文章编号:1001 - 2486(2006)03 - 0124 - 04

动态武器目标分配问题的马尔可夫性※

蔡怀平,刘靖旭,陈英武

(国防科技大学 信息系统与管理学院,湖南 长沙 410073)

摘 要:动态武器目标分配(weapon target assignment,WTA)问题是军事运筹学研究的重要理论问题,也是作战指挥决策中迫切需要解决的现实问题。在对动态 WTA 问题进行描述分析的基础上,运用随机过程理论证明了动态 WTA 过程的马尔可夫性;给出了该马尔可夫决策过程的状态转移概率的解析表达式,并对其状态特点进行了简要分析。研究结果可以为动态 WTA 及相关问题的研究提供理论和方法依据。

关键词:运筹学;动态武器目标分配;马尔可夫决策过程;数学模型

中图分类号:O122　　文献标识码:A

On the Markov Characteristic of Dynamic Weapon Target Assignment Problem

CAI Huai-ping, LIU Jing-xu, CHEN Ying-wu

(College of Information System and Management, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: Dynamic weapon target assignment (WTA) is an important theoretical problem for military operation research and a challenging practical concern in the process of combat command. Based on the description and analysis of dynamic WTA problem, the Markov characteristic of the WTA problem was proved with stochastic theory, and the transition possibility matrix of the Markov decision process was also presented. Meanwhile, the characteristic of the transition probability matrix was analyzed. The result of the research sheds light on the study on the WTA problem in theory and method.

Key words: operations research; dynamic weapon target assignment; Markov decision process; mathematical model

武器目标分配(weapon target assignment,WTA)问题,也称为火力分配问题,主要研究防御作战过程中武器最优配置问题。目前,对静态 WTA 问题的研究较为深入;而动态 WTA 问题,由于需考虑时间等因素,相对较为复杂,目前仍无非常有效的解决方法,因而成为当前研究的一个热点[1-4]。马尔可夫过程理论是研究动态系统的一个有效工具,因而在武器作战效能评估以及作战过程策略优化研究中,通常假设作战过程中的武器与目标的分配过程为马尔可夫过程[5-6]。然而,如何确定马尔可夫过程的转移概率矩阵,一直是研究实际应用问题的一个难点。一般方法是假设该转移概率已知或采用 Petri 网等建模方法进行计算[8]。本文在对动态 WTA 问题作了较为一般性的假设后,证明了该问题可用马尔可夫决策过程进行描述建模,并给出状态转移概率的解析表达式,为动态 WTA 及武器作战效能评估等相关问题的研究提供了理论和方法依据。

1 问题描述与假设

1.1 基本假设

动态 WTA 问题,即研究作战过程中武器的动态最优配置问题。WTA 问题可分为直接对抗式 WTA 或间接对抗式 WTA。所谓直接对抗式 WTA,是指在作战双方直接进行对抗的情况下,进行武器目标分配,双方的作战目的都是为了直接消灭对方,如坦克战中的作战双方均是为了消灭对方的装甲车辆;间接对抗式 WTA,是指攻击方的武器(这里称作防御方的目标)的作战目的是为了摧毁防御方所保护的资

※ 收稿日期:2005 - 12 - 22
基金项目:国家部委资助项目
作者简介:蔡怀平(1971—),男,博士生。

源,而防御方为了使所保护的资源不受损失或损失较小,而对敌目标有选择地分配武器进行打击,如要地防空作战中的导弹部队对袭击所保卫的重要城市或设施的敌方飞机进行导弹拦截等。直接对抗式 WTA 与间接对抗式 WTA 的主要区别在于目标所攻击的对象不同,前者中的目标所打击对象是防御方的武器,后者中的目标所打击的对象是防御方的武器所防护的资源,而不是与防御方的武器直接进行交战。此外,两者的研究方法也不同,前者一般采用对策论、博弈论及兰切斯特方程等方法进行研究[6,7],而后者一般采用排队论及规划方法进行研究[11]。本文侧重于研究间接对抗式 WTA,即防御作战中的动态 WTA 问题。

为了建立数学模型,对实际对抗过程进行适当假设如下:

(1)各个目标到达防御方杀伤区的时间间隔分布记为 $A(t)$。根据威胁程度大小,可以将目标分为 m 类,目标威胁类的集合记为 $\boldsymbol{M}=\{k_i \mid 1\leqslant i\leqslant m, i\in \mathbf{N}\}$。第 k_i 类目标的威胁程度记为 r_{k_i},不妨设 $r_{k_1}\geqslant r_{k_2}\geqslant\cdots\geqslant r_{k_m}$。

(2)防御方的武器系统由 n 个武器单元组成,每个武器一次只能射击一个目标,且每个武器单元射击目标所需要的时间是相互独立、同分布的随机变量,即服从参数为 μ 的负指数分布。当一个目标进入防御方杀伤区时,武器系统根据该目标威胁程度大小及系统状态决定对其是否分配武器单元进行打击。武器系统对目标分配武器单元的原则是,对威胁大的目标优先分配,只要存在空闲的武器单元,即对最大威胁类的目标分配武器。

(3)当新目标到达杀伤区域时,若 n 个武器单元都正在进行射击,则该目标将脱离杀伤区域,即突破防区,将可能对防御方的资源造成损失;若存在空闲武器单元时,作战指挥系统将做出对该目标是否分配武器进行打击的决策。

1.2 变量描述

定义 1 武器系统的状态是指系统中正在进行射击的武器单元数量。如系统中有 i 个武器单元正在进行射击,则该系统的状态记为状态 i。所有的状态 i 构成状态空间 S,即 $S=\{i \mid i=0,1,2,\cdots,n\}$。

定义 2 将目标威胁类的一个子集称为一个方案,记为 $\boldsymbol{a}$。方案 $\boldsymbol{a}_i$ 表示若按威胁度大小排序,其包含的元素为前 $m-i$ 类目标,即 $\{k_1,k_2,\cdots,k_{m-i}\}$,所有的方案 $\boldsymbol{a}_i$ 构成备选方案集 $\boldsymbol{A}=\{\boldsymbol{a}_i \mid 0\leqslant i\leqslant m\}$。

系统处于状态 i 时,从备选方案集中选择的方案记为决策 $\boldsymbol{\pi}_i$,$\boldsymbol{\pi}_i\subseteq \boldsymbol{A}$ $(i=0,1,2,\cdots,n)$。当一个目标 s 进入防御方杀伤区时,武器系统处于状态 i,即指挥系统选择决策 $\boldsymbol{\pi}_i$。若该目标的威胁类型 $k_s\in \boldsymbol{\pi}_i$,则对该目标分配武器单元进行拦截,否则不分配武器单元进行拦截。当武器系统处于状态 n 时,即所有武器单元都正在进行射击,无法对新目标进行分配武器,即 $\pi_n\equiv\varnothing$,此时目标将可能脱离杀伤区域。

定义 3 转移概率 $p(j \mid i,\pi_i)$ 表示武器系统的当前状态为 i,当下一目标 s 到来时,系统选择策略 π_i,系统状态转为 j 的概率,简记为 p_{ij}。所有 p_{ij} 组成转移概率矩阵,简记为 $\boldsymbol{p}$。

2 马尔可夫性证明

2.1 定理的提出与证明

引理 在上述假设条件下,以 n_s 表示新目标 s 进入杀伤区域时系统所处的状态,v_s 为在 (t_s,t_{s+1}) 完成射击的武器单元数,且系统对来袭目标采取的是空闲即射击的原则,则 v_s 在条件 $\{n_s=i\}$ 下的条件分布为

$$P(v_s=h \mid n_s=i,\boldsymbol{\pi}_i)=\int_{0^-}^{\infty} c_{i+1}^{h}(1-\mathrm{e}^{-\mu t})^{h}\mathrm{e}^{-(i+1-h)\mu t}\mathrm{d}A(t),\quad 0\leqslant h\leqslant i+1, i<n \tag{1}$$

证明 记武器单元对目标开始射击的时刻为 0。根据负指数分布的无记忆特性可知,若一武器单元已对一目标射击了一段时间 t 而未结束,在该条件下继续进行射击的时间大于 Δt 的概率为

$$P\{T\geqslant t+\Delta t \mid T\geqslant t\}=\frac{P\{T\geqslant t+\Delta t\}}{P\{T\geqslant t\}}=\frac{\mathrm{e}^{-\mu(t+\Delta t)}}{\mathrm{e}^{-\mu t}}=\mathrm{e}^{-\mu\Delta t} \tag{2}$$

因此该武器单元在 t 以前射击未结束，而在 $[t, t+\Delta t)$ 内完成该次射击的条件概率为 $1-e^{-\mu\Delta t}$。

由于该 n 个武器单元的射击时间相互独立且同分布，因此 v_s 在条件 $\{n_s = i\}$ 下的条件分布为[11-12]

$$P(v_s = h \mid n_s = i, \pi_i) = \int_{0^-}^{\infty} c_{i+1}^{h}(1-e^{-\mu t})^{h} e^{-(i+1-h)\mu t}\, dA(t), \quad 0 \leqslant h \leqslant i+1, i < n \tag{3}$$

定理 在上述假设条件下，以 n_s 表示新目标 S 进入杀伤区域时武器系统所处的状态，则 $\{n_s\}$ 构成一个马尔可夫链，且其转移概率 $p_{ij} = p(n_{s+1} = j \mid n_s = i, \pi_i)$ $(i, j = 0, 1, \cdots, n)$ 表示如下：

$$p(j \mid i, \pi_i) = \begin{cases} \sum\limits_{k \in a_i} p_k \int_{0^-}^{\infty} c_{i+1}^{j}(1-e^{-\mu t})^{i+1-j} e^{-j\mu t}\, dA(t) + \\ \qquad \sum\limits_{k \notin a_i} p_k \int_{0^-}^{\infty} c_{i}^{j}(1-e^{-\mu t})^{i-j} e^{-j\mu t}\, dA(t) & j \leqslant i \leqslant n \quad (4) \\ \sum\limits_{k \in a_i} p_k \int_{0^-}^{\infty} e^{-(i+1)\mu t}\, dA(t) & j = i+1, i < n \quad (5) \\ 0 & j > i+1, i < n \quad (6) \\ \int_{0^-}^{\infty} c_{n}^{j}(1-e^{-\mu t})^{n-j} - e^{-j\mu t}\, dA(t) & j \leqslant i = n \quad (7) \end{cases}$$

证明 设 v_s 为在 (t_s, t_{s+1}) 完成射击的武器单元数，则

$$n_{s+1} = \begin{cases} n_s + 1 - v_s, & q_s < n, k_{s+1} \in \pi_{q_s} \\ n_s - v_s, & q_s < n, k_{s+1} \notin \pi_{q_s} \\ n_s - v_s, & q_s = n \end{cases} \tag{8}$$

v_s 在条件 $\{n_s = i\}$ 下的条件分布如下：

当 $n_s < n, k_{s+1} \in a_{n_s}$ 时，由引理可得

$$P(v_s = h \mid n_s = i, \pi_i) = \begin{cases} \int_{0^-}^{\infty} c_{i+1}^{h}(1-e^{-\mu t})^{h} e^{-(i+1-h)\mu t}\, dA(t), & 0 \leqslant h \leqslant i+1, i < n \\ 0, & h < 0 \text{ 或 } h > i+1 \end{cases} \tag{9}$$

当 $n_s < n, k_{s+1} \notin \pi_{n_s}$ 时，

$$P(v_s = h \mid n_s = i, \pi_i) = \begin{cases} \int_{0^-}^{\infty} c_{i}^{h}(1-e^{-\mu t})^{h} e^{-(i-h)\mu t}\, dA(t), & 0 \leqslant h \leqslant i < n \\ 0, & h < 0 \text{ 或 } h > i \end{cases} \tag{10}$$

当 $n_s = n$，即 $i = n$ 时，

$$P(v_s = h \mid n_s = i, \pi_i) = \int_{0^-}^{\infty} c_{n}^{h}(1-e^{-\mu t})^{h} e^{-(n-h)\mu t}\, dA(t), \quad 0 \leqslant h \leqslant i = n \tag{11}$$

由式(9)～(11)，可知 v_s 的条件分布概率只与 n_s 的状态有关，与 $n_0, n_1, \cdots, n_{s-1}$ 无关。因此由式(8)可知 $\{n_s\}$ 构成一马尔可夫链。

当 $i = n$ 时，由(8)及(11)式可得：

$$\begin{aligned} p_{ij} &= p(n_{s+1} = j \mid n_s = i, \pi_i) \\ &= p(n_s - v_s = j \mid n_s = i, \pi_i) \\ &= p(v_s = n - j \mid n_s = i, \pi_i) \\ &= \int_{0^-}^{\infty} c_{n}^{j}(1-e^{-\mu t})^{n-j} e^{-j\mu t}\, dA(t), \quad 0 \leqslant j \leqslant i = n \end{aligned} \tag{12}$$

即式(7)得证。类似地，由式(8)～(10)可证式(4)～(6)式成立。证毕。

2.2 状态分析

由上述定理可知，该马尔可夫决策过程的状态转移概率形式如下：

$$p(\pi)=\begin{bmatrix} p_{00}(\pi_0) & p_{01}(\pi_0) & 0 & \cdots & 0 & 0 \\ p_{10}(\pi_1) & p_{11}(\pi_1) & p_{12}(\pi_1) & \cdots & 0 & 0 \\ \cdots & & \cdots & & \cdots & \\ p_{n-2,0}(\pi_{n-2}) & p_{n-2,1}(\pi_{n-2}) & p_{n-2,2}(\pi_{n-2}) & \cdots & p_{n-2,n-1}(\pi_{n-2}) & 0 \\ p_{n-1,0}(\pi_{n-1}) & p_{n-1,1}(\pi_{n-1}) & p_{n-1,2}(\pi_{n-1}) & \cdots & p_{n-1,n-1}(\pi_{n-1}) & p_{n-1,n}(\pi_{n-1}) \\ p_{n,0}(\pi_n) & p_{n,1}(\pi_n) & p_{n,2}(\pi_n) & \cdots & p_{n,n-1}(\pi_n) & p_{n,n}(\pi_n) \end{bmatrix} \quad (13)$$

可知该矩阵的零元素为右上三角形,其余元素均大于 0 且小于 1,即其各个状态是相通的,因此该矩阵是不可约的,非周期的。又由于状态是有限的,因此该马尔可夫链是遍历的,且所有状态组成一个正常返类[12]。

由上述马尔可夫链的状态分析可知,该过程具有平稳分布,即存在极限矩阵

$$\lim_{k\to\infty}[p(\pi)]^k = p^*(\pi) > 0 \quad (14)$$

所有行向量均相同,且行元素构成一分布。

3 结束语

本文对动态 WTA 问题进行了描述分析,证明了经过适当抽象,动态 WTA 问题可以用马尔可夫决策过程来描述,从而可以运用马尔可夫决策过程理论来研究动态 WTA 问题及作战效能评估等问题,比如可建立动态武器目标分配问题的马尔可夫决策过程模型来研究其中的方案或策略优化问题。此外,由于马尔可夫链的结构将会影响计算的复杂性,而动态武器目标分配问题的马尔可夫链具有一些特殊性质,在研究这类问题时,注意运用这些性质,将有助于问题解决的简化。

参 考 文 献:

[1] Jay M R, Adnin Y, Hee S H, et al. The Generalized Weapon Target Assignment Problem[A]. 10th International Command and Control Research and Technology Symposium: The Future of C2, McLean, VA, 2005: 2-11.

[2] Cai H P, Liu J X, Chen Y W. Survey of the Research on Dynamic Weapon-Target Assignment Problem [J]. Journal of System Engineering and Electronics, 2006, 3.

[3] Chesney M, et al. Dynamic Allocation of Fires and Sensors (DAFS)[A]. 72nd MORS Symposium[C], Naval Postgraduate School, Monterey, CA, June 2004.

[4] 王正元,谭跃进.坦克会战中动态武器—目标分配问题求解方法[J].国防科技大学学报, 2003, 25 (6): 56-60.

[5] 韩松臣.导弹武器系统效能分析的随机理论方法[M]. 北京:国防工业出版社, 2001: 89-102.

[6] 黄俊,武哲.作战飞机的空—地攻击效能评估[J]. 航空学报, 1999, 20(1): 69-71.

[7] Uryasev S, Pardalos P. Robust Decision Making Addressing Uncertainties in Distributions[R]. F49620-01-1-0338, 2004: 63-82.

[8] 谢文祥,薛钧义. 状态转移服从指数分布控制系统的 SPN 可靠性分析方法[J]. 西安交通大学学报, 1999, 33(2): 11-18.

[9] Gray J E. Global Threat Prioritization[R]. Performer: Naval Surface Warfare Center, Dahlgren, VA. Dahlgren Div. NSWCDD/TR-00/46, 2003: 22-67.

[10] Commander C W. Applications of Operations Research in the United States Air Force [EB/OL]. http://www.afrl.af.mil, 2005-9-15.

[11] Lawton J A. Optimal Thrust Allocation for TBM Interceptor Midcourse Guidance[R]. Naval Surface Warfare Center, Dahlgren, VA. Dahlgren Div. 1996: 40-56.

[12] 徐光辉. 随机服务系统[M]. 北京:科学出版社, 1988: 57-150.