

doi: 10.3969/j.issn.1006-1576.2011.04.009

指挥员作战决策行为的经验加权吸引力学习模型

刘航¹, 程阁²

(1. 解放军炮兵学院 教务处, 合肥 230031; 2. 解放军炮兵学院 5系, 合肥 230031)

摘要: 为更有效地进行作战决策, 提出指挥员作战决策行为的经验加权吸引力 (experience adding weight affinity, EWA) 学习理论模型。通过行为学视角来认识指挥员作战决策行为, 从指挥员的心理和行为因素来分析指挥员的决策质量, 分析指挥员作战决策行为的影响因素, 并基于 EWA 模型对指挥员作战决策行为的学习、改善。实例分析证明, 该模型可有效解决指挥员因缺乏成功历史经验而影响作战决策质量的问题。

关键词: 作战决策行为; 学习; 博弈; EWA 模型

中图分类号: O225 **文献标志码:** A

Experience Adding Weight Affinity Study Model of Commander's Combat Decision-Making Behavior

Liu Hang¹, Cheng Ge²

(1. Office of Teaching Affairs Administration, Artillery Academy of PLA, Hefei 230031, China;

2. No. 5 Department, Artillery Academy of PLA, Hefei 230031, China)

Abstract: In order to better combat decision-making, a new study model, the experience adding weight affinity of the commander's combat decision-making behavior is proposed. From the behavioral point of view, research the commander's combat decision-making behavior. Analyze the commander's combat decision-making quality from his mentality and behavior. Then, analyze the influence factors of the commander's combat decision-making behavior. At last, learn and improve commander's combat decision-making quality based on EWA model. The experiment shows the model can effectively solve the problem of bad combat decision-making quality because of commander's lack of successful experiences.

Keywords: combat decision-making behavior; study; game; EWA model

0 引言

作战决策是指挥员在对作战对手分析判断的基础上, 对作战目标、作战方法、保障计划等重要问题进行筹划和决断的一种博弈行为。长期以来, 人们一直比较重视对决策本身的研究, 但对指挥员决策时的心理素质、行为取向、思维方式的关系研究得不够, 实际上指挥员的决策行为更为重要。因此, 笔者试图从行为科学的角度研究指挥员作战决策, 并通过在经验加权吸引力学习模型中引入吸引力 $A_i(t)$ 和经验权重 $N(t)$ 来改善指挥员的决策行为。

1 指挥员作战决策行为基本知识

1.1 指挥员作战决策行为概念^[1-2]

从行为学的角度看, 指挥员作战决策行为是指指挥员的需要、动机 (为了履行指挥职责)、行为 (执行决策) 和效果 (对作战目标实现的影响) 4 个基本要素构成的有机统一体。因此, 指挥员作战决策行为可以定义为: 在作战指挥活动中, 指挥员为了

正确作出决策并实现决策目标而进行的一系列对决策结果产生实际影响的, 有目的的心理活动及其行动反应。

1.2 指挥员作战决策行为的构成因素^[3]

指挥员作战决策行为是由需要、动机、行为和效果 4 个基本要素构成。

1) 需要。在指挥员作战决策领域, 需要是指指挥员表现出的期望, 反映着指挥员的内部心理和外部环境的特定要求。在决策行为中, 需要指向指挥员的最终目标, 激励着指挥员发挥最大潜能。

2) 动机。在指挥员作战决策行为中, 动机是推动指挥员决策行为的内驱力, 动机控制着决策行为过程。指挥员决策行为的动机有 2 类: 一是内驱动力, 来自于指挥员的需要、欲望等; 二是外驱动力即由外界信息、外界激励等产生。在指挥员作战决策行为中, 动机对决策行为具有引发、导向与激励功能。

3) 行为。指挥员作战决策行为, 是指指挥员在动

收稿日期: 2010-12-20; 修回日期: 2011-02-85

作者简介: 刘航 (1972—), 男, 安徽人, 硕士, 副教授, 从事信息系统研究。

机支配下所进行的活动。按照行为学的理论，我们可以把指挥员的决策行为过程比作一个由信息系统和能量系统组合而成的综合系统。系统中，指挥员首先接受、处理、运用来自外部的信息，达成初步的作战决策方案，而后进入能量系统，通过调动作战力量，最终对作战对手施加影响。

4) 效果。作战效果主要表现在指挥员作战决策行为对作战结果的作用。作战决策行为的效果是指决策行为实施后所产生的结果和影响。在作战活动中，战场环境的复杂性和多变性要求决策被执行后必须实时进行跟踪与反馈。

2 指挥员作战决策行为 EWA 模型描述^[4-5]

2.1 EWA 基本理论及模型

EWA 模型是建立在强化学习模型和信念学习模型^[2]的混合体。其中，参与者对这 2 种类型的信息都会加以利用。EWA 突出了在强化学习和信念学习模型中隐含的 2 个变量，即吸引力 $A_i(t)$ 和经验权重 $N(t)$ 。 $A_i(t)$ 表示 t 时期后参与者对方案 s_i 的偏好， $N(t)$ 表示过去经验的等价观察数，

$$A_i(t) = \frac{\phi \cdot N(t-1) \cdot A_i(t-1) + [\delta + (1-\delta) \cdot I(s_i, s(t))] \cdot \pi(s_i, s'(t))}{N(t)} \tag{1}$$

$$N(t) = \phi(1 - \kappa)N(t-1) + 1 \tag{2}$$

这 2 个变量在每期学习经验后都会得到更新。

2.2 指挥员作战决策行为的 EWA 理论模型

通过总结战例发现，在作战决策行为中，用虚拟博弈信念和对手的历史选择信息所计算出来的期望支付与利用指挥员历史行动的实际支付所计算出的结果几乎是一样的，因此，通过对经验在指挥决策中的影响赋予一定的权重来改善指挥员的决策行为是可行的。

引用博弈论中的 EWA 模型，以期望支付来决定指挥员的初始吸引力，以“效果法则”根据每个方案应该获得的或的确获得的支付来对各方案进行相应的强化，并对前一期的吸引力和本期的“强化”进行加权平均。

在模型中，经验权重的初始值是 $N(0)$ ，并且根据 $N(t) = \phi(1 - \kappa)N(t-1) + 1$ 更新，其限制条件为 $N(t) \leq 1/[1 - \phi(1 - \kappa)]$ ，所以 $N(t)$ 是弱递增的。吸引力的初始值是 $A_i(0)$ ，根据式 (1) 更新，建立式 (3)。

$$P_i(t+1) = e^{\lambda \cdot A_i(t)} / \sum_{k=1}^m e^{\lambda \cdot A^k(t)} \tag{3}$$

并通过将吸引力代入一种 *Logit* 形式来决定概率值。式 (1) 中， I 是一个示性函数，参数 δ 表示相对于实际支付（实际支付的权重为 1）指挥员对假想的支付权重，参数 ϕ 反应了当战场环境不断变化时，由于遗忘或对旧经验的故意放弃而导致的前一期吸引力的减弱。参数 κ 控制的是吸引力的增长率。在 *Logit* 模型中，吸引力的增长率及差异决定了选择概率的分布情况。在 λ 为定值时， κ 越大，指挥员越表现出对某一方案的频繁选择。 $N(0)$ 衡量的是指挥员以“经验等价性”为单位的最初吸引力的强度，所以，当指挥员具有狄利克莱先验概率，并能根据作战模拟进程不断学习时，则其信念近似于贝叶斯式的。由于 $N(t)$ 对经验拟合的作用不是很大，因此取 $N(0)$ 的值为 1。对于初始条件 $A_i(0)$ ，一般通过选择吸引力进行初始相对频率的拟合确定。

3 实例分析^[4,6]

在战争中，通常交战双方都会派至少一名具有实战经验且对战场情况更熟悉的指挥员参加，这里称之为“老练”及经验丰富的指挥员，在双方战场博弈中有着极其重要的作用。“老练”的指挥员在更新学习模型中的参数后，能够以新的参数更新 EWA 参与者的吸引力，根据式 (4) 和式 (5) 来算出期望支付 $E_i^j(t)$ 。

$$E_i^j(t) = \sum_{k=1}^{m_i} [P_{-i}^k(\alpha, t+1)(1 - \alpha') + \alpha' P_{-i}^k(s, t+1)] \cdot \pi_i(s_i^j, s_{-i}^k) \tag{4}$$

$$P_i^j(s, t+1) = \frac{e^{\lambda \cdot E_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot E_i^k(t)}} \tag{5}$$

其中，预期支付 $E_i^j(t)$ 受选择概率 $P_{-i}^k(s, t+1)$ 决定，反过来又通过 *Logit* 反应函数决定选择概率，因此，有经验的指挥员的经验是循环决定的。

现假设指挥员带领参谋群体要在 14 个决策方案中根据敌人态势动态博弈中作出选择。根据 EWA 模型，指挥员的博弈参数估计值如表 1。

(下转第 35 页)

位误差也随着增大。这是因为节点通信半径的变化对 Amorphous 算法中的跳数影响较大, 通信半径不规则性给 Amorphous 算法带来积累的误差随着跳数的增加而增加。

3 结论

1) 通过对 3 种算法比较, MCL 算法误差最小。2) 在节点速度测试中, 速度在 0.4 ~ 1 间对误差没有太大的影响, 所以在配置速度较慢的网络时, 可以延长节点的通信时段从而降低通信代价节约能量。3) 当锚点密度大于 1 时, MCL 算法的误差小且变化幅度小, 在无线传感器网络配置时可优先选

(上接第 28 页)

表 1 指挥员博弈中的参数估计值

参数	策略博弈			
	无经验		有经验	
	老练的 EWA	EWA	老练的 EWA	EWA
ϕ	0.44	0.00	0.29	0.22
δ	0.78	0.90	0.67	0.99
κ	0.00	0.00	0.04	0.00
α	0.24	0.00	0.77	0.00
α'	0.00	0.00	0.41	0.00
d	0.16	0.13	0.15	0.11
LL (内)	-2 095.32	-2 155.09	-1 908.48	-2 128.88
LL (外)	-968.24	-992.47	-710.28	-925.09
\bar{p} (内)	0.06	0.05	0.07	0.05
\bar{p} (外)	0.07	0.07	0.13	0.09

在表 1 中, 给出了更符合实际战争中出现的“老练”指挥员模型的参数估计, 既包括无经验的指挥员, 也包括经验丰富的指挥员。在适应性 EWA 模型 ($\alpha=0$) 中, δ 和 κ 的估计值分别接近于 1 和 0, 因此, 在指挥员仅依靠一股信念作出决策时进行限制是合理的。当加入经验丰富的“老练”后, 老练参与者的估计比例在无经验和富有经验的参与者中的估计值分别是 24% 和 77%。这种由经验而导致的百分比上升, 意味着存在着“对学习的学习”, 即经验丰富的指挥员可以更好的从过去的历史经验中得到学习, 通过了解和熟悉作战对手情况, 可以改善自己的决策行为, 制定出针对对方的更科学、有效的决策。

4 结论

通过构建基于 EWA 的指挥员学习模型, 可以有效解决指挥员因缺乏成功历史经验而影响作战决策质量的问题。

参考文献:

[1] 刘树海. 作战行动学[M]. 北京: 解放军出版社, 2009.

择此密度。4) 当节点密度大于 6 时, MCL 算法的误差小且变化幅度小, 在无线传感器网络配置时可优先选择此密度。5) 当取样个数大于 50 时, MCL 算法的误差小且变化幅度小, 取样个数一般取 50。

参考文献:

[1] 孙利民, 李建中, 等. 无线传感器网络[M]. 北京: 清华大学出版社, 2005.

[2] 郑相全, 等. 无线自组网技术实用教程[M]. 北京: 清华大学出版社, 2004.

[3] 杨毅, 罗德超. 多尺度分析技术在无线传感器网络节点定位中的应用[J]. 四川兵工学报, 2010, 31(4): 93-96.

[2] 刘卫国. 作战指挥行为学[M]. 北京: 解放军出版社, 2005.

[3] 韩志明. 作战决策行为研究[M]. 北京: 国防大学出版社, 2005.

[4] (美) 科林·凯莫勒. 行为博弈—对策略互动的实验研究[M]. 北京: 中国人民大学出版社, 2006: 312-329.

[5] 梁开卷. 决策模式选择的权变性[J]. 领导科学, 1996.

[6] 李超, 刘安惠, 等. 军事行为科学[M]. 北京: 军事谊文出版社, 1998.

[7] (美) 朱·弗登伯格. 博弈学习理论[M]. 北京: 中国人民大学出版社, 2004.

(上接第 30 页)

1) 敌人不会从机场撤离, 敌人将采取坚守的行动策略;

2) 我方应该采取破坏该机场的行动, 而不是夺取该机场, 也不是坐视不管。

3 结语

数字化炮兵作战行动根植于军事行动的复杂系统当中, 具有较大的弹性和不确定性。模糊亚对策理论以量化的思维模式, 综合考虑各种因素对数字化炮兵作战行动的影响, 局势偏好排序合理, 突出己方作战行动受敌方指挥策略的影响, 解决了以往数字化炮兵作战行动多预案间的矛盾性, 客观、全面反映了炮兵作战行动的特征和实际情况, 提高了数字化炮兵作战行动预案科学性, 为做好数字化炮兵行动预案、数字化炮兵行动评估具有借鉴意义。

参考文献:

[1] 陈之宁. 模糊数学及其军事应用[M]. 北京: 海潮出版社, 2003.

[2] 袁宏伟, 张扬. 灰色局势决策理论在炮兵火力计划优选中的运用[J]. 指挥控制与仿真, 2008(6): 20-22.