

基于粗糙集与计划识别的威胁估计方法

王晓帆^{1,2} 王宝树¹ 罗 磊²

(西安电子科技大学 西安 710071)¹ (西安理工大学 西安 710048)²

摘 要 威胁估计过程中随有大量不确定性的信息,但需要对敌方计划进行快速识别。将粗糙集理论与计划识别相结合,提出一种基于粗糙集的计划识别方法。建立了粗糙集的多属性快速计划识别模型,给出了计算方法,并用实例验证了模型的快速有效性和正确性。实验结果说明,该模型可以提高威胁估计的效率,进而为提高模式匹配的效率提供一种新方法。

关键词 威胁估计,计划识别,粗糙集,值约简

中图法分类号 TP14,TP391 **文献标识码** A

Techniques for Thread Assessment Based on Rough Set Theory and Plan Recognition

WANG Xiao-fan^{1,2} WANG Bao-shu¹ LUO Lei²

(School of Computer Science and Technology, Xidian University, Xi'an 710071, China)¹

(School of Computer Science and Technology, Xi'an University of Technology, Xi'an 710048, China)²

Abstract Although there is a mass of uncertain information in process of thread assessment, plan recognition quick is very important. A new method for plan recognition based on rough set was put forward. First the model for plan recognition based multi-attribute of rough set was build and calculating methods were given, then concrete examples were demonstrated to test the validity and correctness of the models. The results of the experiment show that the model can promote the efficiency of thread assessment, and give a new technique of pattern matching quick.

Keywords Thread assessment, Plan recognition, Rough set, Value reduct

1 引言

现代高科技战争中的信息不仅具有海量性,还具有冗余、不确定性等特点,而且敌我双方的态势瞬息万变。因此要求指挥员对敌、我、中立方态势的情况做出迅速准确的判断,对战场发展态势进行预测迫在眉睫。

根据 JDL^[1]描述:威胁评估(Threat Assessment, TA)处于信息融合的第三级,其目的是量化判断敌方的兵力结构部署或武器装备系统对我方构成威胁的能力和敌方可能的行动意图。即威胁估计就是根据当前战场的环境、气候、敌我双方的战场兵力装备部署、发生的一系列事件等诸多因素综合考虑,并对敌方的下一步行动(或目的)进行预测。威胁估计就是一个感知与事件检测、当前态势评估、未来态势预测^[2]的过程。威胁估计的目的就是根据双方的态势状况,为有利于己方的态势发展部署实施下一步的行动计划。目前常用的方法有贝叶斯网络、时空推理模型、模糊推理方法、黑板模型、多代理计划识别、基于直觉模糊与贝叶斯推理等^[3-8]。

计划识别是根据代理的行为序列来推断代理所追求目标的过程,着重于对当前已发生行为的分析和抽象,因而对动态问题有很好的适应性,与威胁估计中通过观察、分析战场中军事单元的动态行为来识别其计划的要求相一致,因此,威胁估计实质上是一个计划识别过程。

2 计划识别(Plan Recognition)

战争中,敌我双方为达到各自的军事或经济目的而制定一系列相应的作战计划。整个作战计划是按照一定层次和关系组织的,如图1所示。

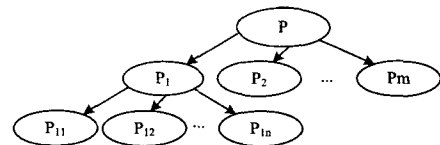


图1 作战计划分解层次图

图1中 P 表示计划,是寻求能够达到所希望目的的行为(或计划)的序列的过程^[6]。一个计划 P 可以表示为:

$$P = (G, (G_1, G_2, \dots, G_n), (C_1, C_2, \dots, C_m)) \quad (1)$$

式中, G 是执行该计划所能达到的目的的描述, $(G_1, G_2, \dots, G_n)$ 是与实现 G 相关的一系列事件(子目的)的集合, $(C_1, C_2, \dots, C_m)$ 是约束集。

计划识别记作对规划 P 的识别,即已知当前识别步,预测下一步规划元组,进而推理出该计划的目的(Goal),记作: $\Gamma(P(s_i) \rightarrow P(s_{i+1}), G(P))$, $i=0, 1, \dots, n$ 。其中, Γ 为规划识别的推理框架(方法), $G(P)$ 为当前规划 P 执行时所能达到的目的描述。

本文受国防科技预研基金(51315080202)资助。

王晓帆 男,博士生,主要研究方向为信息融合、人工智能, E-mail: wangxfok@xaut.edu.cn; 王宝树 教授,博士生导师,主要研究方向为智能信息处理、多传感器融合。

计划识别的过程即是根据领域知识,观察代理行动的时序集合,进行行动集合匹配,进而预测计划的追求目标。传统的计划识别采用的是“等”和“看”的策略,即随着信息采集的过程逐步识别并判断敌方计划。而采集的信息具有残缺性、不确定性以及延迟性等特点,这就要求计划识别在信息不完备情况下,充分利用已有信息进行识别和预测。而粗糙集是处理不确定性信息的数学工具,因此,引入基于粗糙集的计划识别方法。

3 粗糙集(Rough Set)

Rough 集理论自波兰学者 Pawlak 教授 1982 年提出以来,已经在机器学习、数据挖掘等领域中得到了较为广泛的应用^[9,10]。

定义 1 决策表信息系统 $S=\langle U, A, V, f \rangle$, 其中:

(a) $U=\{x_1, x_2, \dots, x_m\}$, $m=|U|$, 且 $m \neq 0$, U 是有限非空的对象全集;

(b) $A=C \cup D$, $|C| \neq 0$, $|D| \neq 0$ 。 A 是属性集合; C 是条件属性集合, D 是决策属性集合;

(c) 对 $\forall a \in A, f: U \rightarrow V_a$, 其中 V_a 是属性 a 的值域, f 是信息函数。

为表述方便,以下设属性集合 C 中有 n 个属性 $\{C_1, C_2, \dots, C_n\}$, 值域为有限离散集合, $|\cdot|$ 为集合的基数。不失一般性,假设仅有一个决策属性 D , 取值为 s 个: $1, 2, \dots, s$ 。对于 $B \subseteq C$, 无差别关系 $IND(B)$ 定义为 $\{(x, y) \in U^2 \mid \forall a \in B, f(x, a) = f(y, a)\}$, 通过 $IND(B)$ 将 U 划分为若干个等价类 X_i , $(1 \leq i \leq |U/IND(B)|)$; 由 D 导出的等价类为 $\{D_1, D_2, \dots, D_s\}$ 。

定义 2 设 $X \subseteq U$ 为论域 U 的一个子集, $P \subseteq C$, X 关于 P 的下近似 $P(X) = \{x \in U: [x]_P \subseteq X\}$; 其中, $[x]_P$ 表示在 U 中所有与 x 在 $IND(P)$ 下等价元素的集合。

定义 3 P, Q 为论域 U 上的两个等价关系簇, Q 的 P 正域记为 $POS_P(Q) = \bigcup_{X \in U/Q} P(X)$ 。

定义 4 P, Q 为论域 U 上的两个等价关系簇, 若 P 的 Q 独立子集 $S (S \subseteq P)$ 有 $POS_S(Q) = POS_P(Q)$, 且 $\forall a \in S, POS_{S-\{a\}}(Q) \neq POS_P(Q)$, 则称 S 为 P 的 Q 约简。

常用的属性约简有基于信息熵、差别矩阵、属性序等方法。由于威胁评估需要快速的预测趋势, 这里根据信息采集种类的时序性, 采用基于属性序^[11]的属性约简方法, 从而避免计算整个属性约简空间, 提高约简效率。

定义 5^[11] 决策表的改进树形表示。给定一个信息系统 $\langle U, C \cup D \rangle$, 其中 U 是论域, $C = \{C_1, C_2, \dots, C_n\}$ 是条件属性集合, D 是决策属性。设 $S(C) = C_1 < C_2 \dots < C_n$ (为任意给定的一个属性顺序), 则决策表的属性-值树 $T(C)$ 表示结构如图 2 (a) 所示。

$T(C)$ 共有 $n+1$ 层, C_i-j 表示第 i 层第 j 个节点。

(1) 第 $i (1 \leq i \leq n)$ 层节点赋予属性 C_i , 其各节点分支为属性 C_i 的取值, 每个节点 C_i-j 的分支(孩子)数目最大为 $|C_i|$ (属性 C_i 的取值个数), 同时含有一个指向其父节点的指针, 如图 2(b) 所示。

(2) 第 $i (1 \leq i \leq n)$ 层每个节点 C_i 都关联一个等价类 $E_i \in U/\{C_1, C_2, \dots, C_i\} (E_i \neq \emptyset)$, 且 $E_i = \bigcup E_{(i+1)p}$, 即 E_i 等于当前节点 C_i 的所有直接子节点的等价类的并, 子节点关联的等价类互不相交。

(3) 第 $n+1$ 层为决策表的等价类划分 $U/C = \{X_1, X_2, \dots, X_m\}$, 每个等价类在决策属性 D 下有单一或多个取值。

算法 1^[11] 计算决策表的属性约简

输入: 属性-值树 $T(C)$;

输出: 决策表的约简 R ;

1. $R=C, i=n$, 令 C_i 为当前属性;
2. 从第 i 层上的第一个节点 C_i-0 开始, 并令 C_i-0 为当前节点 C_i-j ;
3. 对当前节点 C_i-j 下的子树森林 $F(C_i-j)$ 进行合并, 若合并前的单值等价类 X_k 在合并后在 D 上取值增加 (相容变为不相容), 则 C_i 为必要的, $i=i-1$, 进入第 2 步;
4. 重复 3, 遍历第 i 层的所有节点, 如果合并后的所有子树下的等价类在 D 上取值都没增加, 则 C_i 为不必要的, $R=R-\{C_i\}$, 令前一个属性 C_{i-1} 为当前属性, 进入第 2 步;
5. 重复步骤 2-4, 直到 $i=0$;
6. 输出 R, R 为 D 约简。

定义 6 $S=\{C_{i1}, \dots, C_{im}\} (m \leq |C|)$, 为 P 的一个 Q 约简, 定义:

$$(C_i', v_{ij}) = \{x \mid (x \in U) \wedge (f_{C_i'}(x) = v_{ij}), i=1, \dots, m; j=1, \dots, s\}$$

用 $((C_1', v_{11}), (C_2', v_{12}), \dots, (C_m', v_{1m}))$ 表示约简 S 中的一条决策规则。令

$T_i = \{(C'_{k1}, v_{k1}), (C'_{k2}, v_{k2}), \dots, (C'_{kt}, v_{kt})\}, i=1, 2, \dots, m$ 为这样的知识范畴, 对 $\forall G_i \subseteq T_i$, 如果有:

(1) G_i 相对决策 D 是独立的;

(2) $\bigcap \{(a, v) \mid \forall (a, v) \in G_i\} \subseteq [x_i]_D, (x_i \in U)$;

则称 G_i 是 T_i 相对于决策 D 的一个值约简, 记作 $G_i \in \overline{RED}_{T_i}(D)$, 其中 $\overline{RED}_{T_i}(D)$ 表示范畴簇 T_i 的全体相对决策 D 的值约简集合。

类似于属性约简, 值约简也采取基于属性序的方法: 按照信息的易得性(或准确性)进行排序, 根据属性序进行值约简, 避免计算整个值约简空间, 从而提高约简效率。

4 基于粗糙集的多属性决策计划识别模型

计划识别过程如下:

Step1 到领域知识库中进行规则挖掘, 提取出发生的一组计划 $P_1, P_2, \dots, P_l$ 模板;

Step2 利用算法 1 对计划模板进行信息属性约简;

Step3 对约简的计划模板进行值约简, 得到简化的计划模板;

Step4 根据已采集到的信息, 进行计划模板匹配, 预测事件发生趋势。

实例: 假设在领域知识库中提取出一组敌方派出侦察机侦查我沿海军事布防情况, 我方根据获得的情报派出战斗机进行拦截, 双方态势变化的计划模板。条件为 1) 距离 C_1 (远/

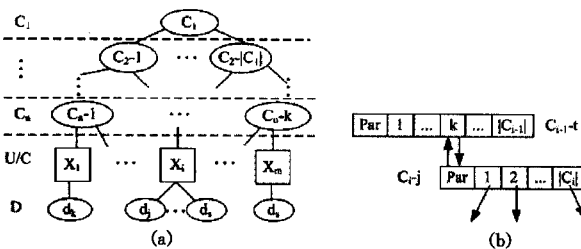


图 2 属性-值树

0,中/1,近/2);2)天气 C_2 (晴/0,阴/1,雨/2);3)释放电磁干扰 C_3 (否/0,是/1);4)释放导弹 C_4 (否/0,释放/1);5)地形 C_5 (平坦/0,崎岖/1)。决策条件 D 为敌方采取的行动:规避/0,相持/1,进攻/2。

1)在领域库中提取信息计划模板如表 1 所列。 $x_1, \dots, x_{17}$ 为知识库中的信息记录。例如: x_1 表示在我方出动战斗机拦截后,在(距离,远)+(天气,晴)+(释放电磁干扰,否)+(释放导弹,否)+(地形,平坦)条件下采取的策略为:规避。

表 1 计划模板

	C_1	C_2	C_3	C_4	C_5	D
x_1	0	0	0	0	0	0
x_2	0	1	0	0	0	0
x_3	1	1	1	0	0	1
x_4	1	2	1	1	0	2
x_5	0	1	1	0	1	0
x_6	0	1	0	0	1	0
x_7	2	2	1	0	1	0
x_8	1	1	1	1	1	2
x_9	0	2	0	0	1	0
x_{10}	2	0	0	0	1	0
x_{11}	0	2	1	0	0	1
x_{12}	0	0	1	0	1	1
x_{13}	2	2	1	1	0	2
x_{14}	2	1	1	0	1	1
x_{15}	2	0	1	0	1	1
x_{16}	0	2	1	1	1	2
x_{17}	2	0	1	1	0	2

2)由算法 1 所给的基于属性值树的属性约简方法,首先建立基于属性序列 $S=C_1 < C_2 < C_3 < C_4 < C_5$ 属性值树模型。根据算法 1 从树的底层开始,逐层判别属性是否为必要,最终得到一个关于属性序 S 的约简 C_2, C_3, C_4 ,即天气、释放电磁干扰和释放导弹为敌方采取规避、相持、攻击策略的决定因素。约简后的计划模板如表 2 所列。

表 2 属性约简后的计划模板

	C_2	C_3	C_4	D
X_1	0	0	0	0
X_2	0	1	0	1
X_3	0	1	1	2
X_4	1	1	0	1
X_5	1	1	1	2
X_6	2	0	0	0
X_7	2	1	0	0
X_8	2	1	1	2

由表 2 可见,对计划模板知识库约简后,计划模板需要匹配的信息减少了,而且模板数目也精简了,但并没有改变信息的决策能力。

3)对约简后的计划模板进行值约简:

①对决策 X_1 ,删除属性 C_2, C_4 后,决策仍是相容的,所以得到决策 $(C_3, 0) \rightarrow (D, 0)$;

②对决策 X_2 ,删除任何属性后,决策变得不相容,所以决策属性与属性值保持不变;

③对决策 X_3 ,删除属性 C_2, C_3 后,决策仍是相容的,所以得到决策 $(C_4, 1) \rightarrow (D, 2)$;

④对决策 X_4 ,删除属性 C_3 后,决策仍是相容的,所以得到决策 $(C_2, 1) + (C_4, 0) \rightarrow (D, 1)$;

⑤对决策 X_5 ,类似于 X_3 ,得到决策 $(C_4, 1) \rightarrow (D, 2)$;

⑥对决策 X_6 ,删除属性 C_2, C_4 后,决策仍是相容的,得到决策 $(C_3, 0) \rightarrow (D, 0)$;删除属性 C_3 后,得到决策 $(C_2, 2) +$

$(C_4, 0) \rightarrow (D, 0)$;

⑦对决策 X_7 ,删除属性 C_3 后,决策仍是相容的,得到决策 $(C_2, 2) + (C_4, 0) \rightarrow (D, 0)$;

⑧对决策 X_8 ,对决策 X_5 ,类似于 X_3 ,得到决策 $(C_4, 1) \rightarrow (D, 2)$;

合并相同选项后如表 3 所列。

表 3 值约简后的计划模板

	C_2	C_3	C_4	D
v_1	—	0	—	0
v_2	—	—	1	2
v_3	1	—	0	1
v_4	0	1	0	1
v_5	2	—	0	0

即进行约简后的计划模板为:

a)(释放电磁干扰,否) $\rightarrow$ (敌方行动,规避)

b)(释放导弹,是) $\rightarrow$ (敌方行动,进攻)

c)(天气,阴)+(释放导弹,否) $\rightarrow$ (敌方行动,相持)

d)(天气,晴)+(释放电磁干扰,是)+(释放导弹,否) $\rightarrow$ (敌方行动,相持)

e)(天气,雨)+(释放导弹,否) $\rightarrow$ (敌方行动,规避)

显然,通过信息的属性约简和值约简过后,计划模板显得更加简洁精炼,因此需要计划识别采集的信息变得更少,但是识别范围与能力没有改变。譬如:我方对敌侦察机出动战斗机拦截,并且没有收到敌方电磁干扰,可以断定,敌方采取规避策略。

结束语 本文将粗糙集理论与计划识别相结合方法应用于威胁估计。建立了基于粗糙集多属性决策的计划识别模型,对已有的计划模板库进行属性约简和值约简,提炼出计划识别的关键因素和取值。实验表明,应用此模型在威胁估计中提高模板匹配效率。

参考文献

- [1] Hall D L, Llinas J. An Introduction to Multisensor Data Fusion [J]. Proceedings of IEEE, 1997, 85 (1): 6-23
- [2] Endsley R. Toward a Theory of Situation Awareness in Dynamic Systems [J]. Human Factors, 1995, 37 (1): 3264
- [3] Das S, Lawless D. Trustworthy Situation Assessment via Belief Networks [C] // Proceedings of the Fifth International Conference on Information Fusion, 2002, 1: 543-549
- [4] Rao B S, Durrant-Whyte H A. Decentralized Bayesian Algorithm for Identification of Tracked Targets [J]. IEEE Transactions on Systems, Man and Cybernetics, 1993, 23(6): 1683-1698
- [5] 雷英杰, 王宝树. 基于直觉模糊决策的战场态势评估方法 [J]. 电子学报, 2006, 34(12): 2175-2179
- [6] 李伟生, 王三民, 王宝树. 基于计划识别的态势估计方法研究 [J]. 电子与信息学报, 2006, 28(3): 532-536
- [7] 雷英杰, 王宝树, 路艳丽. 基于直觉模糊逻辑的近似推理方法 [J]. 控制与决策, 2006, 3: 306-310
- [8] 王晓帆, 王宝树. 基于贝叶斯网络和直觉模糊推理的态势估计方法 [J]. 系统工程与电子技术, 2009, 31(11): 2742-2746
- [9] Pawlak Z. Rough sets [J]. International Journal of Information and Computer Science, 1982, 11(5): 341-356
- [10] Pawlak Z. Rough set approach to multi-attribute decision analysis [J]. European Journal of Operational Research, 1994, 1: 443-459
- [11] 王晓帆, 王宝树. 基于属性值树的粗糙集约简与求核方法 [J]. 西安电子科技大学学报, 2010, 37(6): 1111-1118