

基于影响网络与不完全信息多阶段博弈的 作战行动序列模型及求解方法*

杜正军, 陈超, 姜鑫

(国防科技大学 信息系统工程重点实验室, 湖南 长沙 410073)

摘要: 拟制作战计划的关键是生成作战行动序列。考虑作战行动过程中的不确定性以及激烈对抗性, 以影响网络分析作战行动和战场态势之间的相互影响, 用不完全信息博弈分析敌对双方之间的对抗, 建立了基于影响网络和不完全信息多阶段博弈的作战行动序列模型, 并给出了求解方法。通过计算实例说明了模型及其求解方法。结果显示该模型生成的 COA 更适应于对抗环境。

关键词: 作战行动序列; 影响网络; 不完全信息多阶段博弈

中图分类号: **文献标志码:**A **文章编号:**1001-2486(2012)03-0063-05

Modeling and solution method of course of action based on influence net and multi-stage games with incomplete information

DU Zhengjun, CHEN Chao, JIANG Xin

(Science and Technology on Information Systems Engineering Laboratory,
National University of Defense Technology, Changsha 410073, China)

Abstract: The development of Course of Action (COA) is a key step of military planning. In most existing literature on COA development, the previous studies only take unilateral actions of friendly force into account. Considering the antagonism of war, we propose models that contain actions of both friendly and enemy force based on multi-stage games with incomplete information. Influence net is used to analyze the uncertain relationships between actions and battlefield situation. In the end, numerical examples are presented to illustrate the models and solution, showing that the COA developed in the current study is suitable for the antagonistic situation.

Key words: course of action; influence net; multi-stage games with incomplete information

作战计划拟制的实质是对军事资源进行分配, 合理规划作战行动序列 (Course of Action, COA), 以改变战场态势达到目标状态。Haider 等^[1-6]用影响网络分析行动与态势之间的相互影响并优化行动序列。Wagenhals 等^[7]基于影响网络和 Petri 网开发了 CAESAR II/EB 用于 COA 的评估。Rafi 等^[8]提出了 ATIN (Activation Timed Influence Net) 模型用于对 COA 建模与优化。Zhang 等^[9]采用有限自动机和形式化语言来描述行动方案的逻辑特征以及定量特征, 并开发了一种形式化的建模工具 COAST 用于 COA 的调度。Pousi^[10]运用多准则影响图对 EBO 计划过程进行分析。Yaman 等^[11]运用模糊认知图支持 EBO 的计划过程。Wilkins^[12]、Myers^[13]、Tate^[14-15]分别建立了基于 HTN 的军事计划系统用于辅助 COA 的生成。目前对 COA 的研究主要是应用人工智能和自动规划领域的技术解决作战行动规划问

题, 在计划过程中考虑己方作战行动对战场态势的影响, 通过己方一系列的作战行动实现战场态势的转变, 未考虑到敌方可能采取针对性的作战行动对战场态势的影响。

战争关系着参战双方的生死存亡, 双方都期望战场态势按照己方的意愿发展, 而战场态势的演变取决于双方作战行动的共同作用。在制定作战计划时应该考虑到这种对抗, 这就需要在生成作战行动序列的过程中考虑敌我双方的博弈。其次由于参战双方会极力隐藏己方的真实信息以及战争本身的不确定性, 一方对另一方的兵力、武器装备以及作战目标等缺乏完全信息, 从博弈论的角度讲, 就是对另一方的可行策略、支付函数等博弈参数不确定, 即军事对抗活动是不完全信息博弈。通常情况下, 作战目标通过多个阶段的作战行动完成, 上一阶段博弈的结果是下一阶段博弈的条件, 双方博弈的目标是使最终的战场态势更

* 收稿日期: 2011-11-28

基金项目: 国家自然科学基金资助项目 (71101149); 国防科技大学优秀研究生创新资助项目 (B100502)

作者简介: 杜正军 (1982-), 男, 四川南充人, 博士研究生; E-mail: outstandingdzj@163.com;

陈超 (通信作者), 男, 讲师, 博士, E-mail: chenchao@gmail.com

有利于己方。基于此,本文从作战双方博弈的角度,用不完全信息多阶段博弈分析作战行动序列的生成。影响网络便于描述行动与效果之间的关系,提供了一种有效描述和分析作战行动序列的框架。在目前应用影响网络对作战行动序列的研究工作中^[1-8],并未考虑到资源约束以及双方博弈。本文在影响网络的基础上考虑资源约束以及敌我双方的博弈,建立了一个基于影响网络和不完全信息多阶段博弈的作战行动序列模型。

1 模型及求解方法

1.1 模型描述

文献[2-6,8]利用影响网络模型分析了作战行动与战场态势之间的相互影响,模型中的节点分为状态节点和作战行动节点,但只考虑了单方的作战行动对战场态势的影响。本模型中考虑红蓝双方的作战行动节点,战场态势由红蓝双方的作战行动共同决定。将作战行动过程视为不完全信息多阶段博弈,共有 K 个阶段,在博弈的每一阶段,红蓝双方根据资源约束确定己方和敌方的可行行动,并通过观察对方所选择的行动,根据贝叶斯法则来修正对敌方类型的概率分布,然后选择己方的行动。下面从状态向量、作战行动、作战行动与状态之间的相互影响、资源约束、类型空间以及行动与行动空间 6 个方面对模型进行描述。

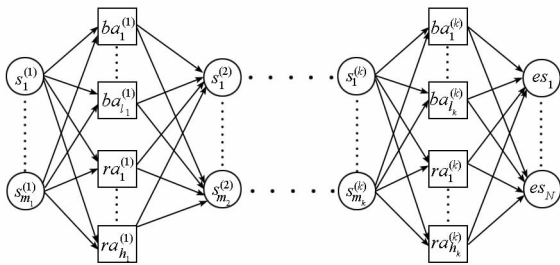


图 1 考虑双方博弈的影响网络 COA 模型
Fig. 1 COA models based on influence net considering actions of both sides

(1) 状态向量

第 k ($1 \leq k \leq K$) 阶段的状态向量 $\mathbf{s}^{(k)} = (s_1^{(k)}, \dots, s_{m_k}^{(k)})$, 最终阶段的状态向量 $\mathbf{s}^e = (s_1^e, \dots, s_{N_e}^e)$ 。第 k 阶段的状态向量既是第 k 阶段行动的前提条件,也是第 $k-1$ 阶段行动的效果。状态向量的每一维分量为 $\{0, 1\}$ 变量,0 表示该情况不出现,1 表示该情况出现。

(2) 作战行动项

作战行动项为最基本的任务单元,在作战过程中的每一阶段,双方都有一些作战行动项可供

选择。设第 k 阶段红方共有 h_k 项可选作战行动 $\{a_{r,1}^{(k)}, \dots, a_{r,h_k}^{(k)}\}$, 蓝方共有 l_k 项可选作战行动 $\{a_{b,1}^{(k)}, \dots, a_{b,l_k}^{(k)}\}$ 。

(3) 作战行动与状态之间的相互影响

第 k 阶段状态向量为 $\mathbf{s}^{(k)}$ 的情况下,红方第 i 项作战行动成功的概率为 $p(a_{r,i}^{(k)} | \mathbf{s}^{(k)})$, 蓝方第 j 项作战行动成功的概率为 $p(a_{b,j}^{(k)} | \mathbf{s}^{(k)})$;红方的行动为 $\varphi_i^{(k)}$, 蓝方的行动为 $\psi_j^{(k)}$ 的条件下,第 $k+1$ 阶段第 i 维状态向量 $s_i^{(k+1)}$ 的概率为 $p(s_i^{(k+1)} | \varphi_i^{(k)}, \psi_j^{(k)})$ 。

①第 k 阶段蓝方第 i 项作战行动的成功概率为:

$$p(a_{b,i}^{(k)} | \mathbf{s}^{(k)}) = \begin{cases} p(a_{b,i}^{(k)}) + hb_i(\mathbf{s}^{(k)})(1-p(a_{b,i}^{(k)})) & \text{if } hb_i(\mathbf{s}^{(k)}) \in [0,1] \\ p(a_{b,i}^{(k)}) + hb_i(\mathbf{s}^{(k)})p(a_{b,i}^{(k)}) & \text{if } hb_i(\mathbf{s}^{(k)}) \in [-1,0] \end{cases}$$

②记第 k 阶段蓝方的作战行动 $(a_{b,1}^{(k)}, \dots, a_{b,l_k}^{(k)})$ 以及红方的作战行动 $(a_{r,1}^{(k)}, \dots, a_{r,h_k}^{(k)})$ 为 $\mathbf{a}^{(k)}$, $\mathbf{a}^{(k)}$ 对状态向量 $\mathbf{s}^{(k)}$ 的第 i 维分量 $s_i^{(k)}$ 的影响参数为 $gs_i^{(k)}(\mathbf{a}^{(k)})$, $s_i^{(k)}$ 的概率为:

$$p(s_i^{(k)} | \mathbf{a}^{(k)}) = \begin{cases} p(s_i^{(k)}) + gs_i^{(k)}(\mathbf{a}^{(k)})(1-p(s_i^{(k)})) & \text{if } gs_i^{(k)}(\mathbf{a}^{(k)}) \in [0,1] \\ p(s_i^{(k)}) + gs_i^{(k)}(\mathbf{a}^{(k)})p(s_i^{(k)}) & \text{if } gs_i^{(k)}(\mathbf{a}^{(k)}) \in [-1,0] \end{cases}$$

(4) 资源约束

在实际的作战过程中,指挥员能调用的资源是有限的。设第 k 阶段蓝方的资源为 $\{s_{b,1}^{(k)}, \dots, s_{b,c_k}^{(k)}\}$, 红方的资源为 $\{s_{r,1}^{(k)}, \dots, s_{r,d_k}^{(k)}\}$ 。第 k 阶段蓝方第 i 项作战行动对第 j 种资源的需求为 $bn_{ij}^{(k)}$, 红方第 i 项作战行动对第 j 种资源的需求为 $rn_{ij}^{(k)}$ 。

(5) 类型空间

红蓝双方的类型空间分别为 $\Theta_r = \{\theta_{r,i}\}$, $\Theta_b = \{\theta_{b,j}\}$, $i \in \mathbf{N}, j \in \mathbf{N}$ 。给定类型空间 $(\theta_{r,i}, \theta_{b,j})$ 联合概率分布 $P(\theta_{r,i}, \theta_{b,j})$ 的情况下,红方根据己方的类型可以推测出蓝方的类型的概率分布 $P(\theta_{b,j} | \theta_r)$, θ_r 为红方实际的类型。同理,蓝方也可以根据 θ_b 推测出红方的类型分布 $P(\theta_{r,i} | \theta_b)$ 。第 k 阶段红方对蓝方所属类型的信念为 $P_r^{(k)} = (p_{r,1}^{(k)}, \dots, p_{r,i}^{(k)}, \dots, p_{r,N_b}^{(k)})$, 蓝方对红方所属类型的信念为 $P_b^{(k)} = (p_{b,1}^{(k)}, \dots, p_{b,j}^{(k)}, \dots, p_{b,N_r}^{(k)})$ 。

(6) 行动、行动空间

①第 k 阶段红方的行动指在 h_k 项作战行动项中选择其中满足资源约束的 n 项 ($0 \leq n \leq h_k$) 执行,定义为向量 $(a_{r,1}^{(k)}, \dots, a_{r,h_k}^{(k)})$, 向量中的分量为 $\{0, 1\}$ 变量,0 表示不执行该作战行动项,1 表示执行该作战行动项,记为 $\varphi_i^{(k)}$ 。红方所有的可行行动构成了第 k 阶段红方的行动集 $\Phi^{(k)} = \{\varphi_1^{(k)}, \varphi_2^{(k)}, \dots, \varphi_{f_k}^{(k)}\}$, $f_k \leq 2^{h_k}$ 。同理定义蓝方第 k

阶段的行动 $\psi_j^{(k)}$ 以及行动集 $\Psi^{(k)} = \{\psi_1^{(k)}, \psi_2^{(k)}, \dots, \psi_{g_k}^{(k)}\}$, $g_k \leq 2^k$ 。

②红方第1阶段行动集 $\Phi^{(1)}$ 到第K阶段的行动集 $\Phi^{(K)}$ 的笛卡尔积 $\Phi^{(1)} \times \Phi^{(2)} \times \dots \times \Phi^{(K)}$ 构成了红方的行动空间 Φ , 即 $\Phi = \Phi^{(1)} \times \Phi^{(2)} \times \dots \times \Phi^{(K)}$ 。同理定义蓝方的行动空间 Ψ , $\Psi = \Psi^{(1)} \times \Psi^{(2)} \times \dots \times \Psi^{(K)}$ 。

1.2 数学模型

综上,作战行动序列可以用六元组 $G = \langle W, \Omega, U, \Theta, P, H \rangle$ 表示,其中:

(1) $W = \{R, B\}$ 表示局中人,即红蓝双方;

(2) $\Omega = \{\Phi, \Psi\}$ 表示红蓝双方的行动空间;

① $\Phi = \Phi^{(1)} \times \Phi^{(2)} \times \dots \times \Phi^{(K)}$

$\Phi^{(k)}$ 的任意一项可行行动 $\varphi_m^{(k)}$ 须满足资源约束:

$$\sum_{i=1}^{h_k} ra_i^{(k)} \times rn_{ij}^{(k)} \leq rs_j^{(k)}$$

$m = 1, \dots, f_k; j = 1, \dots, c_k; k = 1, \dots, K$

② $\Psi = \Psi^{(1)} \times \Psi^{(2)} \times \dots \times \Psi^{(K)}$

$\Psi^{(k)}$ 的任意一项可行行动 $\psi_n^{(k)}$ 须满足资源约束:

$$\sum_{i=1}^{l_k} ba_i^{(k)} \times bn_{ij}^{(k)} \leq bs_j^{(k)}$$

$n = 1, \dots, g_k; j = 1, \dots, d_k; k = 1, \dots, K$

(3) $H = \{H_r^{(k)}, H_b^{(k)}\}$, 表示第k阶段的行动历史,即在作战行动过程中双方实际执行的行动。 $H_r^{(k)} = \{\varphi^{(1)}, \varphi^{(2)}, \dots, \varphi^{(k-1)}\}$, $H_b^{(k)} = \{\psi^{(1)}, \psi^{(2)}, \dots, \psi^{(k-1)}\}$, $1 < k \leq K$

(4) $\Theta = \{\Theta_r, \Theta_b\}$ 表示类型空间,即红蓝双方可能的类型的集合,红方共有 N_r 种类型,蓝方共有 N_b 种类型; $\Theta_r = \{\theta_{r,1}, \theta_{r,2}, \dots, \theta_{r,N_r}\}$, $\Theta_b = \{\theta_{b,1}, \theta_{b,2}, \dots, \theta_{b,N_b}\}$

(5) $P = \{P_r^{(k)}, P_b^{(k)}\}$, $P_r^{(k)}$ 表示第k阶段红方对蓝方所属类型的信念, $P_b^{(k)}$ 表示第k阶段蓝方对红方所属类型的信念,信念需根据作战行动过程中获取的信息进行更新;

① 当 $k = 1$ 时: $P_r^{(1)} = (p_{r,1}^{(1)}, \dots, p_{r,i}^{(1)}, \dots, p_{r,N_b}^{(1)})$, $P_b^{(1)} = (p_{b,1}^{(1)}, \dots, p_{b,j}^{(1)}, \dots, p_{b,N_r}^{(1)})$, $p_{r,i}^{(1)} = P(\theta_{b,i} | \theta_r)$, $p_{b,j}^{(1)} = P(\theta_{r,j} | \theta_b)$, $i = 1, \dots, N_b, j = 1, \dots, N_r$

② 当 $1 < k \leq K$: $P_r^{(k)} = (p_{r,1}^{(k)}, \dots, p_{r,i}^{(k)}, \dots, p_{r,N_b}^{(k)})$, $P_b^{(k)} = (p_{b,1}^{(k)}, \dots, p_{b,j}^{(k)}, \dots, p_{b,N_r}^{(k)})$, $p_{r,i}^{(k)} = P(\theta_{b,i} | \theta_r, H_b^{(k)})$, $p_{b,j}^{(k)} = P(\theta_{r,j} | \theta_b, H_r^{(k)})$, $i = 1, \dots, N_b, j = 1, \dots, N_r$

(6) U 表示支付函数,由双方类型组合以及策

略确定; $U = f(\alpha_{i_1}^{(k)}, \beta_{j_1}^{(k)}, \theta_{r,i_2}, \theta_{b,j_2}), k = 1, \dots, K, i_1 = 1, \dots, (\Delta r_k)^{N_r}, j_1 = 1, \dots, (\Delta b_k)^{N_b}, i_2 = 1, \dots, N_r, j_2 = 1, \dots, N_b$

1.3 模型求解

不完全信息动态博弈的解为完美贝叶斯均衡,完美贝叶斯均衡要求在博弈的每一阶段,局中人的均衡策略在每一个信息集开始的后续博弈上构成贝叶斯均衡,并用贝叶斯法则修正关于其他局中人类型的信息^[17]。在本文中,为便于模型求解,将博弈用策略式表示。在博弈的第一阶段,根据类型的先验概率以及赢得矩阵,用海萨尼转换为标准形式,确定双方的均衡策略以及本阶段的行动。在从第二阶段开始的后续博弈中,双方根据对方的行动修正对方类型的后验概率,并确定该阶段的行动。具体求解步骤如下:

(1) 根据资源约束分析双方行动空间,将博弈用策略式表示,针对双方的每一种类型组合,用影响网络计算出赢得矩阵。

(2) 第1阶段,根据先验概率分布 P_{ij} , 以及赢得矩阵 $A_1^{(ij)}$, 用海萨尼转换转化为标准形式。博弈的标准形式 G_1^* 中红方标准化策略是类型空间 Θ_r 到策略的函数,红方的标准化策略 $x_i^{(1)}$ 表示为 $x_i^{(1)} = (\alpha_{i_1}^{(1)}, \alpha_{i_2}^{(1)}, \dots, \alpha_{i_{N_r}}^{(1)})$, 即红方属于第一种类型则采用策略 $\alpha_{i_1}^{(1)}$, 属于第二种类型则采用策略 $\alpha_{i_2}^{(1)}$, 属于第 N_r 种类型则采用策略 $\alpha_{i_{N_r}}^{(1)}$, 同理蓝方的标准化策略表示为 $y_j^{(1)} = (\beta_{j_1}^{(1)}, \beta_{j_2}^{(1)}, \dots, \beta_{j_{N_b}}^{(1)})$ 。标准化赢得矩阵记为 $\Gamma^{(1)}$, $\Gamma^{(1)}$ 的元素 $\Gamma_{ij}^{(1)}$ 的值等于 $\sum_{n=1}^{N_r} \sum_{m=1}^{N_b} p_{mn} \times A_{i_1, i_n, j_m}^{(nm)}$, i_n, j_m 分别为标准化策略 $x_i^{(1)}, y_j^{(1)}$ 中对应于红方第 n 种类型的策略的下标以及蓝方第 m 种类型的策略的下标。求解得到红蓝双方的均衡策略,双方根据均衡策略执行第一阶段的行动 $\varphi^{(1)}, \psi^{(1)}$ 。

(3) 第 k ($1 < k \leq K$) 阶段,根据 $k-1$ 阶段的行动历史,用贝叶斯法则修正对方类型的后验概率。蓝方根据红方 $k-1$ 阶段的行动 $\varphi^{(k-1)}$, 修正红方属于类型 $\theta_{r,n}$ ($n = 1, 2, \dots, N_r$) 的后验概率为

$$P(r\theta_n | \varphi^{(k-1)}) = \frac{P(\varphi^{(k-1)} | r\theta_n) \times bp_n^{(k-1)}}{\sum_{i=1}^{N_r} bp_i^{(k-1)} \times P(\varphi^{(k-1)} | r\theta_i)}$$

红方根据蓝方 $k-1$ 阶段的行动 $\psi^{(k-1)}$, 修正蓝方属于类型 $\theta_{b,m}$ ($m = 1, 2, \dots, N_b$) 的后验概率为

$$P(b\theta_m | \psi^{(k-1)}) = \frac{P(\psi^{(k-1)} | b\theta_m) \times rp_m^{(k-1)}}{\sum_{i=1}^{N_b} rp_i^{(k-1)} \times P(\psi^{(k-1)} | b\theta_i)}$$

据双方类型的后验概率分布 $P_{ij}^{(k)}$ 以及对应的赢得矩阵 $A_k^{(ij)}$, 用海萨尼转换重新化为标准形式 G_k^* , 红方的标准化策略 $x_i^{(k)}$ 表示为 $x_i^{(k)} = (\alpha_{i_1}^{(k)}, \alpha_{i_2}^{(k)}, \dots, \alpha_{i_{N_r}}^{(k)})$ 。蓝方的标准化策略表示为 $y_j^{(k)} = (\beta_{j_1}^{(k)}, \beta_{j_2}^{(k)}, \dots, \beta_{j_{N_b}}^{(k)})$ 。标准化赢得矩阵记为 $\Gamma^{(k)}$, $\Gamma^{(k)}$ 的元素的 $\Gamma_{ij}^{(k)}$ 值等于 $\sum_{n=1}^{N_r} \sum_{m=1}^{N_b} P_{nm}^{(k)} \times Ak_{i,j,m}^{(nm)}$ 。求解得到红蓝双方的标准化均衡策略。红蓝双方根据均衡策略执行第 k 阶段的行动 $\varphi^{(k)}$ 、 $\psi^{(k)}$ 。

2 计算实例

2.1 算例

考虑一个三阶段的作战行动序列的网络, 如图 2 所示。

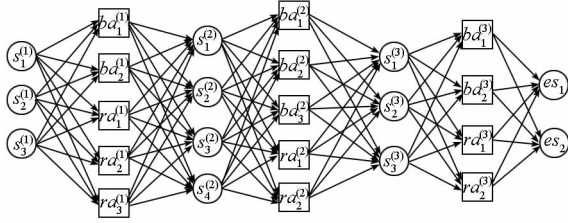


图 2 示例行动影响图

Fig. 2 Example of COA models based on influence net

(1) 每一阶段的状态向量分别为 $(s_1^{(1)}, s_2^{(1)}, s_3^{(1)})$, $(s_1^{(2)}, s_2^{(2)}, s_3^{(2)}, s_4^{(2)})$, $(s_1^{(3)}, s_2^{(3)}, s_3^{(3)})$, 最后一阶段状态向量为 (s_1^e, s_2^e) 。

(2) 红方每一阶段可选作战行动项分别为 $\{a_{r,1}^{(1)}, a_{r,2}^{(1)}, a_{r,3}^{(1)}\}$, $\{a_{r,1}^{(2)}, a_{r,2}^{(2)}\}$, $\{a_{r,1}^{(3)}, a_{r,2}^{(3)}\}$, 蓝方每一阶段可选作战行动项为 $\{a_{b,1}^{(1)}, a_{b,2}^{(1)}\}$, $\{a_{b,1}^{(2)}, a_{b,2}^{(2)}\}$, $\{a_{b,1}^{(3)}, a_{b,2}^{(3)}\}$ 。

(3) 随机生成节点之间的影响参数。

(4) 红方每一阶段的资源分别为 $(s_{r,1}^{(1)}, s_{r,2}^{(1)}, s_{r,3}^{(1)})$, $(s_{r,1}^{(2)}, s_{r,2}^{(2)})$, $(s_{r,1}^{(3)}, s_{r,2}^{(3)}, s_{r,3}^{(3)})$, 蓝方每一阶段资源分别为 $(s_{b,1}^{(1)}, s_{b,2}^{(1)})$, $(s_{b,1}^{(2)}, s_{b,2}^{(2)}, s_{b,3}^{(2)})$, $(s_{b,1}^{(3)}, s_{b,2}^{(3)}, s_{b,3}^{(3)})$ 。随机生成红蓝双方的资源约束以及作战行动项对资源的需求。

(5) 类型空间及其信念概率。

红蓝双方都有两种类型, 分别记为 $\theta_{r,1}$ 、 $\theta_{r,2}$ 和 $\theta_{b,1}$ 、 $\theta_{b,2}$ 。类型的联合概率分布为 $P(\theta_{r,1}, \theta_{b,1}) = 0.1$, $P(\theta_{r,1}, \theta_{b,2}) = 0.4$, $P(\theta_{r,2}, \theta_{b,1}) = 0.3$, $P(\theta_{r,2}, \theta_{b,2}) = 0.2$ 。红方实际类型为 $\theta_{r,1}$, 蓝方实际类型为 $\theta_{b,2}$ 。根据边缘概率公式, 红方对蓝方类型的判断为 $P(\theta_{b,1} | \theta_{r,1}) = 0.2$, $P(\theta_{b,2} | \theta_{r,1}) = 0.8$ 。蓝方对红方的类型判断为 $P(\theta_{r,1} | \theta_{b,2}) = 2/3$, $P(\theta_{r,2} | \theta_{b,2}) = 1/3$ 。

2.2 求解

(1) 分析红蓝双方每一阶段的可行行动集

易知蓝方第一阶段满足资源约束的行动集为 $\{(0,1), (1,0), (1,1)\}$; 第二阶段的行动集为 $\{(0,0,1), (0,1,0), (1,0,0), (0,1,1), (1,0,1)\}$; 第三阶段行动集为 $\{(0,1), (1,0)\}$ 。红方第一阶段行动集为 $\{(0,0,1), (0,1,0), (1,0,0), (1,0,1)\}$; 第二阶段行动集为 $\{(0,1), (1,0)\}$; 第三阶段行动集为 $\{(0,1), (1,0), (1,1)\}$ 。

(2) 红蓝双方博弈过程

① 第一阶段

红蓝双方根据先验概率分布 $P = \begin{bmatrix} 0.1 & 0.4 \\ 0.3 & 0.2 \end{bmatrix}$, 以及在各种类型组合下的赢得矩阵 $A_1^{(11)}, A_1^{(12)}, A_1^{(21)}, A_1^{(22)}$, 用海萨尼转换得到标准形式为 $\Gamma^{(1)}$, $\Gamma^{(1)}$ 为一个 $24^2 \times 30^2$ 的矩阵, $\Gamma_{ij}^{(1)} = \sum_{n=1}^2 \sum_{m=1}^2 P_{nm} \times A_{1,i_n,j_m}^{(nm)}$ 。求解得到

(I) 红方的均衡策略为 $(x_1^{(1)}, x_2^{(1)}) = (0.763, 0.237)$, 其中 $x_1^{(1)} = (\alpha_{11}^{(1)}, \alpha_{12}^{(1)})$, $\alpha_{11}^{(1)} = (0, 1, 0, 0, 1, 1, 1, 1)$, $\alpha_{12}^{(1)} = (1, 0, 0, 0, 1, 0, 1, 1)$, $x_2^{(1)} = (\alpha_{21}^{(1)}, \alpha_{22}^{(1)})$, $\alpha_{21}^{(1)} = (0, 1, 0, 1, 0, 1, 1, 1)$, $\alpha_{22}^{(1)} = (1, 0, 0, 0, 1, 0, 1, 1)$

(II) 蓝方的均衡策略为 $(y_1^{(1)}, y_2^{(1)}) = (0.176, 0.824)$, 其中 $y_1^{(1)} = (\beta_{11}^{(1)}, \beta_{12}^{(1)})$, $\beta_{11}^{(1)} = (1, 0, 0, 1, 1, 0, 1, 1)$, $\beta_{12}^{(1)} = (1, 0, 0, 1, 1, 0, 1, 1)$, $y_2^{(1)} = (\beta_{21}^{(1)}, \beta_{22}^{(1)})$, $\beta_{21}^{(1)} = (1, 0, 0, 0, 1, 0, 1, 1)$, $\beta_{22}^{(1)} = (1, 0, 0, 1, 1, 0, 1, 1)$

分析可知: 红方类型为 $\theta_{r,1}$, 则第一阶段采取行动 $\varphi_1^{(1)} = (0, 1, 0)$, 类型为 $\theta_{r,2}$, 则第一阶段采取行动 $\varphi_2^{(1)} = (1, 0, 0)$ 。红方实际类型为 $\theta_{r,1}$, 因此采取行动 $\varphi^{(1)} = (0, 1, 0)$ 。第一阶段作战行动完成后, 蓝方能判断出红方的类型。 $P(\theta_{r,1} | \theta_{b,2}, \varphi^{(1)}) = 1$, $P(\theta_{r,1} | \theta_{b,2}, \varphi^{(1)}) = 0$ 。蓝方类型为 $\theta_{b,1}$ 、 $\theta_{b,2}$ 在第一阶段都采取行动 $\psi^{(1)} = (1, 0)$ 。第一阶段作战行动完成后, 红方从蓝方的行动中未获取到信息更新对蓝方类型的信念。 $P(\theta_{b,1} | \theta_{r,1}, \psi^{(1)}) = 0.2$, $P(\theta_{b,2} | \theta_{r,1}, \psi^{(1)}) = 0.8$ 。

② 第二阶段

第一阶段行动结束后战场态势 $(s_1^{(2)}, s_2^{(2)}, s_3^{(2)}, s_4^{(2)}) = (0.5651, 0.4449, 0.0372, 0.9914)$ 为第二阶段博弈开始的条件。

(I) 从红方的角度来说, 类型组合为 $(\theta_{r,1}, \theta_{b,1})$ 和 $(\theta_{r,1}, \theta_{b,2})$ 的概率分别是 0.2、0.8。且红方知道蓝方能确知红方的类型为 $\theta_{r,1}$ 。通过影响网络

计算出两种类型组合下的赢得矩阵分别为

$$A_2^{(11)} = \begin{bmatrix} 0.2146 & 0.2952 & 0.2437 & 0.4085 & 0.2119 & 0.2855 & 0.247 & 0.4263 & 0.2202 & 0.3323 \\ 0.1812 & 0.7956 & 0.4043 & 0.9708 & 0.1614 & 0.7806 & 0.4323 & 0.9977 & 0.2327 & 0.8511 \\ 0.5654 & 0.9331 & 0.7732 & 0.9912 & 0.5475 & 0.9278 & 0.8047 & 0.9966 & 0.63 & 0.9408 \\ 0.2045 & 0.2545 & 0.2439 & 0.4098 & 0.2136 & 0.2953 & 0.247 & 0.4265 & 0.211 & 0.2919 \\ 0.1035 & 0.7328 & 0.4063 & 0.9727 & 0.1761 & 0.7953 & 0.4326 & 0.998 & 0.1596 & 0.7892 \\ 0.4911 & 0.9136 & 0.7754 & 0.9916 & 0.5648 & 0.9304 & 0.8051 & 0.9966 & 0.5569 & 0.9235 \\ 0.814 & 0.8531 & 0.8402 & 0.8768 & 0.8111 & 0.851 & 0.8392 & 0.8805 & 0.8064 & 0.8607 \\ 0.1067 & 0.9471 & 0.265 & 0.9929 & 0.0899 & 0.9429 & 0.261 & 0.9977 & 0.0688 & 0.9546 \\ 0.1287 & 0.9207 & 0.1897 & 0.9918 & 0.123 & 0.913 & 0.1947 & 0.9877 & 0.1345 & 0.8968 \\ 0.8064 & 0.8446 & 0.8402 & 0.877 & 0.81 & 0.8531 & 0.8391 & 0.8805 & 0.8019 & 0.8523 \\ 0.06 & 0.9315 & 0.265 & 0.9933 & 0.0848 & 0.9453 & 0.2606 & 0.9977 & 0.0395 & 0.9405 \\ 0.1084 & 0.9006 & 0.1901 & 0.9916 & 0.1253 & 0.9092 & 0.1947 & 0.9874 & 0.117 & 0.8857 \end{bmatrix}$$

利用海萨尼转换转化为标准形式 $\Gamma^{(2)}$ 。由于红方知道蓝方能确知红方的类型为 $\theta_{r,1}$,而红方不能确知蓝方的类型。故求得红方的策略为 $x^{(2)} = (0, 1, 0, 1)$ 。蓝方的标准化策略为 $y^{(2)} = (\beta_1^{(2)}, \beta_2^{(2)})$, $\beta_1^{(2)} = (0, 1, 1, 0, 1)$, $\beta_2^{(2)} = (1, 0, 1, 0, 1)$

(II) 蓝方已经确知红方类型为 $\theta_{r,1}$,即已经知道类型组合为 $(\theta_{r,1}, \theta_{b,2})$,则赢得矩阵为 $A_2^{(12)}$ 。对于蓝方,知道红方分析的结果会采取行动 $\alpha^{(2)} = (0, 1, 0, 1)$,而蓝方的均衡策略为 $\beta_2^{(2)} = (1, 0, 1, 0, 1)$ 。

分析可知:红方第二阶段采取行动 $\varphi^{(2)} = (0, 1)$,蓝方第二阶段采取行动 $\psi^{(2)} = (1, 0, 1)$ 。且红方从蓝方的行动中更新对蓝方类型的信念, $P(\theta_{b,1} | \theta_{r,1}, \psi^{(2)}) = 0, P(\theta_{b,2} | \theta_{r,1}, \psi^{(2)}) = 1$ 。

③第三阶段

红蓝双方都知道类型组合为 $(\theta_{r,1}, \theta_{b,2})$,通过影响网络计算出赢得矩阵为

$$A_3 = \begin{bmatrix} 0.8063 & 0.8607 \\ 0.0687 & 0.9546 \\ 0.1344 & 0.8967 \end{bmatrix}$$

求解得出红方会采取行动 $\varphi^{(3)} = (0, 1)$,蓝方行动为 $\psi^{(3)} = (0, 1)$,赢得值为 0.8063。

3 结论

战争是激烈的对抗活动,战场态势的演变取决于作战双方作战行动的共同作用,于是双方决策、行动的正确性相互依赖。考虑作战计划的对抗性以及不确定性,建立了一个基于多阶段不完全信息博弈的作战行动过程模型,用影响网络分析作战行动与态势之间的相互影响,最后给出了一个计算实例并进行了求解。与只考虑单方作战

行动与战场态势之间影响的方法相比,本文提出的方法更适应于对抗环境,为指挥员制定作战计划提供了一种辅助决策的方法。

参考文献 (References)

- [1] Rosen J A, Smith W L. Influence net modeling with causal strengths: an evolutionary approach [C]// Command and Control Research and Technology Symposium, Naval Post Graduate School, Monterey, 1996:25 - 28.
- [2] Haider S, Zaidi A K, Levis A H. A heuristic approach for best set of actions determination in influence nets [C]// Proceedings of IEEE International Conference on Information Reuse and Integration, Las Vegas, 2004.
- [3] Haider S, Zaidi A K. Transforming timed influence nets into time sliced bayesian networks [C]// Proceedings of Command and Control Research and Technology Symposium, 2004.
- [4] Haider S, Levis H. Modeling time-varying uncertain situations using dynamic influence net [J]. International Journal of Approximate Reasoning, 2008, 49(2): 488 - 502.
- [5] Haider S, Levis A H. Finding effective courses of action using particle swarm optimization [C]// Proceedings of World Congress on Computational Intelligence, Hong Kong, 2008, 1135 - 1140.
- [6] Haider S, Levis A H. Effective courses of action determination to achieve desired effects [J]. IEEE Transactions on System, Man and Cybernetics, 2007, 37(6): 1140 - 1150.
- [7] Wagenhals L W, Levis A H. Modeling support of effects based operations in war games [C]// 7th Command and Control Research and Development Symposium, Naval Post Graduate School, Monterey, Ca, 2002.
- [8] Rafi F M, Zaidi A K, Kazakos P P, et al. Optimization of actions in activation timed influence nets [J]. Informatica, 2009, 285 - 296.
- [9] Zhang L, Mitchell B, Janczura C. COAST: an operational planning tool for course of action development and analysis [C]// Symposium for the 9th International Command and Control Research and Technology Symposium, 2003.

参考文献 (References)

- [1] Goddard J S. Pose and motion estimation from vision using dual quaternion-based extended kalman filtering [D]. Knoxville: the University of Tennessee, 1997.
- [2] 丁尚文, 王惠南, 刘海颖, 等. 基于对偶四元数的航天器交会对接位姿视觉测量算法[J]. 宇航学报, 2009,30(6): 2145-2150.
DING Shangwen, WANG Huinan, LIU Haiying, et al. Algorithm of vision measure for relative position and pose of RVD spacecrafts based on dual-quaternion [J]. Journal of Astronautics, 2009,30(6): 2145-2150. (in Chinese)
- [3] 张世杰. 基于单目视觉的航天器相对导航理论与算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2005.
ZHANG Shijie. Study on theory and algorithms of monocular vision-based navigation for spacecraft [D]. Harbin: Harbin Institute of Technology. (in Chinese)
- [4] 武元新. 对偶四元数导航算法与非线性高斯滤波研究[D]. 长沙: 国防科技大学, 2005.
WU Yuanxin. Research on dual-quaternion navigation algorithm and monocular gaussian filtering [D]. Changsha: National University of Defense Technology, 2005. (in Chinese)
- [5] Olson C F, Matthies L H, Schoppers M, et al. Rover navigation using stereo ego-motion [J]. Robotics and Autonomous Systems, 2003,43(4): 215-229.
- [6] Allen D W, Johnson N, Alison A P. Vision-aided inertial navigation for flight control [J]. Journal of Aerospace Computing, Information and Communication, 2005,(2): 348-360.
- [7] Andreas H. Relative position sensing by fusing monocular vision and inertial rate sensors [D]. Stanford: Stanford University, 2003.
- [8] Mourikis A I, Roumeliotis S I. A multi-state constraint Kalman filter for vision-aided inertial navigation [C]// Proceedings of the IEEE International Conference on Robotics and Automation, Roma, Italy, 2007.
- [9] Stein G P, Mano O, Shashua A. A robust method for computing vehicle ego-motion [C]// Proceedings of the IEEE Intelligent Vehicles Symposium, Dearborn, USA, 2000.
- [10] 刘威, 魏存伟, 赵逢, 等. 基于单目视觉的自行车运动参数鲁棒估计[J]. 机器人, 2009,31(1): 20-26.
LIU Wei, WEI Cunwei, ZHAO Feng, et al. A robust method for vehicle ego-motion estimation based on monocular vision [J]. Robot, 2009,31(1): 20-26. (in Chinese)
- [11] 李万里. 车载组合导航自适应滤波及抗野值算法研究[D]. 长沙: 国防科技大学, 2008.
LI Wanli. Research on adaptive Kalman filter and fault-tolerant algorithm used in in-vehicle integrated navigation system [D]. Changsha: National University of Defense Technology, 2008. (in Chinese)
- [12] Bäumker M, Heimes F J. New calibration and computing method for direct georeferencing of image and scanner data using the position and angular data of an hybrid inertial navigation system [M]. Frankfurt am Main: Bundesamt für Kartographie und Geodäsie, 2002: 197-212.
- [13] Feng G H, Wu W Q, Cao J L, et al. Algorithm for monocular visual odometry/SINS integrated navigation [J]. Journal of Chinese Inertial Technology, 2011,19(3): 210-215.
- [10] Pousi J. Decision analytical approach to effect based operations [D]. Espoo: Helsinki University of Technology, 2009.
- [11] Yaman D, Polat S. A fuzzy cognitive map approach for effect based operations: an illustrative case [J]. Information Sciences, 2009:382-403.
- [12] Wilkins D E, Desimone R V. Applying an AI planner to military operations planning [R]. SRI International Research Report, 1993.
- [13] Myers K, Tyson W M. PASSAT: A user-centric planning framework [C]// Proceedings of the 3rd International NASA Workshop on Planning and Scheduling for Space. Houston, USA. 2002.
- [14] Tate A, Levine J, Jarvis P, et al. Using AI planning technology for army small unit operations [C]// Proceedings of the Fifth International Conference on Artificial Intelligence Planning Systems, 2000.
- [15] Potter S, Tate A, Wickler G. Using I-X process panels as intelligent to do lists for agent coordination in emergency response [C]// Proceedings of the Information Systems for Crisis Response and Management, Newark, New Jersey, USA, 2006.
- [16] Fudenberg D, Tirole J. Game theory [M]. Massachusetts: The MIT Press, 1991.
- [17] John H. Games with incomplete information played by "bayesian" players [J]. Management Science, 1967: 159-182.

(上接第 67 页)