

战术级兵棋实体作战行动智能决策方法

刘 满, 张宏军[†], 郝文宁, 程 恺, 王佳胤

(陆军工程大学 指挥控制工程学院, 南京 210007)

摘 要: 战术级兵棋对抗系统结构严格、对抗性强, 可以作为研究军事智能决策的一个推演平台。鉴于此, 设计一个战术级兵棋实体作战行动智能决策模型框架, 并在该框架指导下设计一款兵棋智能引擎。所设计引擎通过挖掘兵棋历史推演数据, 提取棋子历史位置概率、夺控热度、观察度等评价属性, 利用多属性综合评价软优选算法和兵棋基本规则决策出棋子的下步行动。实验结果表明, 所提出引擎具有实时性强、灵活性高、可移植性好的优点, 可以在兵棋机机对抗和人机对抗挑战赛中取得优异成绩。

关键词: 战术级兵棋; 数据挖掘; 多属性综合评价; 智能引擎

中图分类号: TP13

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0504

开放科学(资源服务)标识码(OSID):



引用格式: 刘满, 张宏军, 郝文宁, 等. 战术级兵棋实体作战行动智能决策方法[J]. 控制与决策, 2020, 35(12): 2977-2985.

Intelligent decision-making method of tactical-level wargames

LIU Man, ZHANG Hong-jun[†], HAO Wen-ning, CHENG Kai, WANG Jia-yin

(College of Command and Control Engineering, Army Engineering University of PLA, Nanjing 210007, China)

Abstract: The tactical-level wargames system has strict structure and strong confrontation, which can be used as a deductive platform for studying military intelligent decision-making. In this paper, an intelligent action decision-making framework of tactical-level wargames is designed, and under the guidance of this framework an wargames intelligent engine is desinged. The engine utilizes a data mining method to extract evaluation attributes such as the historical position probability, the control degree and observation degree from the historical data of the wargames competition, and uses the multi-attribute comprehensive evaluation soft optimization algorithm and the wargames' basic rules to decide actions. The experimental results show that the engine has the advantages of strong real-time performance, high flexibility and good portability, and has achieved excellent results in the machine-machine and human-machine challenge.

Keywords: tactical-level wargames; data mining; multi-attribute comprehensive evaluation; intelligent engine

0 引 言

兵棋是运用规则、数据和阶段描述实际或假定的态势, 对敌对双方或多方的军事行动进行模拟的统称^[1-2], 是分析战争的重要手段。借助于计算机技术的发展, 传统的严格式兵棋开始数字化发展。我国计算机兵棋技术起步较晚, 近年来成为各大军校和民间公司的研究热点, 相关研究主要集中在阐述兵棋的本质和特点^[3-4]、研究兵棋地图技术^[5-7]、棋子机动算法^[8-9]、推演规则体系^[10-11]等。随着以深度强化学习为代表的人工智能技术在围棋上对人类的超越^[12-14], 认知智能技术开始成为当前最热门的研究热点之一。目前, 国际上以《星际争霸》这类大型策略类游戏为平台, 验证相关算法的性能, 从而实现通

用人工智能的目标。鉴于处理非完全信息、多智能体联合决策、超大状态空间上的难点, 深度强化学习等相关技术在处理大型策略类游戏上还未取得实质性突破^[15-17]。在军事应用上, 兵棋可以作为研究智能决策的推演平台, 但兵棋智能决策同样面临非完全信息、多智能体联合决策、超大状态空间上的难点。胡晓峰等^[18-19]从复杂系统的角度, 分析了人工智能技术特别是深度学习技术运用在兵棋系统上面临的机遇与挑战, 相关研究刚刚起步, 但“智能化作战”必然成为未来作战的全新样式。从兵棋数据中挖掘有用信息或提取有效特征用于兵棋智能化推演是行之有效的思路^[20-22], 文献[20]提出了一种栈式稀疏降噪自编码深度学习网络模型, 用于兵棋演习数据特征提

收稿日期: 2019-04-21; 修回日期: 2019-07-25.

基金项目: 国家自然科学基金项目(61806221).

[†]通讯作者. E-mail: jsnjzhj@263.net.

取,得到了较好的分类精度;文献[21]从兵棋数据的采集、处理、分析3个角度,详细介绍了兵棋数据处理的一些方法,并集成这些方法设计了兵棋推演数据分析原型系统;文献[22]为了解决基于纯粹逻辑的人工智能在信息不足、目标不明等矛盾问题上遇到的困难,提出可拓数据挖掘在兵棋推演系统中的实现方法,较好地提高了系统的鲁棒性.以上文献重在兵棋数据的挖掘、聚类、特征提取等数据处理上,以理论研究和探讨为主,在兵棋智能决策模型的构建中未能取得实质性进展.

本文分析基于数据挖掘和融合驱动的战术级兵棋实体作战行动智能决策的基本概念和模型框架,并介绍兵棋数据的数据处理方法和适用于兵棋机动终点位置选择的多属性综合评价软优选算法(multi-attribute comprehensive evaluation soft optimization, MCESO),最后结合具体属性数据和兵棋规则,设计并实现了一款兵棋智能引擎.该引擎通过数据驱动,整个程序提前生成数据表格,能够矩阵运算,运算速度快,实时性强;全程不设任何固定走子点,可以在不同地图和想定条件下运行,可移植性好;增大探索因子后,棋子走子位置将会多变,突破了一般兵棋智能引擎走子固定的缺点.

1 实体行动决策框架

1.1 基本概念

传统兵棋与象棋、围棋等棋类游戏类似,无论哪种兵棋,其基本组成要素均为棋盘(作战环境)、棋子(作战实体)和裁判规则(作战规则)3大部分^[23].计算机兵棋系统是将传统兵棋计算机化,并进行一定的扩展,可以通过“作战想定”对作战环境、作战实体和作战规则具体化,以模仿不同的作战情况.

兵棋推演的过程,实质上是红蓝双方作战实体按推演流程进行作战行动决策的过程.兵棋智能推演模块一般在传统的兵棋推演系统上单独设计,实现兵棋CGF(computer generated forces)行为建模,为训练

人员提供了一个虚拟的作战对手,通常称为兵棋智能引擎.单个作战实体(agent)行动决策的建模主要包括知识获取、行为推理、决策等,模型框架见图1.

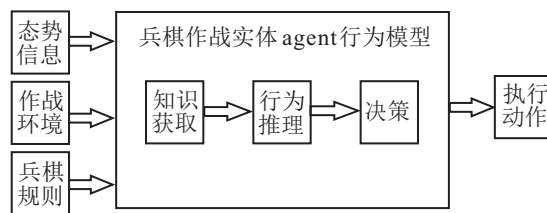


图1 兵棋作战实体 agent 行动决策模型

战术级兵棋系统是指单方兵力规模在营或连级规模的兵棋系统,作战任务一般围绕“夺控点”进行攻防战斗,作战实体可以分为坦克排、战车排、步兵排等,作战行动可以分为步兵上下车、机动、射击、夺控要点等.战术级兵棋系统实体作战行动决策是回合制多 agent 行动序列决策问题,需要解决4个层次问题:1)不同 agent 决策先后次序如何确定;2) agent 有哪些可行行动;3) agent 可行行动的次序如何确定;4)一个行动有多种方案时如何选择.

以上4个层次问题的前3个可以通过兵棋基本规则和程序流程控制简单解决,效果也相对较好.但第4个行动多方案选择问题是最难也是最重要的问题,如作战实体机动,行进的终点和路线的选择很多,这种选择往往是最难也是最关键的.本文重点解决行动多方案优选问题.

1.2 行动决策模型框架

人类在推演兵棋时首先判断双方态势,对敌情作出一定预判,基于历史经验和作战目标的推理后决策出棋子下步的行动.人类的历史经验存在于大脑之中,兵棋推演的历史经验也蕴含在历史推演的数据之中,能够为解决行动多方案优选提供决策依据.

兵棋数据按状态可以分为静态数据和动态数据,静态数据包括环境数据(地图数据)、规则数据和推演历史数据,动态数据主要是推演进程中的态势数据.针对静态数据,可以基于数据挖掘的方法提取与

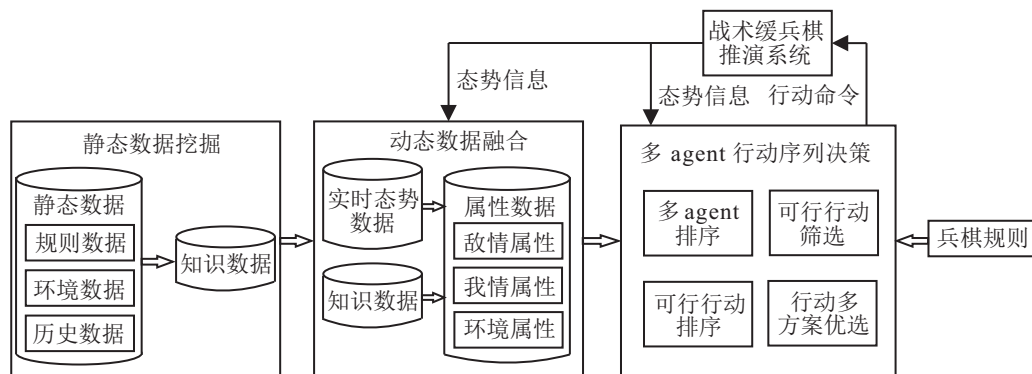


图2 兵棋实体作战行动决策模型

决策相关的数据,获取“知识”;针对动态数据(态势数据),可以基于数据融合的方法将当前态势与知识数据进行融合,生成支持决策的属性数据。

本文设计的兵棋实体作战行动智能决策模型框架如图2所示。静态数据挖掘对兵棋数据中的规则数据、环境数据和历史数据进行挖掘,提取与决策相关的静态知识数据。动态数据融合对知识数据和推演系统的实时态势进行数据融合,形成决策相关的属性数据。多agent行动序列决策综合兵棋基本规则、态势数据和属性数据,完成兵棋作战实体的行动命令生成。行动命令发送到战术级兵棋推演系统后被执行,推演系统会将新的态势数据传递出来。如此循环,实现兵棋推演系统的智能决策。

2 关键技术

2.1 多属性综合评价软优选算法

战术级兵棋实体行动一般分为步兵棋子上下车动作、机动、射击、夺控,这些动作中,机动动作最难决策,因为它与地图位置的选择密切相关,如坦克棋子一个机动阶段最多能机动六格,机动终点的选择可能有96个,一个终点至少能生成一个机动方案。

某个位置对作战实体的重要程度可以通过一些属性进行评价,如敌情属性:棋子移动的终点被敌人观察到的可能性越小越好,我方越不容易被攻击;我情属性:统计以往的走子经验,我方停留次数越多的位置越好;环境属性:消耗体力少且越接近夺控点的位置越好。机动终点的选择是典型的多属性综合评价选择问题,可以用相关理论加以解决。

多属性综合评价(multi-attribute comprehensive evaluation, MCE)根据评价对象和评价目的,从不同的侧面选取刻画系统某种特征的评价指标,建立指标体系,并通过一定的数学模型(或算法),实现系统性综合评价^[24]。加权平均法是将多个评价指标值合成一个整体性的综合评价值^[25]。针对兵棋实体机动终点的多属性综合评价,令

$$U_i = \sum_{j=1}^n \omega_j V_j(x_{ij}). \quad (1)$$

其中: U_i 为第*i*个终点位置(传统上称为第*i*个方案)的加权综合评价值, x_{ij} 为第*i*个终点的第*j*个属性值, $V_j(x_{ij})$ 为 x_{ij} 的标准化评价值, ω_j 为第*j*个属性值的权系数。

传统方案优选的标准是选择综合评价值最高的方案 $\max(U_i)$,但该优选方法不适宜军事问题选择。

“兵无常势,水无常形”,优秀的军事决策必须兼顾质量与变化。为了避免传统多属性综合评价优选

算法在策略优选上固定的缺陷,本文引入机器学习学科中软概率的优选方法^[26],提出多属性综合评价软优选算法,即将综合评价值转换为能够调控热度的概率值,并依据概率的大小选择方案。优选准则改为

$$p_i = \frac{e^{\frac{U_i}{\tau}}}{\sum_{j=1}^n e^{\frac{U_j}{\tau}}}, \quad (2)$$

$$k = \text{random_choice}(\mathbf{P}). \quad (3)$$

其中: p_i 为由评价值 U_i 转换的概率值; $\tau > 0$,称为“温度因子”,具有调控概率的作用; \mathbf{P} 为由 p_i 组成的离散概率分布数组;random_choice()为依概率选择坐标函数; k 为最后优选出的方案序号。

当 $\tau \rightarrow 0$ 时,不同方案的概率值差别较大,最终选出的方案趋向于综合评价值最大的方案;当 $\tau \rightarrow +\infty$ 时,不同方案的概率值差距较小,最终选出的方案趋向于随机选择;当 $1 \leq \tau \leq e$ 时,可以较好地兼顾质量与变化, τ 越大变化性越强。

可以看出,只要对兵棋上每个位置属性进行科学评价,利用多属性综合评价软优选算法,就能算出兵棋实体机动的终点位置,然后调用已有的路径生成算法(如限制距离后的泛洪法),即可输出完整的机动命令。

2.2 兵棋数据挖掘与融合

数据挖掘(data mining)是KDD(knowledge discovery in database)中的重要技术,可以从复杂的、多尺度的、信息丰富的海量数据中提取或挖掘出隐藏在背后的有价值的知识和信息^[27]。兵棋数据挖掘是指利用数据挖掘技术对兵棋推演系统相关的各类数据进行收集、处理,并提取或挖掘有价值的知识或信息的过程。兵棋数据按类型可以分为环境数据、规则数据、历史数据和态势数据。环境数据是指兵棋系统对战场环境的量化表示而产生的数据;规则数据是指兵棋系统对作战实体、作战规则的描述而产生的数据;历史数据是指兵棋推演系统在多次推演过程中记录下来的推演数据的总和;态势数据是指在兵棋推演过程中产生的描述战场动态情况的数据。其中环境数据、规则数据、历史数据是静态不变的,态势数据是动态时变的。

以多属性综合评价软优选算法需求的属性数据为目标,按文献[27]总结的过程模型,兵棋数据挖掘与融合可以分为5个步骤:1)数据选择;2)数据预处理;3)数据转换;4)数据挖掘;5)数据融合。数据挖掘与融合的过程见图3。最后生成多个与决策相关的属性数据,分别为敌情、我情和环境3个属性数据集。

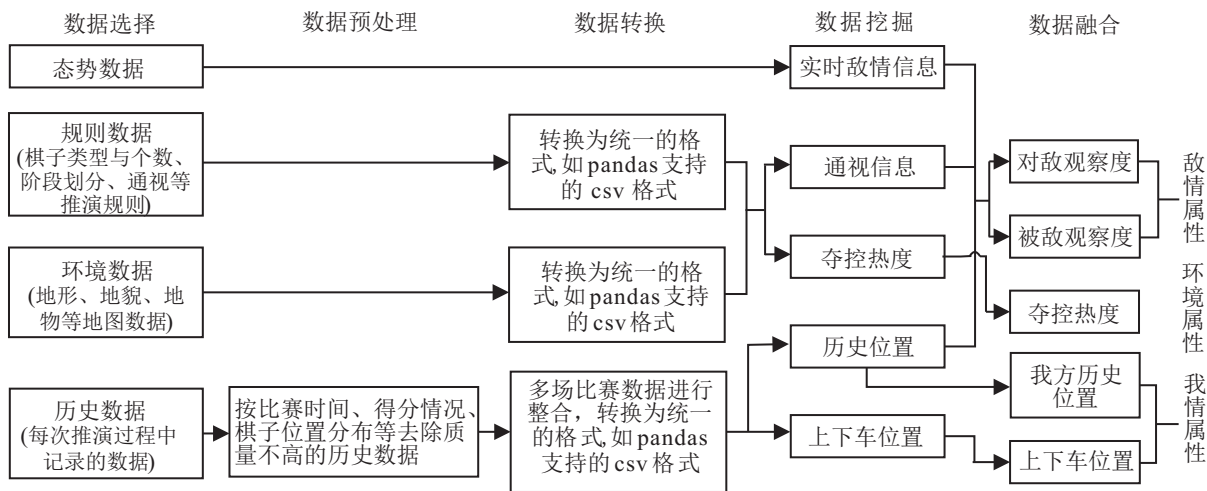


图 3 兵棋数据挖掘过程

敌情属性. 对可观察敌方棋子进行确定性分析, 对不可观察敌方棋子通过比赛历史数据进行概率分析, 融合之后形成实时敌方位置预测. 综合通视情况, 计算出对敌观察度和被敌观察度这两个评价属性.

我情属性. 通过对比赛数据中我方棋子位置进行概率分析, 形成我方棋子历史位置概率值, 作为我方下步移动位置的一个评价属性.

环境属性. 针对下步可机动到的所有位置, 按地形和当前位置与夺控点距离对我有利程度进行评价, 形成夺控热度值, 作为我方下步移动位置的一个评价属性.

3 兵棋智能引擎设计与实现

3.1 兵棋推演平台与兵棋数据介绍

按照以上兵棋实体行动模型和关键技术, 针对不同的战术级兵棋推演系统设计兵棋智能引擎. 以 2018 第 2 届全国兵棋推演大赛使用的“铁甲突击群”兵棋推演系统为推演平台^[28], 以比赛过程中记录的约 3 000 盘比赛数据 (按 7 幅地图分为 7 个作战想定) 和地图数据作为数据源, 设计并实现具体的兵棋智能引擎. 每场比赛记录的数据包含 7 张数据表格, 主要记录初始兵力状态、结束兵力状态和比赛过程中的态势等数据. 地图数据包含地图上每个六角格的编号、高程、地貌等信息, 文中“六角格”和“点”指代的对象相同.

3.2 兵棋数据挖掘和融合

3.2.1 棋子历史位置概率挖掘和上下车位置挖掘

比赛数据表信息比较零散, 必须经过筛选、整合、提取过程. 筛选是考虑比赛完整性、比赛得分等因素, 选出质量高的比赛数据; 整合是将所有比赛数据的多张信息表整合为 1 张, 主要记录的数据类别有想定 ID、比赛 ID、棋子 ID、棋子颜色、棋子阶段

初始位置、棋子阶段终止位置、比赛最终得分等要素; 提取就将筛选后的数据进行统计分析, 提取每个棋子在每个阶段终点位置的概率分布, 最终形成表格 $S[c, n, l, i]$, c 为棋子颜色索引, n 为棋子名称索引, l 为阶段索引, i 为棋子位置索引.

局部数据如表 1 所示, 第 1、2、3 列分别表示棋子颜色、名称、阶段, 第 1 行中的数字为六角格编号, 表格内容为棋子出现的概率. 表 1 既可以指导我方棋子的走子位置, 也可以预测敌方可能的走子位置.

表 1 棋子位置概率分布表 (局部)

colour	name	stage	...	50 032	50 034	50 036	...
0	10	1	...	0	0.043 10	0	...
1	10	2	...	0	0	0	...
0	10	3	...	0.017 24	0.008 62	0.025 86	...

根据比赛数据, 滤除质量不高的比赛数据后, 提取所有上车、下车的位置, 生成上车位置列表 L_{up} 和下车位置列表 L_{down} .

3.2.2 夺控热度表计算

夺取夺控点是战术级兵棋系统得分的重要因素, 棋子距离夺控点越近, 越容易完成夺控动作, 越可能赢得比赛. 夺控热度表是根据某点与夺控点距离和地形因素综合形成的与地图表格等大的数据表, 表征地图上各点位置的重要程度. 因地形对车辆 (这里特指坦克和装甲车的统称) 和士兵的体力消耗影响不同, 在生成夺控热度表时区分车辆夺控热度表和士兵夺控热度表.

车辆夺控热度表. 地图上有主要夺控点和次要夺控点, 可以计算出车辆主夺控热度表、车辆次夺控热度表和车辆主次夺控热度表. 针对地图上的某点 $P(x, y)$, 主夺控点为 $O_{first}(x_{first}, y_{first})$, 次夺控点为 $O_{second}(x_{second}, y_{second})$, 该点的 3 种控热度依次为

$$\varphi_{\text{motor}}^{\text{first}}(x, y) = \frac{\tau_{\text{common}}}{\tau(x, y)} \left(1 - \frac{L(P, O_{\text{first}})}{L_{\text{max}}^{\text{first}}} \right), \quad (4)$$

$$\varphi_{\text{motor}}^{\text{second}}(x, y) = \frac{\tau_{\text{common}}}{\tau(x, y)} \left(1 - \frac{L(P, O_{\text{second}})}{L_{\text{max}}^{\text{second}}} \right), \quad (5)$$

$$\varphi_{\text{motor}}^{\text{fir-sec}}(x, y) = \frac{1}{2}(\varphi_{\text{motor}}^{\text{first}}(x, y) + \varphi_{\text{motor}}^{\text{second}}(x, y)), \quad (6)$$

其中: τ_{common} 为车辆通过一格普通地形需要消耗的体力值, $\tau(x, y)$ 为车辆通过 $P(x, y)$ 点消耗的体力值, $L(P, O_{\text{first}})$ 和 $L(P, O_{\text{second}})$ 为 P 点到主要夺控点和次要夺控点的格数量距离, $L_{\text{max}}^{\text{first}}$ 和 $L_{\text{max}}^{\text{second}}$ 为主要夺控点和次要夺控点与边界最大格数量距离. 可以看出: 当地形因素相同时, 距夺控点越近的点其夺控热度越大; 当距离相等时, 体力消耗大的地形, 其夺控热度越小.

以上为单点的夺控热度, 计算地图上所有点的夺控热度, 即可形成车辆主夺控热度表 $\Phi_{\text{motor}}^{\text{first}}$ 、车辆次夺控热度表 $\Phi_{\text{motor}}^{\text{second}}$ 和车辆主次夺控热度表 $\Phi_{\text{motor}}^{\text{fir-sec}}$.

士兵夺控热度表. 士兵体力消耗受地形影响不大, 夺控热度的计算只考虑距离因素, 不考虑地形因素, 计算方法与车辆夺控热度表类似.

3.2.3 观察表计算

若棋子位于隐蔽地形内(如村庄、丛林等), 其观察距离不变, 但被观察距离减半, 则可通视的两点之间敌我双方的观察距离是不对等的. 以我方棋子类型 n 和位置编号 i 为行索引, 敌方棋子位置编号 j 为列索引, 若满足敌对我观察条件, 则相应交叉点为 1, 否则为 0. 遍历所有地图六角格, 计算出敌对我观察表 $C_{\text{E2O}}[n, i, j]$.

以敌方棋子类型 n 和我方棋子位置编号 i 为行索引, 敌方棋子位置编号 j 为列索引, 若满足我对敌观察条件, 则相应交叉点为 1, 否则为 0. 遍历所有地图六角格, 计算出我对敌观察表 $C_{\text{O2E}}[n, i, j]$.

3.2.4 对敌观察度和被敌观察度计算

对敌观察度是我方棋子在某个位置对敌观察程度的度量, 被敌观察度是我方棋子在某个位置被敌观察程度的度量. 我方坦克在行进的过程中一般倾向于移动到对敌观察度高的位置, 从而能够发现敌人并射击敌人, 而终点位置倾向于被敌观察度低的位置, 从而可以更好地隐藏自己, 不被敌人射击. 某点的对敌观察度与某点的观察范围和敌人可能的位置分布有关, 某点的被敌观察度与某点的被观察范围和敌人可能的位置分布有关. 车辆在点 $P(x, y)$ 的对敌观察度和被敌观察度可以表示为

$$m_{\text{motor}}(i) = \sum_{n=1}^4 \tilde{S}_{c,n,l,:} C_{\text{O2E1},i,:}^T +$$

$$\sum_{n=5}^6 \tilde{S}_{c,n,l,:} C_{\text{O2E2},i,:}^T, \quad (7)$$

$$\tilde{m}_{\text{motor}}(i) = \sum_{n=1}^6 \tilde{S}_{c,n,l,:} C_{\text{E2O1},i,:}^T. \quad (8)$$

其中: i 为 P 点的六角格编号; $\tilde{S}_{c,n,l,:}$ 为 l 阶段 c 颜色(敌方)6 个棋子(前 4 个为车辆, 后 2 个为士兵)经实时态势数据融合后形成的敌方位位置概率矩阵, 由后续 3.3 节介绍的“敌情和敌方位位置更新模块”生成; $C_{\text{O2E1},i,:}$ 为我方在 i 位置对敌车辆(标识为 1)观察矩阵, $C_{\text{O2E2},i,:}$ 为我方在 i 位置对敌士兵(标识为 2)观察矩阵; $C_{\text{E2O1},i,:}$ 为敌方对我方车辆的观察矩阵; T 为矩阵转置.

对地图上的多个点求对敌观察度和被敌观察度, 然后进行线性归一化处理即可形成对敌观察度表 M 和被敌观察度表 \tilde{M} .

3.3 敌情和敌方位位置更新模块

此模块主要完成 2 个功能. 敌我力量更新: 根据态势信息统计敌我双方兵力变化等信息作为程序决策依据; 敌方位位置更新: 将敌方历史概率表与比赛实时态势中的敌方算子位置进行数据融合, 形成敌方位位置概率矩阵 \tilde{S} .

3.4 主要动作生成规则和算法

棋子有上车、下车、夺控、射击、机动等动作, 下面介绍主要动作的生成规则和算法.

3.4.1 射击、夺控动作的产生规则

射击夺控动作产生的规则相对简单, 当棋子满足射击或夺控条件后执行该动作.

3.4.2 上车、下车动作的产生规则

上车和下车动作的难点在于选择上下车的位置. 上车的规则是在满足上车条件后, 若人和战车的位置位于上车列表 L_{up} 内, 则执行上车动作. 下车动作的规则是在满足下车条件后, 若载人战车的位置位于下车列表 L_{down} 内, 或者载人战车距离夺控点小于等于一定数值(本文设为 3), 则执行下车动作.

3.4.3 士兵移动算法

士兵移动, 参考夺控热度和历史位置概率 2 个因素. 因为士兵目标较小, 不易被敌方攻击, 且需要士兵到前沿探查敌情, 所以不考虑被敌观察度这一因素. 依据多属性综合评价软优选算法, 考虑地图上所有点的评价, 士兵移动综合评价矩阵可表示为

$$B_{\text{soldier}} = \omega_1^{\text{soldier}} F \Phi_{\text{soldier}}^{\text{fir-sec}} + \omega_2^{\text{soldier}} F \hat{S}_{c,n,l,:} \quad (9)$$

其中: F 为士兵可机动范围矩阵(可到达为 1, 不可到达为 0), $\hat{S}_{c,n,l,:}$ 为士兵历史位置概率矩阵(从历史位

置概率表 $S[c, n, l, i]$ 中抽取), $\Phi_{\text{soldier}}^{\text{fir-sec}}$ 为士兵主次夺控热度矩阵, $\omega_1^{\text{soldier}}$ 和 $\omega_2^{\text{soldier}}$ 为加权因子. 然后进行软概率优选, 有

$$P = \frac{e^{\frac{B_{\text{soldier}}}{\tau}}}{\text{sum}\left(\frac{B_{\text{soldier}}}{\tau}\right)}, \quad (10)$$

$$(x, y) = \text{random_choice}(P). \quad (11)$$

其中: $\text{sum}()$ 为矩阵所有元素求和, P 为与地图范围等大的离散概率分布, $\text{random_choice}()$ 为依概率选择序号函数, (x, y) 为最后优选出的移动终点. 最后, 调用路径生成函数, 生成棋子移动路径.

3.4.4 战车移动算法

战车移动, 参考夺控热度、历史位置概率和被敌观察度3个因素, 依据多属性综合评价软优选算法, 考虑地图上所有点的评价, 战车移动综合评价矩阵可表示为

$$B_{\text{vehicle}} = \omega_1^{\text{vehicle}} F \Phi_{\text{motor}}^{\text{fir-sec}} + \omega_2^{\text{vehicle}} F \hat{S}_{c,n,l,:} + \omega_3^{\text{vehicle}} F(1 - \tilde{M}). \quad (12)$$

其中: F 为战车可机动范围矩阵, $\Phi_{\text{motor}}^{\text{fir-sec}}$ 为车辆主次夺控热度矩阵, $\hat{S}_{c,n,l,:}$ 为战车历史位置概率矩阵(从历史位置概率表 $S[c, n, l, i]$ 中抽取), \tilde{M} 为被敌观察度矩阵(与地图幅度等大, 由式(8)计算可机动范围内每点的被敌观察度, 不可到达的点填充1), $\omega_1^{\text{vehicle}}$ 、 $\omega_2^{\text{vehicle}}$ 、 $\omega_3^{\text{vehicle}}$ 为加权因子. 然后进行软概率优选(式(10)和(11)), 优选出移动终点. 最后调用路径生成函数, 生成战车机动路径.

3.4.5 坦克移动算法

坦克在机动的过程中可以进行机动射击, 控制过程相对复杂, 需分别计算观察射击点和终点, 在观察

射击点的选择上还要区分有目标观察射击和无目标观察射击.

当态势信息中没有敌情信息时, 坦克执行无目标观察射击, 策略分为以下步骤.

step 1: 按多因素综合评价软优选算法, 考虑夺控热度、历史位置概率和被敌观察度3个因素, 按照战车移动终点的计算方法计算坦克移动终点 $T_{\text{end}}^{\text{tank}}$.

step 2: 搜索坦克从起始点到终点 $T_{\text{end}}^{\text{tank}}$ 中所有可能经过的点. 按式(7)求出所有可能经过点的对敌观察度, 不能到达的对敌观察度为0, 归一化后形成对敌观察度矩阵 M , 只考虑这一个因素, 按多因素综合评价软优选算法, 求取观察射击点 $T_{\text{observe}}^{\text{tank}}$.

step 3: 调用路径生成函数, 生成坦克从起始点到观察射击点的路径, 并执行机动动作.

step 4: 更新态势信息, 发现敌人棋子则调用射击函数进行射击, 否则执行 step 5.

step 5: 调用路径生成函数, 生成坦克从观察射击点到终点的路径, 并执行机动动作.

当态势信息中有敌情时, 坦克执行有目标射击, 策略分为以下步骤.

step 1: 根据专家经验按威胁程度对敌人棋子进行排序, 形成敌情列表.

step 2: 选择射击目标. 按顺序选择敌情列表中第1个棋子作为目标, 计算坦克可移动范围内是否存在可射击点, 如不存在, 则从敌情列表中选择下一个棋子作为目标, 并判断是否存在可射击点. 如此循环直到找到可以射击的目标, 执行 step 3, 若敌情列表目标都不可射击, 则停止执行有目标射击, 开始执行无目标观察射击.

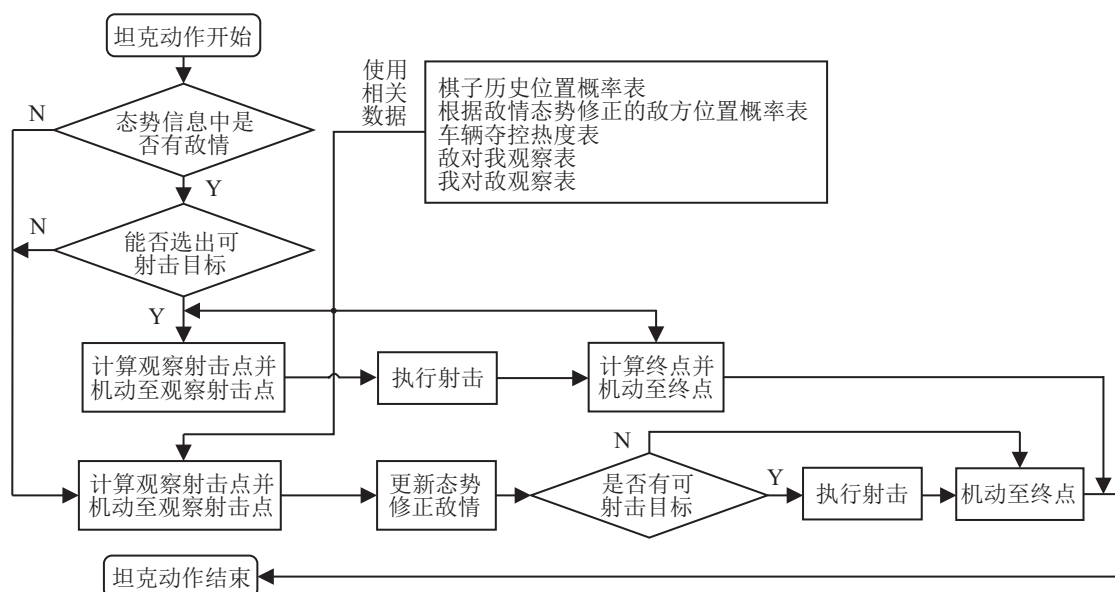


图4 坦克移动射击动作流程

step 3: 计算观察射击点. 对于选择的射击目标, 计算分别与坦克和目标距离之和最小的点集合, 从中选择离坦克距离最近的点作为观察射击点.

step 4: 调用路径生成函数, 生成到观察射击点的路径, 执行机动动作, 到达观察射击点后, 执行射击动作.

step 5: 按多因素综合评价软优选算法, 计算坦克行进终点 T_{end}^{tank} , 调用路径生成函数, 生成到终点的路径, 执行机动动作.

整个坦克移动算法综合无目标移动射击和有目标移动射击两种情况, 流程如图4所示.

3.5 兵棋智能引擎实现

兵棋智能引擎基于python语言实现, 除敌方位置

概率表外, 其他数据表都在数据预处理阶段提前生成, 敌情位置概率表在态势处理步骤中进行数据融合生成. 兵棋智能引擎的核心模块(主程序)流程如图5所示. 首先初始化对战环境, 更新态势信息并更新敌情, 生成新的敌情位置概率表并放入数据库中. 然后判断游戏是否结束, 未结束则进入主要动作判断环节. 每个动作依次查询我方所有棋子是否可以执行这个动作, 可以执行则调用相应算法模块, 生成棋子动作. 最后程序返回态势更新模块, 执行下一个循环. 针对不同agent(作战实体)优先级排序使用经验法, 按照士兵、战车 and 坦克的顺序进行; 针对可行动作优先级排序也使用经验法, 按照夺控、上车、下车、射击、机动的顺序进行.

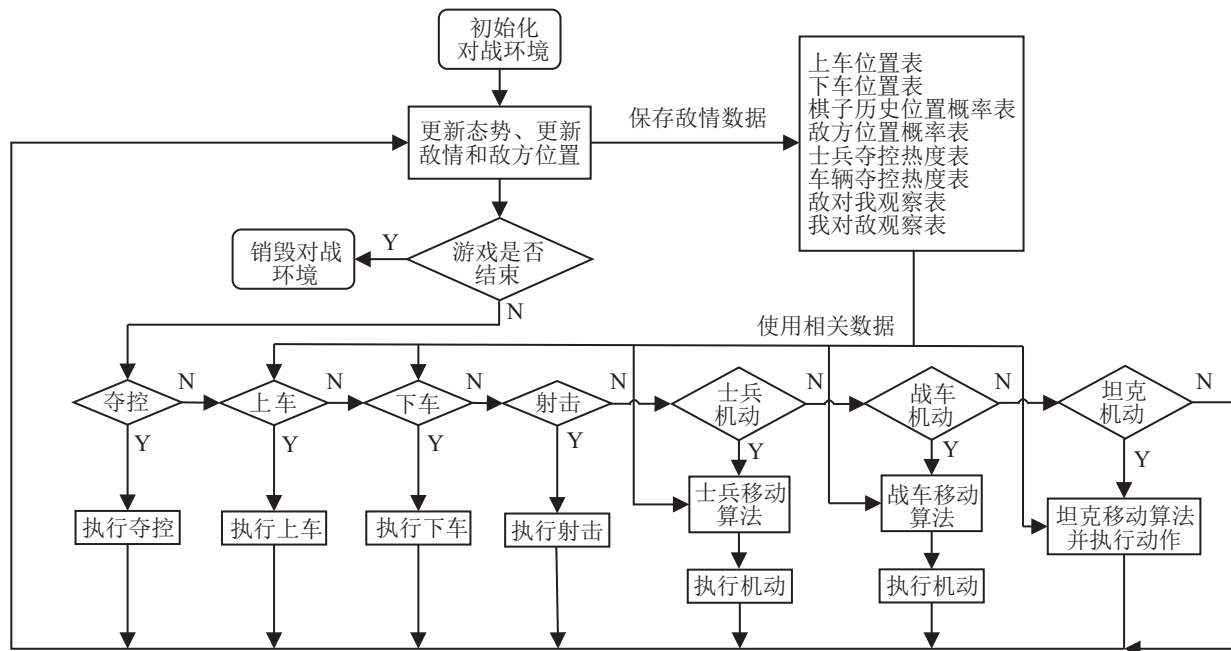


图5 兵棋智能引擎的核心模块流程

4 兵棋智能引擎效果

4.1 实验平台简介

系统运行环境如图6所示, 分3大模块: 远程兵棋推演系统、接口平台和兵棋智能引擎.

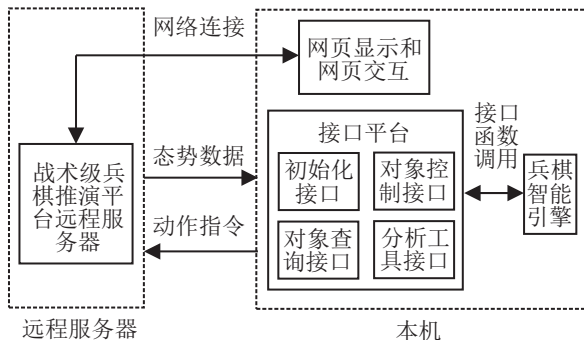


图6 兵棋智能推演系统运行环境

推演系统运行在服务器端, 通过浏览器访问观看. 系统为连级规模回合制推演系统, 分红蓝双方, 每方有坦克、战车、步兵3类6个棋子, 在特定地形上围绕2个夺控点(主次各1个)进行攻防战斗.

接口平台在兵棋推演服务器与兵棋智能引擎之间架起一座桥梁, 提供4类接口函数的调用, 分别为: 初始化接口、对象查询接口、对象控制接口、分析工具接口.

兵棋智能引擎为整个推演系统的核心, 通过判断态势给出棋子有效动作, 从而驱动整个系统有效运行.

4.2 兵棋智能引擎效果

实验中, 概率热度设为1, 这样生成的策略质量较优, 同时也兼顾了策略的多样性. 另外, 在3类夺控热

度表的选择上,可以根据不同地图不同阶段的主要进攻方向灵活进行选择. 调节参数根据不同的地图进行设定,依据比赛胜率进行手工调节,主要参数设定如表2所示.

表2 主要参数设定表

参数类型	城镇居民地	高原通道	山岳丛林地	水网稻田地	中等起伏地	岛上地形
夺控热度表	主次	次	主	次	主	主次
$\omega_1^{\text{soldier}}$	1	1	1	1	1	1
$\omega_2^{\text{soldier}}$	0.5	0.5	0.4	1	0.5	0.5
$\omega_1^{\text{vehicle}}$	1	1	1	0.2	1	1
$\omega_2^{\text{vehicle}}$	0.4	1	0.4	1	0.3	0.5
$\omega_3^{\text{vehicle}}$	0.2	0	0	0.2	0.1	0
ω_1^{tank}	0.2	0.5	0.5	0.5	0.3	0.3
ω_2^{tank}	0.2	0.2	0.2	0.5	0.3	0.3
ω_3^{tank}	1	1	1	1	1	1

本兵棋智能引擎与全国著名高校、研究所等组成的31支比赛队伍进行循环机机对战,最终排名第5,总决赛中与人类6名冠军选手进行人机对战,比分约为1:5,不敌人类选手. 经赛后总结,基于兵棋数据挖掘的兵棋智能引擎不敌基于人类经验设计的兵棋智能引擎,但是优于基于深度强化学习的兵棋智能引擎,因为基于人类经验的引擎很好地总结了人类精英玩家的规则智慧,而深度强化学习目前没有解决多智能体协作和不完全信息这两个技术难点,对战效果相对较差. 但基于人类经验的兵棋智能引擎只能针对特定的想定单独设计,开发周期较长,缺乏可移植性.

5 结 论

本文提出了基于数据挖掘和多属性综合评价软优选算法的战术级兵棋实体行动智能决策模型框架,可以基于不同的数据挖掘方法和评价算法实现系统功能,具有较好的普适性和一定的扩展性. 以该框架为指导,介绍了兵棋数据挖掘融合的流程和多属性综合评价软优选算法原理,最后设计并实现了一款兵棋智能引擎.

下一步将从4个方面作深入研究:

- 1) 完善多属性综合评价软优选算法. 目前选择的评价因素较为简单,加权因子根据比赛成绩人为设定,这两个方面可以进一步探索改进.
- 2) 分类走子风格. 对比赛数据进行聚类,分类出防守型、进攻型、综合型等走子风格,对不同走子风格的数据进行知识挖掘.
- 3) 挖掘不同棋子间的关系. 不同棋子间有较强的协作关系,在数据处理中不再简单挖掘单个棋子的位置信息,而要整体考虑挖掘棋子间的关系,实现棋子间的联动.
- 4) 进一步细化规则判断. 借鉴人类棋手的经验,

将态势判断进一步细化,实现数据与规则双驱动.

参考文献(References)

[1] James F Dunnigan. The complete wargames handbook[M]. New York: William Morrow & Company, 1997: 1-2.

[2] The joint theater level simulation(JTLS) analyst's guide[Z]. 116 Lake View Parkway Suffolk, 2007.

[3] 胡晓峰, 罗批. 兵棋推演: 复杂系统管理的创新与实践[J]. 指挥控制与仿真, 2016, 38(6): 1-5.
(Hu X F, Luo P. Wargaming: Innovation and practice for complex system management[J]. Command Control & Simulation, 2016, 38(6): 1-5.)

[4] 杨光, 胡习霜. 海军兵棋演习系统研究[J]. 指挥控制与仿真, 2016, 38(4): 96-101.
(Yang, G, Hu X S. Research of navy war game exercise system[J]. Command Control & Simulation, 2016, 38(4): 96-101.)

[5] 汤奋, 武志强, 张欣, 等. 基于模板的六角格兵棋地图可视化设计研究[J]. 测绘工程, 2016, 25(4): 55-59.
(Tang F, Wu Z Q, Zhang X, et al. Research on the visualization design of hexagonal war game map based on template[J]. Engineering of Surveying and Mapping, 2016, 25(4): 55-59.)

[6] 汤奋, 武志强, 张欣, 等. 六角格兵棋地图表示方法及其设计探析[J]. 测绘与空间地理信息, 2016, 39(5): 76-78.
(Tang F, Wu Z Q, Zhang X, et al. A nalysis on the expression method and design of hexagonal war game map[J]. Geomatics & Spatial Information Technology, 2016, 39(5): 76-78.)

[7] 张欣, 张寅宝, 李坤伟. 兵棋地图概念框架研究[J]. 地理空间信息, 2017, 15(8): 5-7.
(Zhang X, Zhang Y B, Li K W. Research on the conceptual framework of war chess map[J]. Geospatial Information, 2017, 15(8): 5-7.)

[8] 张俊恒. 计算机兵棋中兵力机动路径规划研究[D]. 长沙: 国防科学技术大学研究生院, 2010.
(Zhang J H. Research on force maneuvering path planning in computer warfare[D]. Changsha: School of Graduate, National University of Defense Technology, 2010.)

[9] 张欣, 王麒麟. 计算机兵棋越野机动路径网络模型研究[J]. 军民两用技术与产品, 2017, 14: 191-192.
(Zhang X, Wang Q L. Research on network model of computer warrior off-road maneuvering path[J]. Dual Use Technologies & Products, 2017, 14: 191-192.)

[10] 任航, 王斌, 胡波. 雷达对抗兵棋棋子与规则设计[J]. 电子信息对抗技术, 2016, 31(4): 74-78.
(Ren H, Wang B, Hu B. Research on radar countermeasure wargame elements[J]. Electronic Warfare Technology,

- 2016, 31(4): 74-78.)
- [11] 吴坚, 闫科, 闫丽, 等. 基于群体共识达成法的兵棋推演规则优化方法[J]. 装甲兵工程学院学报, 2016, 30(3): 106-110.
(Wu J, Yan K, Yan L, et al. Optimization method of wargame rules based on group consensus approach[J]. Journal of Academy of Armored Force Engineering, 2016, 30(3): 106-110.)
- [12] Silver D, Huang A. Mastering the game of go with deep neural networks and tree search[J]. Nature, 2016, 529(7587): 484-489.
- [13] Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge[J]. Nature, 2017, 550(7676): 354-359.
- [14] Silver D, Hubert T, Schrittwieser J, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play[J]. Science, 2018, 362(6419): 1140-1144.
- [15] Vinyals O, Ewalds T, Bartunov S, et al. Starcraft II: A new challenge for reinforcement learning[J]. 2017, arXiv: 1708.04782, 2017.
- [16] Shao K, Zhu Y, Zhao D. Starcraft micromanagement with reinforcement learning and curriculum transfer learning[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2018.
- [17] Peng P, Wen Y, Yang Y, et al. Multiagent bidirectionally-coordinated nets: Emergence of human-level coordination in learning to play starCraft combat games[J]. 2017, arXiv: 1703.10069, 2017.
- [18] 胡晓峰, 荣明. 智能化作战研究值得关注的几个问题[J]. 指挥与控制学报, 2018, 4(3): 195-200.
(Hu X F, Rong M. Several important questions of intelligent warfare research[J]. Journal of Command and Control, 2018, 4(3): 195-200.)
- [19] 胡晓峰, 贺筱媛, 陶九阳. AlphaGo的突破与兵棋推演的挑战[J]. 科技导报, 2017, 35(21): 51-62.
(Hu X F, He X Y, Tao J Y. AlphaGo's breakthrough and challenges of wargaming[J]. Science & Technology Review, 2017, 35(21): 51-62.)
- [20] 郑书奎, 吴琳, 贺筱媛. 基于深度学习的兵棋演习数据特征提取方法研究[J]. 指挥与控制学报, 2016, 2(3): 194-201.
(Zheng S K, Wu L, He X Y. The feature extraction method of wargame exercises data based on deep learning[J]. Journal of Command and Control, 2016, 2(3): 194-201.)
- [21] 石崇林. 基于数据挖掘的兵棋推演数据分析方法研究[D]. 长沙: 国防科学技术大学研究生院, 2012.
(Shi C L. Research on data analysis method of war chess deduction based on data mining[D]. Changsha: School of Graduate, National University of Defense Technology, 2012.)
- [22] 刘长亮, 鲍传美, 包化, 等. 可拓数据挖掘在高性能兵棋推演系统中的应用[J]. 指挥信息系统与技术, 2018, 9(1): 62-67.
(Liu C L, Bao C M, Bao H, et al. Application of extensible data mining in high performance wargaming system[J]. Command Information System And Technology, 2018, 9(1): 62-67.)
- [23] 谭鑫. 基于规则的计算机兵系统技术研究[D]. 长沙: 国防科学技术大学研究生院, 2010.
(Tan X. Research on the technology of computer-based chess system based on rules[D]. Changsha: School of Graduate, National University of Defense Technology, 2010.)
- [24] 郭亚军. 综合评价理论、方法及应用[M]. 北京: 科学出版社, 2007: 6-8.
(Guo Y J. Comprehensive evaluation theory, method and application[M]. Beijing: Science Press, 2007: 6-8.)
- [25] 叶义成, 柯丽华, 黄德育. 系统综合评价技术及其应用[M]. 北京: 冶金工业出版社, 2006: 51-52.
(Ye Y C, Ke L H, Huang D Y. System comprehensive evaluation technology and its application[M]. Beijing: Metallurgical Industry Press, 2006: 51-52.)
- [26] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 375-376.
(Zhou Z H. Machine learning[M]. Beijing: Tsinghua University Press, 2016: 375-376.)
- [27] Exner F. Advances in knowledge discovery and data mining[J]. Journal of the Association for Information Science and Technology, 1998, 49(4): 386-387.
- [28] 刘禹. 征集“先知兵圣”战术级人机对抗挑战赛AI参赛团队[EB/OL]. (2018-09-04). <http://www.chinawargame.net/report/details.do?id=97>.

作者简介

刘满(1986—), 男, 博士生, 从事军事智能决策、系统仿真的研究, E-mail: daman.043@163.com;

张宏军(1963—), 男, 教授, 博士生导师, 从事数据与知识工程、计算机仿真理论等研究, E-mail: jsnjzhj@263.net;

郝文宁(1971—), 男, 教授, 博士生导师, 从事海量高维数据归约、作战效能评估等研究, E-mail: hwnbox@163.com;

程恺(1983—), 男, 副教授, 博士, 从事智能规划、效能评估等研究, E-mail: chengkai911@126.com;

王佳胤(1995—), 男, 硕士生, 从事军事训练与仿真的研究, E-mail: 1614699005@qq.com.

(责任编辑: 郑晓蕾)