

文章编号: 1006-1576 (2009) 01-0004-04

无人机攻防对抗不完全信息动态博弈方法研究

惠一楠, 朱华勇, 沈林成

(国防科学技术大学 3 院, 湖南 长沙 410073)

摘要: 以不完全信息动态博弈理论为基础, 无人机攻防对抗问题为背景, 建立无人机攻防对抗问题的不完全信息动态博弈模型; 针对博弈树的求解问题, 设计免疫进化算法, 求解问题的贝叶斯纳什均衡解, 得到无人机的最优策略序列。仿真结果表明该模型能够合理的应用于无人机攻防对抗问题当中, 具有较好的应用价值。

关键词: 攻防对抗; 不完全信息动态博弈; 免疫进化算法; 贝叶斯纳什均衡

中图分类号: V279; O225 **文献标识码:** A

Study on Dynamic Game Method with Incomplete Information in UAV Attack-Defends Campaign

HUI Yi-nan, ZHU Hua-yong, SHEN Lin-cheng

(No. 3 College, National University of Defense Technology, Changsha 410073, China)

Abstract: Based on dynamic game with incomplete information theory, and took UAV attack-defends campaign as background, the dynamic game with incomplete information model of UAV attack-defends campaign was established. According to the problem of game tree solution, the immune evolutionary algorithm was designed, the Bayesian Nash equilibrium was carried out, and the best strategy sequence was achieved. The simulating result shows that it can be applied in the model of the UAV attack-defends campaign appropriately and has a nice value in application.

Keywords: Attack-defends campaign; Dynamic game with incomplete information; Immune evolutionary algorithm; Bayesian Nash equilibrium

0 引言

现代高技术局部战争, 无人机以其高隐身、高机动、无人化等优越性, 成为未来夺取制空权的重要法宝。由于战场环境的高复杂性和瞬时性, 导致敌我双方所获得的信息不完全。同时, 无人机在执行作战任务中, 还应能够根据不断获取的最新信息, 及时更新和调整自己的策略, 这是一个动态博弈的过程。目前, 应用博弈的思想研究攻防对抗, 国内外已经有了相应的研究成果^[1-3], 但是这些研究都是基于完全信息的静态博弈理论, 与实际问题的相差较大, 应用价值不大。故对无人机攻防对抗的不完全信息动态博弈方法进行研究。

1 攻防对抗不完全信息动态博弈建模

1.1 无人机攻防对抗问题说明

将敌我双方看作一个整体作为博弈的 2 个参与者, 由于敌我双方所能采取的对抗策略和对抗过程有限, 基于动态规划的思想, 假定时间和空间都可离散化, 以一定的单位时间为基准, 将对抗博弈阶段划分为若干个时域离散的子过程 K , 如图 1。

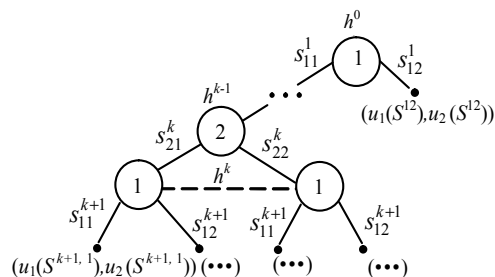


图 1 无人机进入作战区域后攻防对抗想定博弈树

在第 1 阶段, 首先由无人机选择策略 s_{11}^1 或者 s_{12}^1 , 选择 s_{12}^1 博弈结束, 敌我双方收益分别为 $((u_1(s_{12}^1), u_2(s_{12}^1)))$; 选择 s_{11}^1 , 继续博弈; 到第 k 阶段, 此时敌方做出决策 s_{21}^k 或者 s_{22}^k ; 到第 $k+1$ 阶段我方在不知道敌方策略的情况下做出决策 s_{11}^{k+1} 或者 s_{12}^{k+1} , 此时, 敌我双方收益分别为 $(u_1(s_{11}^{k+1}, 1), u_2(s_{11}^{k+1}, 1))$, $u_i(s^k)$ 表示参与者 i 对当前战场态势的评估。

1.2 不完全信息动态博弈模型标准式表述

基于上述分析, 应用博弈论的相关理论^[4], 可得到 SEAD 任务的不完全信息动态博弈模型的标准式表述为:

收稿日期: 2008-08-28; 修回日期: 2008-09-10

作者简介: 惠一楠 (1984-), 男, 陕西人, 国防科学技术大学在读硕士研究生, 从事无人机任务管理技术、智能控制、任务规划研究。

$$G = \langle N, H, F, P, S^k, (\succ_i) \rangle$$

其中:

1) $N=\{1,2\}$, 为博弈的 2 个参与者, 1 代表我方, 2 代表敌方。

2) $H^k=\{s^1, s^2, \dots, s^{k-1}\}$ 代表历史集, 是到达第 k 阶段时参与者策略序列的集合 (敌我双方的对抗策略肯定是有限的), 满足以下 2 个条件:

(1) 空序列 $\phi \in H$;

(2) 如果有限长参与者策略序列 $(s^k)_{k=1}^L \in H$, 且 $L < K$, 则 $(s^k)_{k=1}^L \in H$;

用 Z 表示终极历史 (1 条从起始节点到终极节点的策略序列)。

其中, $s^k = s_{ij}^k$, 表示在第 k 个对抗阶段, 参与者 i 选择在第 k 阶段的 j 策略。 $\{s_{ij}^k\}$ 表示参与者 i 选择在第 k 阶段所能选择的策略集合。

3) F 代表参与者函数, $F(H)$ 表示紧接 H 之后采取策略的参与者 (敌我双方都可能连续采取策略)。

4) P 代表概率函数, $P(s^k|H)$ 表示在历史 H 之后, 选择策略 s^k 的概率。其中, $s^k \in \{s_{ij}^k\}$, $\{s_{ij}^k\}$ 表示 H 之后的可行策略集。

5) S^k 表示第 k 阶段的战场态势。 $S^k = S^{k-1} \times s^k$, 表示第 k 阶段的战场状态由第 $k-1$ 阶段的战场状态执行策略 s^k 后得到, $s^k = s_{ij}^k$ 。

6) \succ_i 代表参与者 i 对于 Z 上不确定事件的偏好关系, 用收益函数 $u_i(S^k)$ 表示, $u_i(S^k)$ 为参与者 i 对当前战场效能的评估。即若收益函数 $u_i(S^{k1}) \geq u_i(S^{k2}) \Leftrightarrow$ 战场态势 $S^{k1} \succ S^{k2}$, 当 $S^{k1} = S^{k-1} \times s_{i1}^k$, $S^{k2} = S^{k-1} \times s_{i2}^k$ 时, 即有策略 $S_{i1}^k \succ S_{i2}^k$ 。

1.3 最优策略序列即精炼贝叶斯纳什均衡

根据博弈论相关知识, 求解该不完全信息动态博弈问题的精炼贝叶斯纳什均衡, 可转化为求解:

$$\max_{s_i \in S_i} \sum \tilde{P}_i(H | s^k) u_i(s^k)$$

其中: $\tilde{P}_i(H | s^k)$ 为参与者 i 执行策略 s^k 后对历史集 H 的推断, $u_i(s^k)$ 为参与者 i 对战场效能的评估。

由此, 即可求得精炼贝叶斯纳什均衡解为 $s^* = (s^{1*}, s^{2*}, \dots, s^{k*}, \dots, s^{K*})$, 表示在整个对抗过程中的最优策略序列, 即 $\forall s_{ij}^k$:

$$s^* = (s^{1*}, s^{2*}, \dots, s^{k*}, \dots, s^{K*}) \geq (s^{1*}, s^{2*}, \dots, s_{ij}^k, \dots, s^{K*})$$

其中 $s^{k*} \in \arg\{s_{ij}^k\}$, 表示参与者 i 在第 k 阶段的最优策略。

2 模型各部分函数之间相互转换关系

根据第 1 部分的分析, 当执行完第 $k-1$ 阶段, 到第 k 阶段时, 已经执行过的历史集 $H=\{s^1, s^2, \dots, s^{k-1}\}$, 该阶段的执行者为 $i=F(h)$, 表示该阶段轮到参与者 i 执行策略。此时, 参与者 i 选择策略 s^k 的先验概率为 $P_i(s^k|H)$, 执行 s^k 时对历史集 H 的推断为 $\tilde{P}_i(H | s^k)$, 执行策略 s^k 前战场状态为 S^{k-1} , 执行策略 s^k 后预期战场状态为 S^k , $S^k = S^{k-1} \times s^k$ 。

设定战场态势 S^k 由武器状态 $We_i^j(s^k)$ 和平台状态 $Pl_i^j(s^k)$ 组成, 并由此设定此战场态势下的效能评估函数为 $Ev_i^j(s^k) = We_i^j(s^k) + Pl_i^j(s^k)$, $Ev_i^j(s^k)$ 为参与者 i 对参与者 j 的效能评估, $We_i^j(s^k)$ 和 $Pl_i^j(s^k)$ 分别为对武器效能和平台效能的评估。由于作战目的、信息获取等不同, 敌我双方对自身以及对方的效能评估可能不同。

随着时间的推移, 当执行完策略 s^k , 到达第 $k+1$ 步时, 参与者 i 的状态转移^[5,6]为:

$$We_i^j(s^k) = We_i^j(s^{k-1}) + \Delta We_i^j(s^k)$$

$$Pl_i^j(s^k) = Pl_i^j(s^{k-1}) * Tr_i^j(s^k)$$

其中, $\Delta We_i^j(s^k)$ 和 $Tr_i^j(s^k)$ 分别表示执行策略 s^k 时武器效能变化量和平台效能转移系数。

同时, 收益函数 $u_i(s^k)$ 可用战场态势评估信息表示, $u_1(s^k) = Ev_1^1(s^k) - Ev_1^2(s^k)$, $u_2(s^k) = Ev_2^2(s^k) - Ev_2^1(s^k)$, 且一般情况下, $u_1(s^k) = -u_2(s^k)$ 。否则, 必然有一方存在判断失误或者双方对某事件评判存在严重差别, 这在强敌作战时很难出现。

3 应用免疫进化算法求解贝叶斯纳什均衡

尽管 Nash 证明了纳什均衡解的存在性, 但是 Nash 并没有给出 1 个通用的求解 Nash 均衡解的算法。现有的求解纳什均衡解的算法有很多, 例如 Min-Max、 α - β 剪枝、小窗口搜索^[7,8]等算法。这些算法在求解博弈问题时各有优缺点, 但有一个共同之处就是对求解的指数复杂度问题。

通过模拟生物进化现象, 设计仿生算法求解博弈问题是当前研究的方向。进化算法是建立在模仿生物遗传与自然选择基础上的一种非数值并行随机优化算法, 其性能优异、鲁棒性强、应用广泛。但是, 传统进化算法存在早熟、退化和后期局部收敛速度慢等弊端。受生物免疫系统的启发, 在原有进化算法的基础上加入免疫因子, 研究基于免疫系统机理的进化算法—免疫进化算法^[9,10] (Artificial

Immune System, AIS), AIS 集免疫机制和进化机制于一体, 具有并行计算、种群多样性好、局部搜索能力强、收敛速度快等优点。

3.1 编码方案的设计

以当前位置和当前策略作为基因, 采用十进制和二进制混合方式编码, 组成染色体串, 如图 2。

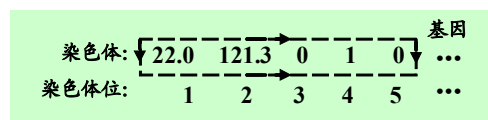


图 2 染色体编码示意图

其中, 以 5 位为 1 个基因, 前 2 位采用十进制编码表示无人机当前位置, 后 3 位采用二进制编码表示无人机当前将要采取的策略 (000—111 可表示 8 种不同策略)。染色体长度为 $5K$, K 为搜索深度。

3.2 适应度函数的计算

对于每个染色体串 i , 分别计算到达各个基因位 j 时的战场态势 $|u_1(s^k)|$, 求得 $|u_1(s^k)|$ 最大值所在的基因位。判断该基因位的参与者 i , 若 $i=1$, $u_1(s^k) > 0$, 说明该染色体整体对我方有利; 若 $i=2$, $u_1(s^k) < 0$, 说明该染色体整体对对方有利。因而, 染色体 i 的适应度函数值可表示为 $Fitness_i = u_1(s^k)$ 。

3.3 种群的多样性控制

为避免早熟现象, 免疫进化算法中引入多样性控制这一操作。采用基于浓度的多样性控制方法, 使得低密度个体生存几率较大, 而高密度个体生存受到抑制, 体现个体之间的相互促进和抑制作用, 达到控制种群多样性的目的。

假设群体规模 M , 每个个体包含 K 个基因, 则第 j 个基因的信息熵可表示为:

$$H_j(M) = \sum_{i=1}^M (-p_{ij} * \log p_{ij})$$

p_{ij} 是第 i 个个体的等位基因源于第 j 个基因的概率。如果在位置 j 上所有个体的等位基因都相同, 则 $H_j(M)=0$ 。从而种群的多样性可用平均信息熵 $H(M)$ 表示为:

$$H(M) = \sum_{j=1}^K H_j(M) / K$$

由此可得个体 i 和个体 j 之间的相似度为 $Similar_{ij} = 1 / [1 + H(2)]$, 个体 i 在当前种群中的浓度 (用与个体 i 相近的个体数量衡量) 为 $C_i = \left(\sum_{j=1}^M Flag_{ij} \right) / M$, 其中:

$$Flag_{ij} = \begin{cases} 1 & Similar_{ij} \geq Threshold \\ 0 & Similar_{ij} < Threshold \end{cases}$$

$Threshold$ 为设定的相似度阈值。

3.4 进化算子的设计

通过编码组成初始种群后, 对种群中的每个个体进行适应度评估, 然后通过选择、交叉和变异等进化算子来实现优胜劣汰的进化过程。其中, 选择算子采用适应度比例加个体浓度的方法, 个体 i 被选择的概率 $SelP_i$ 可表示为:

$$SelP_i = \frac{Fitness_i / \left(C_i * \sum_{i=1}^M Fitness_i \right)}{\sum_{i=1}^M \left[Fitness_i / \left(C_i * \sum_{i=1}^M Fitness_i \right) \right]}$$

这样, 可以在适当鼓励与最优抗体接近个体的同时抑制浓度很高的个体, 充分体现了抑制浓度与鼓励优良个体相结合的算法特点; 交叉算子采用基于基因的单点交叉, 交叉概率 Pc ; 变异算子采用基于基因值的随机变异, 变异概率 Pm 。

3.5 设置结束条件

适应度函数值最大值若干代内不再增加, 迭代停止; 已经达到最大进化代数, 迭代停止; 由此, 得到的适应度函数值最大的染色体串即为问题的贝叶斯纳什均衡解。

4 算例仿真及结果分析

4.1 各单元仿真初始参数设定

设定作战想定为我方 1 架无人机 (高度 10000m, 速度 0.8Ma) 打击敌方 1 个导弹阵地, 各个组件参数设置如表 1。

表 1 各单元初始参数配置表

参与者	类型	数量	单位价值	最大齐射数	击中概率	击中效果
1	UCAV	1	1000	1	0.65	0.5
1	导弹	10	50		0.33	0.99
2	目标	1	5000	2	0.9	0.30
2	导弹	30	50		0.33	0.99

简化作战过程, 设定作战双方所能采取的策略集合如表 2。

表 2 第 k 步时双方策略配置表

	1	2	3
1	发射 $Num_i^j(s^k)$ 导弹打击 导弹阵地	发射 $Num_i^j(s^k)$ 导弹拦截 地空导弹	躲避
2	发射 $Num_i^j(s^k)$ 导弹打击 无人机	发射 $Num_i^j(s^k)$ 导弹拦截 空空导弹	不发射

因此, 若第 k 阶段参与者 i 对参与者 j 执行策

略 s^k 时发射导弹, 则:

$$Tr_i^j(s^k) = Hit_i^j(s^k) * (1 - St_i^j(s^k)) * Num_i^j(s^k)$$

$$\Delta We_i^j(s^k) = -Num_i^j(s^k)$$

其中, $Hit_i^j(s^k)$ 、 $St_i^j(s^k)$ 和 $Num_i^j(s^k)$ 分别为被击中概率、击中后效果、齐射数 (单位时间发射导弹数目)。

若第 k 步时不发射导弹, 则 $Tr_i^j(s^k) = 1$, $\Delta We_i^j(s^k) = 0$ 。

4.2 仿真结果及分析

采用免疫进化算法求解, 得到最优控制序列如表 3。

表 3 若干作战阶段策略选择表

阶段	无人机位置	参与者 1 策略	参与者 2 策略
1	(115.00,27.00)	起飞	无
165	(119.32,24.86)	导弹 1 攻击导弹阵地	无
166	(119.36,24.84)	导弹 2 攻击导弹阵地	导弹 1,2 打击飞机
167	(119.39,24.83)	导弹 3 攻击导弹阵地	导弹 3,4 拦截红方导弹 1
168	(119.42,24.81)	无	导弹 5,6 拦截红方导弹 2
169	(119.46,24.80)	无	导弹 7,8 拦截红方导弹 3
183	(119.93,24.56)	躲避蓝方导弹 1, 2	无
188	(120.09,24.48)	导弹 4 攻击导弹阵地	导弹 9,10 打击飞机
189	(120.13,24.46)	导弹 5 攻击导弹阵地	导弹 11,12 拦截红方导弹 4
190	(120.16,24.44)	无	导弹 13,14 拦截红方导弹 5
202	(120.22,24.33)	躲避蓝方导弹 9, 10	无
207	(120.01,24.33)	无	导弹 15 打击飞机
211	(119.84,24.33)	导弹 6 攻击导弹阵地	无
212	(119.80,24.33)	无	导弹 16 拦截红方导弹 6

根据表 3, 绘制各阶段的简要作战过程示意图如图 3, 并统计双方战果如表 4。

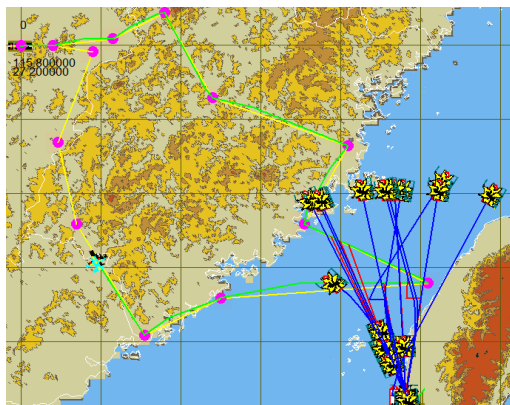


图 3 简要作战过程示意图

从表 4 可看出, 我方在无人机受损 50% 的情况下, 耗费 6 枚导弹, 就能有效地打击敌方导弹阵地 (毁损率 60%), 并且消耗敌方导弹 16 枚。其中, 我方 6 枚导弹中 2 枚命中目标, 4 枚被敌方拦截; 敌方 16 枚导弹中有 11 枚用于拦截我方导弹 (4 枚命中), 5 枚用于打击我方无人机 (1 枚命中, 其余

4 枚被我方有效躲避)。

表 4 作战双方战果统计表

参与者	导弹损失	折合价值	双方受损	折合价值	总损失	总收益
红方	6	300	50%	500	800	3000
蓝方	16	800	60%	3000	3800	-3000

从表 4 可看出, 在此次作战过程中, 我方明显占优, 因为我方作为进攻方, 掌握着作战的主动权。经过对问题的不完全信息动态博弈分析后, 得出结论: 不管敌方采取何种应对策略, 我方都可以制定相应的作战策略并取得不小于 3000 的总收益。

5 结束语

为了建立有效的无人机攻防对抗不完全信息动态博弈模型, 为无人机能够在敌方行动和意图不完全明了的情况下, 通过分析敌方历史信息和当前信息, 得到当前敌方属性的概率分布信息, 并应用免疫进化算法求解贝叶斯纳什均衡解, 得到当前的最优行动策略。仿真结果表明, 模型设计切合实际, 求解算法切实可行, 对无人机的自主化设计有一定的促进作用, 具有良好的应用前景。

参考文献:

- [1] Thomas Hamilton, Richard Mesic. A Simple Game-Theoretic Approach to Suppression of Enemy Defenses and Other Time Critical Target Analyses[R]. Santa Monica: Published by RAND Corporation, 2004.
- [2] Rene Vidal, Omid Shakernia, H. Jin Kim, etc. Probabilistic Pursuit-Evasion Games: Theory, Implementation and Experimental Evaluation[J]. IEEE Transactions On Robotics And Automation. 2002, 5(31): 100-107.
- [3] 姚宗信, 李明, 陈宗基. 基于博弈论模型的多级协同对抗多目标任务决策方法[J]. 航空计算技术, 2007, 3(37): 7-10.
- [4] 侯定丕. 博弈论导论[M]. 合肥: 中国科学技术大学出版社, 2004.
- [5] 张莉, 张安, 汤志荔, 等. 多编队对地攻防对抗多层决策分析仿真与研究[J]. 系统仿真学报, 2007, 1(19): 106-109.
- [6] 张莉, 张安, 张永芳, 等. 编队协同对地攻防对抗多级动态决策方法研究[J]. 系统仿真学报, 2006, 2(18): 794-796.
- [7] LUGER G F. 人工智能—复杂问题求解的结构和策略 [M]. 史忠植, 译. 北京: 机械工业出版社, 2004.
- [8] 李红, 吴粉侠, 刘小豫. 博弈树搜索算法研究[J]. 长春工程学院学报 (自然科学版), 2007, 8(2): 59-62.
- [9] 陈士俊, 孙永广, 吴宗鑫. 一种求解纳什均衡解的遗传算法[J]. 系统工程, 2001, 19(5): 67-70.
- [10] 王磊, 潘进, 焦李成. 免疫算法[J]. 电子学报, 2000, 28(7): 74-78.