolutionary Computation*, 2013, 17(2): 241 – 258.
- [15] 马飞, 曹泽阳, 刘晖. 基于博弈论的目标分配策略空间构建与搜索 [J]. *系统工程与电子技术*, 2010, 32(9): 1941 – 1945. (MA Fei, CAO Zeyang, LIU Hui. Construction and search of strategy space of target assignment based on game theory [J]. *Systems Engineering and Electronic*, 2010, 32(9): 1941 – 1945.)
- [16] 范如国. 博弈论 [M]. 武汉: 武汉大学出版社, 2011. (FAN Ruguo. *Game Theory* [M]. Wuhan: Wuhan University Press, 2011.)
- [17] 曾松林, 王文辉, 丁大春, 等. 基于动态博弈的目标分配方法研究 [J]. *光电与控制*, 2011, 18(2): 26 – 29, 72. (ZENG Songlin, WANG Wenyun, DING Dachun, et al. Target allocation method based on dynamic game [J]. *Electronics Optics & Control*, 2011, 18(2): 26 – 29, 72.)
- [18] 贾文生, 向淑文, 杨剑锋, 等. 基于免疫粒子群算法的非合作博弈 Nash均衡问题求解 [J]. *计算机应用研究*, 2012, 29(1): 28 – 31. (JIA Wensheng, XIANG Shuwen, YANG Jianfeng, et al. Solving Nash equilibrium for N-persons' non-cooperative game based on immune particle swarm algorithm [J]. *Application Research of Computers*, 2012, 29(1): 28 – 31.)
- [19] 史峰, 王辉, 郁磊, 等. MATLAB智能算法30个案例分析 [M]. 北京: 北京航空航天大学出版社, 2010. (SHI Feng, WANG Hui, YU Lei, et al. *Cases Analysis of MATLAB Intelligent Algorithm* [M]. Beijing: Beihang University Press, 2010.)
- [20] DEB K. *Multi-objective Optimization Using Evolutionary Algorithms* [M]. New Jersey, America: John Wiley & Sons, Inc, 2001.
- [21] 程适, 史玉回. 基于 L_1 范式的粒子群算法群体多样性研究 [J]. *计算机科学*, 2011, 38(7): 190 – 193. (CHENG Shi, SHI Yuhui. Measurement of PSO diversity based on L_1 norm [J]. *Computer Science*, 2011, 38(7): 190 – 193.)

作者简介:

王昱 (1980–), 女, 讲师, 博士研究生, 目前研究方向为无人机空战态势估计、目标分配, E-mail: yuki1120@sina.com;

章卫国 (1956–), 男, 教授, 博士生导师, 目前研究方向为现代控制方法与智能控制方法及应用、先进与智能飞行控制及优化技术、容错控制方法及应用、基于无人机平台的自主控制与飞行验证技术, E-mail: zhangwg@nwpu.edu.cn;

傅莉 (1968–), 女, 教授, 硕士生导师, 目前研究方向为混杂对策建模方法、智能控制技术以及磁流变阻尼器应用技术, E-mail: ffulli@163.com;

黄得刚 (1986–), 男, 博士研究生, 目前研究方向为飞行器路径规划导航、制导与控制, E-mail: hdg0216@163.com;

李勇 (1980–), 男, 博士, 目前研究方向为多目标优化算法、轧制流程优化与控制, E-mail: liyong.neu@gmail.com.