

# 未知环境下多机器人协作追捕算法

李 ■<sup>1,2</sup>, 潘启树<sup>2</sup>, 周浦城<sup>3</sup>, 洪炳 ■<sup>2</sup>

(1. 东北林业大学机电工程学院, 黑龙江哈尔滨 150001; 2. 哈尔滨工业大学计算机科学与技术学院, 黑龙江哈尔滨 150001;  
3. 解放军炮兵学院信息工程系, 安徽合肥 230031)

**摘 要:** 研究了在未知环境下一种多机器人协作追捕多个移动目标方法, 将追捕过程分为环境地图创建、构建追捕团队, 及团队成员协作快速捕获目标三部分. 设计抽象传感器模型用于追捕者探索障碍物. 综合考虑与捕获目标相关的各种因素建立属性关系样本数据集, 提出了利用关联规则数据挖掘技术构建追捕团队的方法. 通过对追捕目标位置预测, 将追捕问题类比为多机器人协作路径规划问题, 在环境地图基础上求解追捕团队到达目标位置的最优路径. 仿真结果表明, 多机器人能够高效地协作捕获移动目标, 证明了所提方法在复杂动态环境下实现的可行性及有效性.

**关键词:** 多机器人; 追捕问题; 目标搜索; 数据挖掘; 关联规则; 路径规划

**中图分类号:** TP242.6 **文献标识码:** A **文章编号:** 0372-2112 (2011) 03-0567-08

## Multi-Robot Cooperative Pursuit Algorithm for in An Unknown Environment

LI Jun<sup>1,2</sup>, PAN Qi-shu<sup>2</sup>, Zhou Pu-cheng<sup>3</sup>, HONG Bing-rong<sup>2</sup>

(1. College Mechanical and Electrical Northeast Forestry University, Harbin, Heilongjiang 150001, China;

2. School of Computer Science and Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China;

3. Department of Information Engineering, Artillery Academy of PLA, Hefei, Anhui 230031, China)

**Abstract:** An approach of cooperative hunting for multiple mobile targets by multi-robot is presented in an unknown environment, which divides the pursuit process into creating the environment map, forming the pursuing teams and capturing the targets. The abstract sensor model is designed to search obstacles for hunters. The data sets of attribute relationship is built by consulting all of factors about capturing evaders, then the association rule data mining is used to build the pursuit groups. Through doping out the positions of targets, the pursuit game can be transformed into multi-robot path planning problems. The best path can be found in the environment map. The simulation results show that the mobile evaders can be captured effectively and efficiently, and prove the feasibility and validity of the given algorithm under dynamic environment.

**Key words:** multi-robot; pursuit game; target search; data mining; association rule; path planning

## 1 引言

多机器人追捕问题是多机器人领域一个重要的研究课题, 通过多个移动机器人(追捕者)之间的协作和协调完成追捕多个移动目标(逃跑者)的任务.

未来战争将是人机共存的高科技战争, 为减少人员伤亡, 人们希望采用军用机器人团队来代替人类执行诸如战场情报搜索与侦察作业, 核、生、化污染地带的清消作业等, 这对未来军用机器人的研制提出了新的要求. 多机器人追捕问题的研究目标就在于构建合理、高效、鲁棒的开放式机器人团队协作与协调机制, 通过机器人在动态复杂环境中针对当前局势的实时智能决策, 使机

器人团队正常有序、经济高效地完成人类赋予的任务.

目前将逃跑者与追捕者置于同等条件下进行研究的成果中, 单目标的追捕有如下几种典型方法: 将包含反馈控制律用于协调多移动机器人的运动, 采用队形矢量控制群体队形实现了对单目标的捕获和包围<sup>[1]</sup>; 根据追捕者与逃跑者初始位置与方向开发的 SPHERES 算法成功完成单目标追捕任务<sup>[2]</sup>; 基于强化学习方法探索多机器人的协调行为实现对单目标的追捕<sup>[3-5]</sup>; 文献[6]中提出了将地图探索和追捕行为合并为一个问题, 并在概率框架下完成了未知环境下单目标追捕任务. 多目标追捕有如下几种代表方法: Grinton<sup>[7]</sup>利用承诺与公约合作机制研究追捕者之间相容承诺的生成算法实现了多

个追捕者合作捕获多个静止猎物;Schenato<sup>[8]</sup>等人利用传感器网络特点设计了一种在未知环境下分层控制结构完成多移动目标追捕任务;李淑琴<sup>[9]</sup>等人提出一种基于动态角色的多移动目标围捕算法实现了多机器人有效地合作捕获多智能移动目标;文献<sup>[10,11]</sup>从经济学角度利用任务拍卖机制对传统的合同网协议进行了改进,实现了一种允许动态联盟的多机器人合作追捕算法。

然而,多追捕者作为一种复杂的动态机器人系统,由于其感知系统的动态不确定性、执行机构的不确定性,加上其所处环境的动态未知性,直接研究追捕联盟形成方法存在耗时长、计算代价高、联盟结果不准确等问题,而将强化学习技术应用到多机器人追捕问题中需要许多假设条件,例如环境是否符合马尔科夫特性等,这些都与实际追捕问题发生条件不能完全吻合.本文在前人基础上提出一种在未知环境下多移动机器人协作追捕多个移动目标算法,算法分为三部分:(1)基于栅格环境模型,追捕者在搜索逃跑者的过程中根据获得的观测数据对追捕区域内的障碍物的位置与分布进行估计,创建环境概率地图;(2)根据追捕者与逃跑者之间的属性关系建立样本数据集,利用关联规则数据挖掘技术找到有趣规则,以此为依据针对每个逃跑者组建追捕团队;(3)追捕团队在追捕过程中对逃跑者移动位置进行预测确定追捕者的目标位置从而将追捕问题转化成路径规划问题,基于环境地图选择追捕团队成员捕获逃跑者的最优路径。

## 2 追捕问题描述

假定有限的  $n_c$  个二维方形栅格环境  $X$  (包含障碍物) 内存在  $n$  个追捕者追捕  $m$  个逃跑者,用  $R = \{R_1, \dots, R_n\}$  表示  $n$  个追捕者的集合,  $T = \{T_1, \dots, T_m\}$  表示  $m$  个逃跑者的集合. 追捕者和逃跑者统称为智能体,且逃跑者有自己的类型  $R_{T_j}, j \in m, R_{T_j} \in \{I, II, III, IV\}$  指定至少需要几个追捕者才能将其捕获. 场地内有形状和大小任意的固定障碍物,其位置由映射  $m: \chi \rightarrow \{0, 1\}$  指定.  $\forall x \in \chi, m(x) = 1$  表示栅格  $x$  是障碍物. 如果两个栅格有一个公共边,则称这两个栅格相邻.  $x$  的相邻栅格记为  $A(x)$ , 在矩形平面上  $|A(x)| \leq 4$ . 在  $t \in T = \{1, 2, \dots\}$  时刻,追捕者的位置  $X_R(t) = (x_{R1}(t), \dots, x_{Rn}(t))$ , 逃跑者的位置  $X_T(t) = (x_{T1}(t), \dots, x_{Tm}(t))$ , 并规定任意两个智能体不能占据同一位置,且  $t$  时刻各智能体只能执行一个动作,原地不动或者移动到其相邻并未被占据栅格,即  $X_k(t+1) = A(X_k(t)) \cup \{X_k(t)\}$ ,  $k \in R \cup T$ . 追捕者与逃跑者知道追捕区域的边界,但对追捕区域内的环境信息未知. 逃跑者被捕获定义为其相邻栅格内追捕者的数量大于等于其类型值,或被包围无路可走,即成功捕获逃跑者. 捕获逃跑者  $T_j$  后追捕

团队得到报酬  $S_{T_j}$ .

## 3 多机器人协作地图创建

由于搜索区域的环境信息未知,因此追捕者在搜索与追捕逃跑者的过程中应当不断根据获得的观测数据对搜索区域内的障碍物的位置与分布进行估计,用于搜索与追捕. 此处采用一种抽象的传感器模型,用于追捕者探测目标和障碍物. 其感知模型是由两个参数构成的函数:误判率  $p \in [0, 1]$  是追捕者探测到目标(障碍物)位于某个栅格,而实际上在该栅格并没有目标(障碍物)的概率,亦即  $p = P(y(t) | m(x) = 0)$ ;漏报率  $q \in [0, 1]$  是追捕者探测到某个栅格没有目标(障碍物),而实际上在该栅格却存在目标(障碍物)的概率,亦即  $q = 1 - P(y(t) | m(x) = 1)$ . 在任意时刻  $t \in T$ ,追捕者的传感器观测值  $y(t)$  由一个三元组组成:  $y(t) = \{v(t), e(t), o(t)\}$ , 其中  $v(t)$  表示观测到的追捕者的位置,  $e(t), o(t)$  分别表示被检测的逃跑者以及障碍物所在栅格的集合. 追捕者直到  $t$  时刻的各次观测构成了一个观测序列,记为  $Y_t = \{y(1), y(2), \dots, y(t)\}$ . 因为障碍物固定不动,所以当追捕者已经具有  $t-1$  时刻的障碍物位置  $P_o(x | Y_{t-1})$ , 那么障碍物位于栅格  $x$  的概率  $P_o(x | Y_t)$  是由给定直到  $t-1$  的观测序列  $Y_{t-1} = Y_t$  和新的观测  $y(t) = y(t)$  的情况下,障碍物在  $t$  时刻位于栅格  $x$  的概率决定. 假设栅格  $x$  被障碍物占用的事件为  $H$ , 用  $\bar{H}$  表示栅格  $x$  为空闲的事件. 根据 Bayes 规则有:

$$P(H | Y_t) = \frac{P(Y_t | H)P(H)}{P(Y_t | H)P(H) + P(Y_t | \bar{H})P(\bar{H})} \quad (1)$$

由于前面已经假设各次观测相互独立, 即有  $P(Y_t | H) = P(Y_{t-1} | H)P(y(t) | H)$ ,  $P(Y_t | \bar{H}) = P(Y_{t-1} | \bar{H})P(y(t) | \bar{H})$ , 将上式代入式(1)可得  $P(H | Y_t) =$

$$\frac{P(y(t) | H)P(Y_{t-1} | H)P(H)}{P(y(t) | H)P(Y_{t-1} | H)P(H) + P(y(t) | \bar{H})P(Y_{t-1} | \bar{H})P(\bar{H})} \quad (2)$$

由于其中

$$\begin{aligned} P(Y_{t-1} | H)P(H) &= P(H | Y_{t-1})P(Y_{t-1}), P(Y_{t-1} | \bar{H})P(\bar{H}) \\ &= P(\bar{H} | Y_{t-1})P(Y_{t-1}), \end{aligned} \quad \text{因此式(2)变为}$$

$$P(H | Y_t) =$$

$$\frac{P(y(t) | H)P(H | Y_{t-1})}{P(y(t) | H)P(H | Y_{t-1}) + P(y(t) | \bar{H})P(\bar{H} | Y_{t-1})} \quad (3)$$

在  $t$  时刻,对于任意的栅格  $x \in X$ , 有下面几种可能的情况:

(1) 栅格  $x$  既属于追捕机器人在  $t$  时刻的感知区域,同时又属于追捕机器人探测到的障碍物集合. 记追捕机器人在  $t$  时刻的感知区域为  $D_R(t)$ . 若  $t-1$  时刻根

据观测序列  $Y_{t-1}$  得到栅格  $x$  处是障碍物的概率为  $P_o(x|Y_{t-1})$ , 那么对于栅格  $x \in D_R(t) \cap o(t)$ , 根据式 (3), 栅格  $x$  被障碍物占用的概率为:

$$P_o(x|Y_t) = \frac{\beta}{\beta + P(y(t)|m(x)=0)P_o(m(x)=0|Y_{t-1})} \\ = \frac{\beta}{\beta + P(y(t)|m(x)=0)(1 - P_o(x|Y_{t-1}))} \quad (4)$$

式中,

$$\beta = P(y(t)|m(x)=1)P_o(x|Y_{t-1})$$

根据传感器的感知模型, 显然有:

$$P_o(x|Y_t) = \frac{(1-q)P_o(x|Y_{t-1})}{(1-q)P_o(x|Y_{t-1}) + P(1 - P_o(x|Y_{t-1}))} \quad (5)$$

(2) 栅格  $x$  属于追捕机器人在  $t$  时刻的感知区域, 但不属于追捕机器人探测到的障碍物集合. 对于栅格  $x \in D_R(t) \setminus o(t)$ , 根据传感器的感知模型, 栅格  $x$  处是障碍物的概率应按下式进行更新:

$$P_o(x|Y_t) = \frac{\beta'}{\beta' + P(\bar{y}(t)|m(x)=0)P_o(m(x)=0|Y_{t-1})} \\ = \frac{qP_o(x|Y_{t-1})}{qP_o(x|Y_{t-1}) + (1-P)(1 - P_o(x|Y_{t-1}))} \quad (6)$$

式中,

$$\beta' = P(\bar{y}(t)|m(x)=1)P_o(x|Y_{t-1})$$

(3) 栅格  $x$  既属于追捕机器人  $t$  时刻所在的栅格, 同时属于追捕机器人探测到的障碍物集合. 此时根据传感器的感知模型, 用于探测障碍物的感知信息只有当追捕机器人与障碍物处在同一栅格位置时才是完美的, 即对于栅格  $x \in v(t) \cap o(t)$ , 有:  $P_o(x|Y_t) = 1$ .

(4) 栅格  $x$  属于追捕机器人  $t$  时刻所在的栅格, 但在追捕机器人探测到的障碍物集合内, 与情况 (3) 类似, 此时对于栅格  $x \in v(t) \setminus o(t)$ , 有:  $P_o(x|Y_t) = 0$ .

(5) 对于其它情况, 由于没有获得有关障碍物分布的有用信息, 因此障碍物分布的地图信息无法更新, 即:  $P_o(x|Y_t) = P_o(x|Y_{t-1})$ . 综合上面的几种情况, 对于栅格  $x$  的障碍物分布信息按下式更新:

$$P_o(x|Y_t) = \begin{cases} \frac{(1-q)P_o(x|Y_{t-1})}{(1-q)P_o(x|Y_{t-1}) + p(1 - P_o(x|Y_{t-1}))}, & x \in d_R(t) \cap o(t) \\ 1, & x \in v(t) \cap o(t) \\ 0, & x \in v(t) \setminus o(t) \\ \frac{qP_o(x|Y_{t-1})}{qP_o(x|Y_{t-1}) + (1-p)(1 - P_o(x|Y_{t-1}))}, & x \in d_R(t) \setminus o(t) \\ P_o(x|Y_{t-1}), & \text{其他} \end{cases} \quad (7)$$

追捕者将各自得到障碍物分布的概率信息共享形成完整的环境地图信息. 本文根据地图融合的一般过程<sup>[12]</sup>采用加权平均法综合全部追捕者已有的环境知识, 若追捕者  $R_i$  对某个栅格的观测次数越多, 说明其得到的有关障碍物的信息越可靠, 那么在融合操作时相应的权重越大. 故此权重  $\lambda_i$  按下式计算:

$$\lambda_i = \frac{\text{追捕者 } R_i \text{ 对栅格 } x \text{ 的观测次数}}{\sum_{i=1}^n \text{追捕者 } R_i \text{ 对栅格 } x \text{ 的观测次数}} \quad (8)$$

## 4 关联规则数据挖掘技术组建追捕团队

追捕区域内分布多个追捕者与逃跑者, 多个追捕者采用分区局部最大搜索策略搜索目标<sup>[13]</sup>. 当某个追捕者发现逃跑者后, 通知其它追捕者. 本文提出一种利用关联规则数据挖掘技术从多个追捕者中选择合适的追捕成员组成追捕团队的方法.

### 4.1 结合追捕问题的关联规则数据挖掘定义

关联规则 (association rule) 的概念是 R. Agrawal 等人于 1993 年提出的, 其核心思想是通过描述数据库中数据项之间所存在的潜在关系的规则, 找出满足给定支持度和置信度阈值的多个域之间的依赖关系, 即寻找给定数据集中的有趣联系. 为了便于问题的讨论, 结合追捕问题给出了关联规则挖掘问题的几个定义.

**定义 1** 追捕者  $R_i (i \leq n)$  追捕逃跑者  $T_j (j \leq m)$  的关联规则挖掘数据集记为  $D_{T_j}$ ,  $D_{T_j} = \{R_1, R_2, \dots, R_i, \dots, R_n\}$ ,  $R_i$  表示追捕机器人,  $R_i = \{k_1, k_2, \dots, k_r, \dots, k_p\} (i = 1, 2, \dots, n)$  作为一个事务;  $R_i$  中的元素  $k_r (r = 1, 2, \dots, p)$  称为项目 (Item), 表示追捕者  $R_i$  与逃跑者  $T_j$  之间相关的属性关系描述.

**定义 2** 设  $I = \{k_1, k_2, \dots, k_q\} (p \leq q)$  是  $D_{T_j}$  中全体项目组成的集合,  $I$  的任何子集  $X$  称为  $D_{T_j}$  中的项目集 (itemset), 关联规则是形如  $X \Rightarrow Y$  的蕴涵式, 其中  $X \subset I$ ,  $Y \subset I$ , 且  $X \cap Y = \emptyset$ .  $X$  称为前项,  $Y$  称为后项; 关联规则反映  $X$  中的项目出现时,  $Y$  中的项目也跟着出现的规律. 追捕问题中, 有趣的关联规则为选择某一追捕者追捕特定目标提供可靠依据.

**定义 3** 数据集  $D_{T_j}$  中包含项目集  $X$  的事务数称为项目集  $X$  的支持数, 记为  $\sigma_x$ . 项目集  $X$  的支持率, 记作:  $\text{support}(X)$ ,

$$\text{support}(X) = \frac{\sigma_x}{|D_{T_j}|} \times 100\% \quad (9)$$

其中  $|D_{T_j}|$  是数据集  $D_{T_j}$  的事务数. 若  $\text{support}(X)$  不小于用户指定的最小支持率记作:  $\text{minsupport}$ , 则称  $X$  为频繁项目集, 否则称  $X$  为非频繁项目集. 那么项目集  $(X \cup Y)$  的支持率称为关联规则  $X \Rightarrow Y$  的支持率, 即  $\text{support}(X \Rightarrow Y) = \text{support}(X \cup Y)$ , 关联规则  $X \Rightarrow Y$  的置

信度记作:  $\text{confidence}(X \Rightarrow Y)$ ,

$$\text{confidence}(X \Rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X)} \times 100\% \quad (10)$$

通常用户根据挖掘需要指定的最小置信度记为  $\text{minconfidence}$ .

支持率和置信度是描述关联规则的两个重要概念,能够比较直接地形容关联规则的性质.前者用于衡量关联规则在整个数据集中的统计重要性,后者用于衡量关联规则的可信程度.为了发现有意义的关联规则即追捕者与逃跑者之间对应属性关系的出现规律,需要给定两个阈值:最小支持率和最小置信度.一般来说,只有支持率和置信度均满足阈值条件的关联规则才可能是用户感兴趣、有用的关联规则.

## 4.2 追捕团队的挖掘过程

**过程 1** 针对逃跑者  $T_j$  建立追捕样本数据集  $D_{T_j}$ , 如表 1 所示,  $k_r$  为  $D_{T_j}$  的项目. 本文设置如下:

表 1 追捕样本数据集  $D_{T_j}$

| 事务    | $k_1$    | $k_2$    | — | $k_r$    | — | $k_q$    |
|-------|----------|----------|---|----------|---|----------|
| $R_1$ | $V_{11}$ | $V_{12}$ | — | $V_{1r}$ | — | $V_{1q}$ |
| $R_2$ | $V_{21}$ | $V_{22}$ | — | $V_{2r}$ | — | $V_{2q}$ |
|       |          |          | — |          | — |          |
| $R_i$ | $V_{i1}$ | $V_{i2}$ | — | $V_{ir}$ | — | $V_{iq}$ |
|       |          |          | — |          | — |          |
| $R_n$ | $V_{n1}$ | $V_{n2}$ | — | $V_{nr}$ | — | $V_{nq}$ |

$k_1$  表示  $R_i$  所受  $T_j$  的引力与障碍物的斥力共同作用的合力  $F_{ij}$ ,  $F_{ij} = F_{\text{引}} + F_{\text{斥}}$ . 根据文献[14]中势场函数定义,引力势函数  $U_{\text{引}}(R_i) = 1/2\xi d^m(R_i, T_j)$ ,  $\xi$  是位置增益系数,  $d(R_i, T_j)$  是  $R_i$  与  $T_j$  之间的距离,这里令  $m=2$ ,相应的引力由引力势的负梯度方向决定,即  $F_{\text{引}} = -\nabla U_{\text{引}}(P_i) = -\xi d(R_i, T_j)$ . 斥力势函数  $U_{\text{斥}}(R_i)$  的表达式形式如下:

$$U_{\text{斥}}(R_i) = \begin{cases} \frac{1}{2} \eta \left( \frac{1}{\rho(R_i, Obs)} - \frac{1}{\rho_0} \right)^2, & \rho(R_i, Obs) \leq \rho_0 \\ 0, & \rho(R_i, Obs) > \rho_0 \end{cases} \quad (11)$$

其中  $\rho(R_i, Obs)$  是  $R_i$  与障碍物  $Obs$  边缘的最短距离,  $\rho_0$  是障碍物的影响距离阈值,一般地  $\rho_0 \leq \min(d_1, d_2)$ ,  $d_1$  是各障碍物间相互距离中最小距离的一半,  $d_2$  是  $T_j$  到各障碍物的最小距离.  $\eta$  是位置增益系数. 对应的斥力为:

$$F_{\text{斥}}(R_i) = -\nabla U_{\text{斥}}(R_i) = \begin{cases} \eta \left( \frac{1}{\rho(R_i, Obs)} - \frac{1}{\rho_0} \right)^2 \frac{\nabla \rho(R_i, Obs)}{\rho^2(R_i, Obs)}, & \rho(R_i, Obs) \leq \rho_0 \\ 0, & \rho(R_i, Obs) > \rho_0 \end{cases} \quad (12)$$

$k_2$  设为  $C_{ij}$  表示其他追捕者对  $R_i$  追捕  $T_j$  的可靠性的认同,由于某种原因  $R_i$  没能完成追捕任务,可靠性降低的同时直接影响其它追捕者选择团队决策,故其更新规则分为两种情况,即:

$$C_{ij} = \begin{cases} \frac{C_s}{C_t} + \epsilon_1, & \text{成功捕获目标} \\ 1 - \frac{C_f}{C_t} - \epsilon_2, & \text{未能捕获目标} \end{cases} \quad (13)$$

其中  $C_s$  为成功捕获  $T_j$  的次数,  $C_f$  是未能捕获  $T_j$  的次数,  $C_t$  是参加追捕  $T_j$  的次数,  $\epsilon_1, \epsilon_2$  分别是更新阈值.

$k_3$  设为  $Pro_{ij}$  表示  $R_i$  能够捕获  $T_j$  的概率,此处选择最近一段时间内  $R_i$  捕获  $T_j$  的频度及时效影响来刻画其大小,其更新规则定义为:

$$Pro_{ij} = \begin{cases} \frac{T}{t-t_0} \frac{C_s}{C_a}, & t-t_0 \leq T \\ \frac{C_s}{C_a}, & t-t_0 > T \end{cases} \quad (14)$$

其中  $t$  为当前时刻,  $t_0$  为最近一次成功捕获  $T_j$  的时间,  $T$  为追捕  $T_j$  后的最长时效,  $C_a$  为  $R_i$  参加追捕不同逃跑者的次数.

$k_4$  设为  $Rev_{ij}$  表示  $R_i$  捕获  $T_j$  后的奖励 ( $Rev_{ij}$ ) 与所需花费成本 ( $Co_{ij}$ ) 的差额,其计算表达式为:  $Rev_{ij} = Rew_{ij} - Co_{ij}$ ,  $Co_{ij}$  为追捕所需的通信费用、投入的资源消耗、无法捕获目标的惩罚以及追捕者放弃目前追捕目标的代价等成本之和.

$k_5$  设为  $E_{ij}$  表示  $R_i$  对自身能否捕获  $T_j$  的能力估计,其值根据下面公式计算:  $E_{ij} = \alpha \cdot Cab_{ij} + \beta$ , 其中  $Cab_{ij}$  为  $R_i$  对自己是否有能力追捕  $T_j$  的评价值,  $\alpha, \beta$  为相应的影响系数.

$k_6$  即为  $R_i$  当前的负载状态  $R\_S_i$ , 若  $R_i$  没有参加任何追捕任务,  $R\_S_i$  取值“闲”,  $\beta > 0$ ; 若  $R_i$  正在追捕某一目标,  $R\_S_i$  取值“忙”,  $\beta < 0$ .

$R_i$  表示追捕者  $i$ ;  $V_{ir}$  是与  $R_i, k_r$  对应的值. 根据属性的设置可以看出表 1 数据值  $V_{ir}$  包含连续值和离散值,为了简化本文挖掘布尔关联规则,需要将类别属性值及量化属性值转换成布尔类型. 具体方法如下:

① 量化属性离散化.  $k_1 \sim k_5$  均是量化属性,下面对这五个量化属性加以离散化分组:

$$\begin{aligned} k_1 \text{ 分为: } & a1 \{ V_{i1} \in [0, \infty) \}, a2 \{ V_{i1} \in (-\infty, 0) \} \\ k_2 \text{ 分为: } & b1 \{ V_{i2} \in [0.9, 1) \}, b2 \{ V_{i2} \in [0.5, 0.9) \}, \\ & b3 \{ V_{i2} \in [0.1, 0.5) \} \\ k_3 \text{ 分为: } & c1 \{ V_{i3} \in [0, 0.3) \}, c2 \{ V_{i3} \in [0.3, 0.6) \}, \\ & c3 \{ V_{i3} \in [0.6, 1) \} \\ k_4 \text{ 分为: } & d1 \{ V_{i4} \in (0, 3] \}, d2 \{ V_{i4} \in (3, 9] \}, \\ & d3 \{ V_{i4} \in (9, 16] \} \end{aligned}$$

$k_5$  分为:  $e1 \{ V_{i5} \in [0.5, 1] \}, e2 \{ V_{i5} \in [0, 0.5] \}$

②类别属性转化.  $R\_S_i$  是分类属性, 须进一步化为布尔类型. 那么  $R\_S_i$  分为:  $f1 \{ V_{i6} = \text{"闲"} \}, f2 \{ V_{i6} = \text{"忙"} \}$ .

以  $R_i$  为例, 若其各项取值  $V_{i1} > 0, V_{i2} = 0.63, V_{i3} = 0.35, V_{i4} = 15, V_{i5} = 0.61, V_{i6} = \text{"闲"}$ , 经过以上预处理步骤后, 将表 1 的样本数据转化成事务数据, 如表 2 所示. 类似过程可得到表示全部追捕者状态的事务数据库.

表 2 从表 1 转化成的事务数据

| 事务    | 事务项目明细                        |
|-------|-------------------------------|
| $R_i$ | $a1 \ b2 \ c2 \ d3 \ e1 \ f1$ |

**过程 2** 找出所有频繁项集. 比较著名的方法有: Apriori 算法<sup>[15]</sup>, Eclat 算法<sup>[16]</sup> 和 FP-growth<sup>[17]</sup> 算法. 为简便起见, 本文采用传统的 Apriori 算法, 根据定义, 这些项集出现的频繁性至少和预定义的最小支持计数一样. 这里将逃跑者至少需要多少追捕者才能被捕获作为设定最小支持率依据, 以便于得到的追捕团队不至于因为规模不够而无法捕获逃跑者, 即若某一逃跑者的类型  $R_{T_j} = IV$ , 至少需要 4 个追捕者才能被捕获, 若事务数据库中共有 12 条样本记录, 则  $\text{minsupport} = 4/12 \approx 33\%$ .

**过程 3** 由频繁项目集产生关联规则. 为了保证得到规则的可信程度, 最小置信度至少应该大于全部追捕者中  $C_{ij}$  的最大值, 由于  $C_{ij}$  值随着追捕过程不断更新, 若  $C_{ij}$  值过低得到过多无趣规则反而不利于组建追捕团队, 因此此处指定最小置信度最低阈值为 0.5, 即  $\text{minconfidence} = \max(0.5, \max C_{ij})$ . 设任意频繁项集  $l$ , 其非空子集  $s$ , 若满足下式

$$\frac{\text{support}(l)}{\text{support}(s)} \geq \text{minconfidence} \quad (15)$$

则输出规则  $s \Rightarrow (l - s)$ .

**过程 4** 根据得到的关联规则判断  $R_i$  是否选择  $T_j$  作为追捕目标, 作为追捕的条件依据, 本文假设  $k_1, k_2, k_5$  为前项属性,  $k_3, k_4$  为后项属性, 且理论上应在  $\{c2, c3, d2, d3\}$  内取值以保证有效地完成任务,  $k_6$  为参考属性, 在追捕实验参数约定下, 若挖掘出有趣规则  $A: a1 \wedge b2 \wedge e1 \Rightarrow c3 \wedge d3 [35\%, 72\%]$ , 根据离散化时相应的数据映射关系, 我们可以将规则  $A$  解释为: 存在 35% 的追捕者,  $T_j$  的吸引力对其作用大于障碍物的排斥力作用,  $R_i$  自身具备追捕目标的能力并且可靠性较强, 如果进行追捕捕获  $T_j$  的几率在 0.6 以上并且获得 9~16 的净利润的可能性有 72%. 根据找到的关联规则重新扫描原始样本数据集  $D_{T_j}$ , 从中找出 35% 的符合  $a1, b2, e1$  条件的追捕者, 如上例  $12 \times 0.35 \approx 4$ , 那么这 4 个追捕者组成追捕团队进行合作追捕目标. 但若出现目标只需

要 3 个追捕者即可被捕获而符合条件的追捕者数目大于 3, 那么以不破坏系统稳定性为原则进行选择, 如果  $R\_S_i$  是“闲”则优先考虑, 若  $R\_S_i$  相同则判断  $Pro_{ij}$ 、 $Rev_{ij}$ , 令  $V_{ij} = \omega_1 \cdot Pro_{ij} + \omega_2 \cdot Rev_{ij}$ ,  $\omega_1, \omega_2$  为两者的衡量系数,  $V_{ij}$  值大者优先选择; 若出现同一追捕者同时符合不同追捕团队要求情况, 则选择置信度高的关联规则对应的追捕团队. 规则  $A$  是理想挖掘结果, 若没有挖掘出形如  $A$  的有趣规则而是挖掘出的关联规则后项属性中存在某个不属于假设范围的项, 如规则  $B: a2 \wedge b2 \wedge e1 \Rightarrow c1 \wedge d2 [40\%, 72\%]$ , 同样扫描数据集  $D_{T_j}$ , 从中找出符合规则  $B$  条件的追捕者并暂时将他们忽略, 在支持率和置信度不变情况下从剩余的追捕者 ( $12 - 12 \times 0.4 \approx 8$  个) 中重新挖掘后项属性合理的有趣规则, 若发现规则  $C: a1 \wedge b2 \wedge e2 \Rightarrow c2 \wedge d2 [40\%, 72\%]$ , 从中选择符合规则  $C$  条件的  $8 \times 0.4 \approx 3$  个追捕者组成追捕团队, 如果逃跑者类型  $R_{T_j} = IV$ , 则令  $U_{ij} = \omega_3 \cdot F_{ij} + \omega_4 \cdot C_{ij} + \omega_5 \cdot E_{ij}$ ,  $\omega_3, \omega_4, \omega_5$  为衡量系数, 从数据集  $D_{T_j}$  剩余 ( $12 - 3 = 9$  个) 追捕者中选择一个  $U_{ij}$  值最大的机器人加入该团队, 循环以上过程可以组成有效的针对  $T_j$  的合作追捕团队  $CG(T_j)$ .

## 5 多机器人协作追捕算法

### 5.1 追捕目标移动位置预测

形成追捕团队后每个团队成员协调它们的行为合作追捕目标, 追捕团队根据逃跑者的状态变化协调团队成员行为选择, 本文提出一种多机器人协作追捕目标算法. 该算法的基本过程是当追捕者搜索到逃跑者后根据上述的关联规则数据挖掘方法组成多个针对不同逃跑者的追捕团队, 团队中各成员之间相互协作追捕本队目标. 假设逃跑者也是具有智能的机器人, 没有追捕者威胁时随机移动反之进行智能躲避, 为了有效追捕运动目标, 追捕团队通过对逃跑者移动位置的预测来辅助团队决策. 追捕者  $R_i \in CG(T_j)$  的目标是尽快使自己到达  $T_j$  的相邻栅格位置, 使得  $R_i$  与  $T_j$  的距离  $d(x_{R_i}(t), x_{T_j}(t))$  最小.  $T_j$  则尽力使自己被捕获的可能最小. 根据前面的捕获定义,  $A(x_{T_j}(t))$  内每个  $R_i$  贡献的捕获可能为  $1/R_{T_j}$ . 假定  $T_j$  的感知半径  $d_{T_j}$  内追捕成员的距离贡献  $DC = \sum 1/d(x_{R_i}(t), x_{T_j}(t))$ , 阻碍贡献  $OC = 1/N_{A(x_{T_j}(t))}$ , 其中  $N_{A(x_{T_j}(t))}$  表示  $t$  时刻  $T_j$  一步可达的相邻栅格数目, 包围贡献  $EC = \sum \theta_{i, i+1}(t)/2\pi$ , 其中  $\theta_{i, i+1}(t)$  是  $t$  时刻相邻两个追捕成员与  $T_j$  连线构成的逆时针夹角.  $T_j$  被捕获概率可按下式计算:

$$Pro(R_i)_{cap(T_j)} = \begin{cases} 0 & \forall R_i, d(x_{R_i}(t), x_{T_j}(t)) > d_{T_j} \\ (\mu_1 \cdot DC + \mu_2 \cdot OC + \mu_3 \cdot EC) / R_{T_j} & \\ R_i, d(x_{R_i}(t), x_{T_j}(t)) \leq d_{T_j}, \mu_1 + \mu_2 + \mu_3 = 1 & \end{cases} \quad (16)$$

根据上文定义可知  $T_j$  的  $t+1$  时刻可能的位置为  $X_{T_j}(t+1) = A(X_{T_j}(t)) \cup \{X_{T_j}(t)\}$ , 若  $T_j$  在某一追捕者  $R_i$  的感知范围内则预测  $E_j$  将选择向  $Pro(R_i)_{cap(T_j)}$  最小的位置移动, 根据预测选择使得  $d(x_{R_i}(t+1), x_{T_j}(t+1))$  最小的栅格作为下一时刻目标位置; 反之  $R_i$  认为  $T_j$  随机移动从而选择使得  $d(x_{R_i}(t+1), x_{T_j}(t))$  最小的栅格作为下一时刻目标位置。

## 5.2 追捕团队协作追捕目标

若同一追捕团队成员  $R_i \in CG(T_j)$ 、 $R_r \in CG(T_j)$  选定了同一个栅格作为下一时刻自己的目标位置, 为避免冲突, 对其当前时刻位置与  $T_j$  位置连线距离进行比较, 若  $d(x_{R_i}(t), x_{T_j}(t)) < d(x_{R_r}(t), x_{T_j}(t))$ , 则  $R_i$  具有优先的选择权,  $R_r$  重新选择其他栅格。当不同追捕团队成员  $R_i \in CG(T_j)$ 、 $R_r \in CG(T_j)$  选定了同一个栅格作为下一时刻自己的目标位置时, 追捕目标在其感知范围内的追捕者具有优先权; 若  $R_i$ 、 $R_r$  各自追捕目标均不在其感知范围内, 令  $P = \mu_1 \cdot \Delta DC + \mu_2 \cdot \Delta OC + \mu_3 \cdot \Delta EC$  表示追捕者的移动对捕获目标的贡献变化, 那么  $P$  值大者具有优先的选择权; 若  $R_i$ 、 $R_r$  各自追捕目标均不在其感知范围内且  $d(x_{R_i}(t), x_{T_j}(t)) < d(x_{R_r}(t), x_{T_j}(t))$ , 则  $R_i$  具有优先的选择权,  $R_r$  重新选择其他栅格。确定了追捕目标点之后, 追捕问题可以类比为动态环境下多移动机器人路径规划问题。其基本过程是: 追捕者  $R_i$  将观测其他机器人的行为选择, 利用 2.1 节中得到的静态障碍物信息分布、目标移动位置创建其追捕概率地图  $M_i$ , 即为追捕区域内的每个栅格赋一个占据概率  $p_h^i$ , 将路径规划信息编码到环境空间的栅格中, 然后进行搜索, 生成追捕的最优路径。 $p_h^i$  的确定方法如下: 若栅格  $x$  由静态障碍物占据, 那么  $p_h^i = p_o(x|Y_t)$ ; 若栅格  $x$  由  $R_i$  的追捕目标  $T_j$  占据, 那么  $p_h^i = 1$ , 其上、下、左、右四个栅格位置占据概率  $p_h^i = Pro(P_i)_{cap(T_j)}$ ; 其他栅格的占据概率  $p_h^i = 0$ 。将每个追捕者生成的概率地图相对应的栅格占据概率叠加生成一个全局概率地图  $M$ , 即其栅格占据概率  $p_h = \sum \lambda_i p_h^i$ 。根据地图每个追捕者的所在栅格作为参考栅格, 从参考栅格的四个相邻栅格中选择占据概率值最小的栅格, 再将所选栅格作为新的参考栅格, 重复上述过程直到到达目标栅格。当有两个可选栅格占据概率相同时, 选择与目标位置距离最近的栅格。

## 6 仿真实验

为了测试本文所提方法的效果, 检验其可行性和有效性, 进行了如下仿真实验: 场地是  $420 \times 500$  大小的矩形区域, 将其划分成尺寸为  $10 \times 10$  的  $42 \times 50$  个栅格, 障碍物是任意形状连续实心区域, 追捕者与逃跑者均为半径为 5 的圆形躯体。追捕者与逃跑者的感知半

径  $d$  相同。追捕者与逃跑者均分配了唯一的 ID 号, 追捕者和逃避者速度相同, 每次移动一个栅格。每捕获一个逃跑者追捕团队获得的报酬为该逃跑者类型的三次方, 并由团队成员平均分配所得报酬。一旦目标被捕获, 它就从原位置消失。仿真试验中参数设置如表 3。

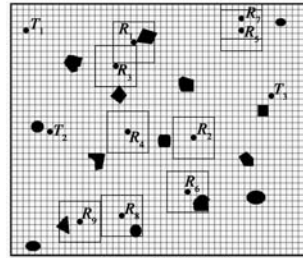


图1 初始位置

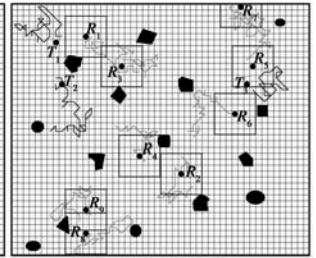


图2 发现目标  $T_3$

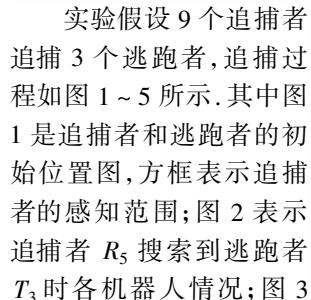


图3 对应图3的障碍物分布

给出了搜索到  $T_3$  时创建的障碍物地图快照, 其中白色区域是已搜索区域、灰色区域是未搜索区域, 黑色矩形是发现的障碍物区域; 图 4、图 5 分别给出捕获第一个及最后一个逃跑者时的快照 (为清晰起见, 两图中捕获逃跑者时追捕者及逃跑者未标出 ID)。

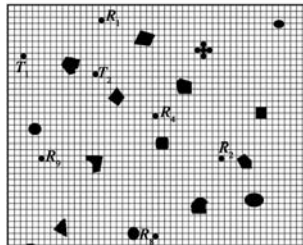


图4 捕获第一个目标示意图

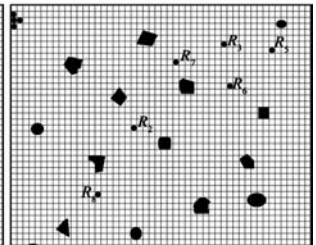


图5 捕获全部逃跑者示意图

表 3 实验控制参数

| 参数           | 取值     | 备注              | 参数      | 取值  | 备注     |
|--------------|--------|-----------------|---------|-----|--------|
| $P$          | 0.1    | 误判率             | $w_1$   | 0.5 | 衡量系数   |
| $q$          | 0.1    | 漏报率             | $w_2$   | 0.5 | 衡量系数   |
| $\xi$        | 1      | 位置增益系数          | $w_3$   | 1/3 | 衡量系数   |
| $\eta$       | 1      | 位置增益系数          | $w_4$   | 1/3 | 衡量系数   |
| $\epsilon_1$ | 0.05   | 更新阈值            | $w_5$   | 1/3 | 衡量系数   |
| $\epsilon_2$ | 0.05   | 更新阈值            | $\mu_1$ | 1/3 | 距离贡献权重 |
| $T$          | $2t_0$ | 追捕 $T_j$ 后的最长时效 | $\mu_2$ | 1/3 | 阻碍贡献权重 |
| $\alpha$     | 0.5    | $E_{ij}$ 的影响系数  | $\mu_3$ | 1/3 | 包围贡献权重 |
| $\beta$      | 0.5    | $E_{ij}$ 的影响系数  | $d$     | 10  | 感知半径   |

本文以捕获全部逃跑者的总时间作为性能指标, 对于组建追捕团队采取随机选择策略、基于改进合同网协议的选择策略<sup>[10]</sup>以及本文提出的关联规则选择策

略来测试三者对捕获结果的影响. 实验条件不变情况下, 其中追捕者个数为 9 个, 逃跑者个数分别由 1 个增加到 10 个, 类型由系统随机确定. 图 6 给出了三种选择策略方式对应的结果.

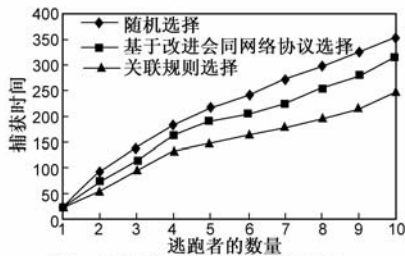


图6 不同选择策略对应的结果

可以看出, 关联规则选择策略得到的结果最好, 花费的时间最少, 而随机选择策略性能最差. 这是因为本文提出的选择策略在组建追捕团队时将追捕者的当前状态与历史经历进行综合判断后针对特定目标组成合理的追捕团队, 有助于提高捕获的效率. 基于改进合同网络协议的选择策略, 在开放复杂的多机器人环境中追捕者完全依据历史范例进行决策而导致其适应能力变差, 因而效果不如关联规则选择策略理想. 同时若采用相同的通讯机制(如基于消息), 本文从发现目标发送消息通知追捕者进入关联规则选择系统到得到所创建追捕团队结果发出开始追捕指令, 相比于基于改进合同网络协议的选择策略省却了投标及中标消息发送, 极大地减少了创建团队过程中的通信开销, 避免了资源的浪费. 此外本文方法根据目标产生动态规则, 无需存储历史案例, 从而减轻了系统的负担.

表 4 不同协作机制下追捕效益比较

| 追捕策略            | 策略一    | 策略二    |
|-----------------|--------|--------|
| 追捕团队获得平均报酬值 $S$ | 1439.5 | 1396.2 |
| 追捕者平均运行栅格数 $N$  | 187    | 216    |
| 追捕效益 $S/N$      | 7.70   | 6.46   |

实验条件不变情况下, 采用相同追捕组合分别执行本文提出的追捕策略(以下称策略一)与贪心追捕策略(以下称策略二)两组算法, 每组进行 1000 次试验, 每次试验持续 1000 个仿真周期. 考察的指标是追捕者每运行一个栅格所得到的平均报酬值, 此处称其为追捕效益. 表 4 给出了实验结果比较. 从中可以看出, 策略一的追捕效益将策略二的追捕效益提高了 19 个百分点. 这是由于追捕者本身为自利机器人, 贪心行为机制使其行为选择是为了使自己的利益获得最大, 团队成员之间没有协作而不能保证追捕团队获得利益最大, 而本文类比多机器人路径规划问题设计的追捕策略考虑了团队成员之间相互协调, 使其各自选择到达目标的最优路径的同时避免相互之间行为冲突从而获得团队最优的报酬值.

## 7 结论

本文提出一种在未知环境下多移动机器人合作追捕多个运动目标方法. 搜索逃跑目标的同时创建环境的概率地图, 发现逃跑者后综合考虑与捕获逃跑者相关的各种因素, 利用关联规则数据挖掘技术针对每个具体追捕目标组建相应既能保证追捕团队追捕能力又能避免资源浪费的追捕团队, 追捕过程中根据逃跑者采取的不同躲避方式对其移动位置进行预测, 确定追捕的目标位置, 类比多机器人路径规划问题协调各追捕者快速找到到达目标位置的最优路径捕获逃跑者. 实验结果表明所提算法追捕效率有了较大提高. 然而, 利用关联规则数据挖掘技术组建追捕团队时置信度的确定对结果有明显的影响, 置信度过高可能找不到符合的有趣规则, 无法组建追捕团队, 过低可能得到并不合适的追捕团队. 此外, 仿真环境中矩形区域相对简单, 未考虑环形区域及交叉路口等情况, 环境中没有设置动态障碍物, 这些问题在未来工作中有待于进一步研究.

## 参考文献

- [1] Yamaguchi H. A cooperative hunting behavior by mobile-robot troops[J]. International Journal of Robotics Research, 1999, 8(8): 931 - 940.
- [2] Kopparty S, Ravishankar C V. A framework for pursuit evasion games in  $R^n$ [J]. Information Processing Letters 2005, 96(3): 114 - 122.
- [3] Kok J R, Vlassis N. Sparse Cooperative Q-learning[A]. Proceedings of the 21st International Conference on Machine Learning[C]. Banff, Alberta, Canada: MIT Press, 2004. 61 - 68.
- [4] Yuko Ishiwaka, Takamasa Sato, Yukinori Kakazu. An approach to the pursuit problem on a heterogeneous multiagent system using reinforcement learning[J]. Robotics and Autonomous Systems, 2003, 3(4): 245 - 256.
- [5] 周浦城, 洪炳 ■, 黄庆成. 一种新颖的多 agent 强化学习方法[J]. 电子学报, 2006, 34(8): 1488 - 1491.  
Zhou Pu-cheng, Hong Bing-rong, Huang Qing-cheng. A novel multi-agent reinforcement learning approach[J]. Acta Electronica Sinica, 2006, 34(8): 1488 - 1491. (in Chinese)
- [6] Vidal R, Shakernia O, Kim H J, Shim D H, Sastry S. Probabilistic pursuit-evasion games: theory, implementation and experimental evaluation[J]. IEEE Transactions on Robotics and Automation, 2002, 18(5): 662 - 669.
- [7] Grinton C. A Tested for Investigating Agent Effectiveness in a Multiagent Pursuit Game[D]. Victoria, Australia: The University of Melbourne, 1996.

- [8] Luca Schenato, Songhwa Oh, Shankar Sastry, Prasanta Bose. Swarm coordination for pursuit evasion games using sensor networks[A]. Proceedings of the 2005 IEEE International Conference on Robotics and Automation[C]. Barcelona, Spain: IEEE Press, 2005. 2493 – 2498.
- [9] 李淑琴, 王欢, 李伟, 杨静宇. 基于动态角色的多移动目标围捕问题算法研究[J]. 系统仿真学报, 2006, 18(2): 362 – 365.  
Li Shu-Qin, Wang Huan, Li Wei, Yang Jing-Yu. Multiple mobile targets capturing algorithm based on dynamic role[J]. Journal of System Simulation, 2006, 18(2): 362 – 365. (in Chinese)
- [10] 周浦城, 洪炳■, 王月海. 动态环境下多机器人合作追捕研究[J]. 机器人, 2005, 27(4): 289 – 295.  
Zhou Pu-Cheng, Hong Bing-Rong, Wang Yue-hai. Multi-robot cooperative pursuit under dynamic environment [J]. Robot, 2005, 27(4): 289 – 295. (in Chinese)
- [11] Cai Ze-Su, Sun Li-Ning, Gao Hai-Bo, Zhou Pu-Cheng, Piao Song-Hao, Huang Qing-Cheng. Multi-robot cooperative pursuit eased on task bundle auctions[A]. Proceedings of the First International Conference on Intelligent Robotics and Applications[C]. Berlin, Heidelberg: Springer-Verlag, 2008. 35 – 244.
- [12] Carpin S, Birk A, Jucikas V. On map merging[J]. Robotics and Autonomous Systems, 2005, 53(1): 1 – 14.
- [13] 周浦城. 群机器人追捕目标问题的研究[D]. 哈尔滨: 哈尔滨工业大学, 2006.  
Zhou Pu-Cheng. Research on collective robots pursuit problem [D]. Harbin: Harbin Institute of Technology, 2006. (in Chinese)
- [14] Rimon E. Exact robot navigation using artificial potential function[J]. IEEE Transactions on Robotics and Automation, 1992, 8(2): 501 – 518.
- [15] Agrawal R, Srikant R. Fast algorithms for mining association

rules[A]. Proceeding of the 20th International Conference On Very Large Databases [C]. Santiago, Chile: Morgan Kaufmann, 1994. 478 – 499.

- [16] Zaki M J. Scalable algorithms for association mining[J]. IEEE Transactions on Knowledge and Data Engineering, 2000, 12 (3): 372 – 390.
- [17] Jiawei Han, Jian Pei, Yiwen Yin, Runying Mao. Mining frequent patterns without candidate generation[J]. Data Mining and Knowledge Discovery, 2004, 8(1): 53 – 87.

#### 作者简介



李 ■ 女, 1978 年生于齐齐哈尔, 现为东北林业大学机电工程学院教师, 主要研究方向为多智能体协作, 多智能体系统。  
E-mail: xlmeng@hit.edu.cn

潘启树 男, 教授, 博士生导师 1943 年出生于安徽. 分别于 1968 年和 1983 年获得哈尔滨工业大学机械制造工艺专业和管理工程学科获硕士学位, 主要研究领域为智能决策支持系统, 知识工程.



周浦城 男, 1977 年生于江西宜春, 分别于 1999 年和 2002 年获解放军炮兵学院军事指挥专业学士学位和军事运筹专业硕士学位, 现为解放军炮兵学院信息工程系教师, 主要从事多 agent 系统、强化学习、智能机器人等方面的研究.

洪炳 ■ 男, 教授, 博士生导师, 1937 年生于吉林延边, 1983 年获日本早稻田大学工学博士学位. 现为哈尔滨工业大学计算机科学与技术学院教授. 主要研究领域为: 分布式人工智能、智能机器人、多机器人系统、虚拟现实等.