

文章编号: 1001-0920(2001) 03-291-05

# 基于 RMM 和贝叶斯学习的城市交通多智能体系统

欧海涛, 张卫东, 许晓鸣  
(上海交通大学自动化系, 上海 200030)

**摘要:** 提出一种基于递归建模方法(RMM)的多智能体协调方法,使智能体在分布式环境下对其它智能体的决策建模选择合理的行动。对RMM中的信念更新采用贝叶斯学习方法,使智能体可以确定其它智能体的准确模型并实时更新信息。在城市交通控制领域建立多智能体交通控制系统,仿真结果表明了该方法的有效性,对实现智能交通系统具有重要意义。

**关键词:** 多智能体; 城市交通控制; 递归建模方法; 贝叶斯学习; ITS

中图分类号: TP 273

文献标识码: A

## Urban Traffic Multi-agent System Based on RMM and Bayesian Learning

OU Hai-tao, ZHANG Wei-dong, XU Xiao-ming

(Department of Automation, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract:** A multi-agent coordination is addressed in urban traffic control, which uses the recursive modeling method(RMM) that enables an agent to select its rational action by examining with other agents by modeling their decision making in a distributed multi-agent environment. Bayesian learning is used in conjunction with RMM for belief update. Based on this method, a multi-agent traffic control system is established and the results demonstrated its effective.

**Key words:** multi-agent; urban traffic control; recursive modeling method; Bayesian learning; ITS

## 1 引言

在城市交通控制中,对一组大量的信号进行优化控制是一个十分复杂的问题。尽管现有的研究已得到广泛应用,但对于大量剧烈变化的交通情况(如交通事故),控制效果并不理想<sup>[1]</sup>。其主要原因是:1) 交通信号数目庞大,若对这些信号都进行监督控制,就会导致海量数据传输以及中心控制计算机信息处理的滞后;2) 缺乏准确的参考模型,以至无法预测控

制的效果;3) 局部交通信号的优化有可能导致其它区域的交通恶化<sup>[2]</sup>。为克服这些困难,本文引入了多智能体的方法。在交通领域中,由于其物理拓扑结构的分布式特性,使它很适合应用多智能体方法。例如: Burmeister<sup>[3]</sup>提出的未来汽车多智能体联运系统; Claudia<sup>[4]</sup>提出用交通多智能体系统的增量相互学习方法来协调交叉口的两个控制器; Findler<sup>[5]</sup>给出了交通网络的分级结构; Weib<sup>[6]</sup>则给出多智能体在交通领域的应用。

收稿日期: 1999-10-25; 修回日期: 2000-01-19

基金项目: 国家 863/ CIMS 基金项目(9845-008)

作者简介: 欧海涛(1973—),男,四川成都人,博士生,从事智能控制、多智能体系统等研究;许晓鸣(1957—),男,上海人,教授,博士生导师,从事智能控制、预测控制等研究。

本文在城市交通控制中引入了基于递归建模方法(RMM)和贝叶斯学习的多智能体方法。RMM可使一个智能体根据其行为产生的效用大小选取合理的行动,并能在多智能体环境中通过对其它智能体决策建模来协调与它们之间的行为。这种方法只需少量甚至不需要通讯量,这正是传统交通网络控制中的一个瓶颈问题。而通过贝叶斯学习与RMM的结合,则可根据实际观测到的行为变化在线调整它对其它智能体的信念。

## 2 递归建模方法

RMM是描述一个智能体关于其它智能体信念、能力和意图的简要表达<sup>[7]</sup>,其定义如下: $P_{R_i}$ 是智能体 $R_i$ 的一个三元支付矩阵 $(R, A, U)$ 。其中, $R$ 是此环境中的智能体集合,标号为 $R_1 \sim R_n (n-1)$ ,它包括所有可能对 $R_i$ 产生影响的决策智能体; $A$ 是 $A_j$ 行动集组成的集合, $A_j = \{a_1^j, a_2^j, \dots\}$ 表示智能体 $R_j$ 所有可能的行动; $U$ 是支付函数,表示对每个可能的智能体行动组合 $U: A_1 \times A_2 \times \dots \times A_n$ 分配一系列支付,这里 $\mathcal{R}$ 是实数集合。

设 $U_{a_k^i, \dots, a_m^i, \dots, a_n^i}^{R_i}$ 表示智能体 $R_i$ 执行行动 $a_m^i$ 时其它智能体分别执行各自的行动( $R_1$ 执行 $a_k^1, \dots, R_n$ 执行 $a_n^i$ )时的支付。智能体 $R_i$ 的递归模型结构 $RMS_{R_i}$ 定义为

$$RMS_{R_i} = (P_{R_i}, RM_{R_i}) \quad (1)$$

其中, $P_{R_i}$ 是上述定义的支付矩阵, $RM_{R_i}$ 是 $R_i$ 对环境其它 $n-1$ 个智能体建模的递归模型,其定义如下。

$$RM_{R_i} = ((p_{\alpha^i}^{R_i}, M_{\{-R_i\}}^{(R_i, \alpha^i)}), \dots, (p_{\alpha^i}^{R_i}, M_{\{-R_i\}}^{(R_i, \alpha^i)}), \dots, (p_{\alpha^i}^{R_i}, M_{\{-R_i\}}^{(R_i, \alpha^i)})) \quad (2)$$

其中 $M_{\{-R_i\}}^{(R_i, \alpha^i)}$ 是 $R_i$ 关于其它智能体的 $m$ 个可能模型中的一个( $\{-R_i\}$ 表示除了智能体 $R_i$ 以外的智能体集合)。智能体 $R_i$ 对每个可能模型分配以概率 $p_{\alpha^i}^{R_i}$ 来表示其准确度,这些概率称为建模概率,其和为1,即 $\sum_{\alpha=1}^m p_{\alpha^i}^{R_i} = 1$ 。

## 3 贝叶斯学习

由于分布式环境中的动态变化特性,应在RMM方法中对智能体的信念在其观测行为的基础上不断更新。我们提出采用贝叶斯学习的方法<sup>[8]</sup>,不断更新智能体中关于其它智能体合理模型的分配概率,以实现这一动态过程。信念更新过程主要有以下

步骤:

Step1: 在该智能体的知识库中识别其它智能体的可能模型;

Step2: 对初始信念为其它智能体的模型分配先验概率(如果无先验知识则分配相同的概率);

Step3: 在一给定时间点处理观测其它智能体的行为;

Step4: 使用贝叶斯学习更新与其它智能体相关的信念,并将这些信念存储到其知识库中,返回Step3。

Step4中的贝叶斯学习过程如下:设智能体 $i$ 有一系列可能模型 $M_{ij} = \{M_{ij}^1, M_{ij}^2, \dots, M_{ij}^n\}$ ,用以预测智能体 $j (j \neq i)$ 的行为。定义 $BEL(M_{ij}^k)$ 为智能体 $i$ 中有关智能体 $j$ 的模型 $M_{ij}^k$ 的信念,可由分配给该模型的概率表示;并定义 $A_i = \{a_1^i, a_2^i, \dots, a_n^i\}$ 为智能体 $i$ 的可能行动集。模型 $M_{ij}^k$ 正确的先验概率为 $P(M_{ij}^k)$ 。假设智能体 $i$ 观测到智能体 $j$ 执行了一个行动 $a_j^l$ ,于是智能体 $i$ 更新其分配给模型 $M_{ij}^k$ 的概率为

$$BEL(M) = \tilde{P}(M_{ij}^k | a_j^l) = \frac{P(a_j^l | M_{ij}^k) P(M_{ij}^k) / P(a_j^l)}{P(a_j^l)} \quad (3)$$

上述信念更新是在对其它智能体观测行为的基础上,通过一系列相互作用不断地执行。

## 4 城市交通多智能体系统

我们基于RMM和贝叶斯学习建立了城市交通多智能体系统。智能体心智状态的组成如图1所示。智能体的框架描述了智能体的信念、能力和偏好,并用面向对象的示例对知识按概念和类别进行分级。理性智能体根据其知识做出优化决策,并在动态环境中通过信念更新相互作用来维护其知识库。

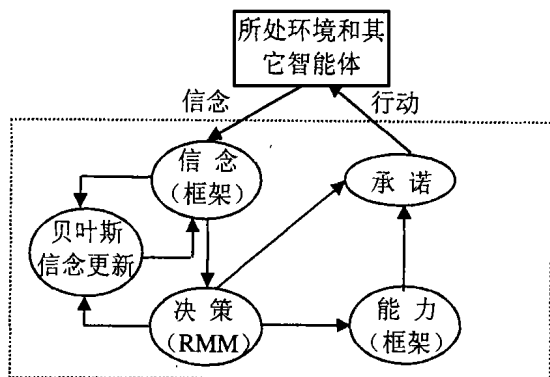


图1 智能体心智状态的组成

### 4.1 特性分析

考虑由两个交叉口组成的简单交通网络, 如图 2 所示, 图中数字表示等待车辆数。智能体 A1 和 A2 分别代表这两个交叉口控制器。行动集为  $\{a_0, a_1, a_2\}$ , 其中, 行动  $a_0$  表示东西向通行, 南北向禁止通行; 行动  $a_1$  表示东西向禁止通行, 南北向通行; 行动  $a_2$  表示东西向和南北向都禁止通行。

每个道路尽头都与外部交通相连, 车辆在各路口的到达按随机分布。道路中车辆以相同的速度  $d = 30\text{km/h}$  行驶, 车辆由静止加速到速度  $d$  的时间为  $t_{\text{acc1}} = 4\text{s}$ 。当车辆经过某一交叉口时, 它只能前向行驶或右转行驶, 即  $\{d_i\} = \{\text{Forward}, \text{Right}\}$ , 转向概率为  $\{P_{d_i}\} = \{0.5, 0.5\}$ ,  $i = 1, 2$ 。

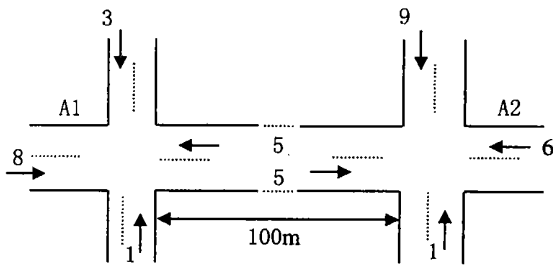


图 2 两个交叉口的简单交通网络

### 4.2 产生支付矩阵

应用面向智能体的方法, 在图 2 的场景中根据埋在交叉口下面的车辆传感器检测的车流信息进行决策。决策问题由对策论中的支付矩阵表示(参见图 3), 其元素表示每个信号周期两个智能体采取相应行动时交通网络中车辆排队等待总队长。显然, 优化控制的目标是使交通网络中的等待车辆队长最短。在每个交叉口处, 交通信号灯的每个信号控制动作都有一个最小周期, 如果太小就会引起交通混乱, 我们取为 24s。在每个周期开始时, 由 RMM 计算应选取何种行动。两个交叉口之间距离为 100m。

### 4.3 对其它智能体用递归建模结构建模

在递归建模方法中, 其它理性智能体的行动由描述其决策情形的模型确定。如果已知智能体 A2 是理性决策者, 则其决策便可按支付矩阵求解。然而, 一般情况并不知道智能体 A2 是否为理性的, 例如当 A2 处发生交通事故时, 智能体 A2 可能无法获得信息进行决策。因此, 考虑两种 A1 对 A2 建模的形式: 一种是支付矩阵的形式; 另一种是对 A2 行动的无信息(No-Info)模型。在 RMM 中, 每种模型都分配一个描述其准确程度的概率。

根据前述 RMM 建模结构, 可得对于图 2 情形的三层递归模型结构, 如图 3 所示。在递归模型结构

中, 第 1 层表示智能体 A1 观测自身的决策情况, 用 A2 的支付矩阵表示; 第 2 层表示 A1 具有 A2 决策情形的模型; 第 3 层是 A2 考虑了 A1 决策情形得到的模型对 A1 建模的影响。

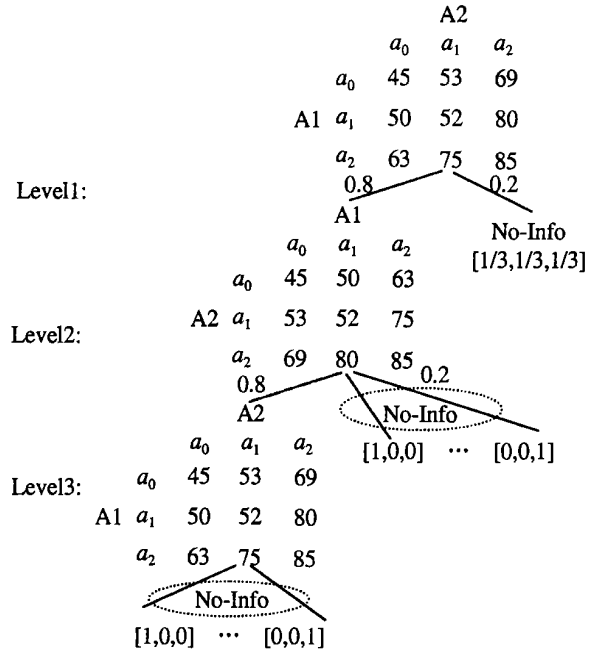


图 3 A1 的 3 层递归建模结构

分层模型均以 No-Info 模型结束, 即考虑 A1 在只具有 A2 的有限信息情况下进行推理决策。一般说, 智能体之间信息了解程度的差异决定了不同的模型差异。图 3 第 2 层中的 No-Info 表示当 A1 处发生异常情况(如交通事故)时, A1 不知道 A2 如何对其行动建模; 第 3 层中的 No-Info 表示在 A1 处正常情况下, A1 认为 A2 不具备任何 A1 对 A2 行动建模的信息。

图 3 示出的建模结构表示了智能体的优化决策递归地依赖于其它智能体的决策, 可用动态规划的方法由下至上地进行求解。本文实现了一种样本算法求解模型, 其步骤如下:

定义  $A_i = \{a_i^1, a_i^2, \dots, a_i^n\}$  为智能体  $i$  的可能行动集合,  $P_i = [p_i^1, p_i^2, \dots, p_i^n]$  为这些行动的概率分布。  $P_i$  的计算取决于每个行动被选作最优行动的频率  $F_i = (f_i^1, f_i^2, \dots, f_i^n)$ ,  $f_i^n$  的计算步骤如下:

input: 智能体  $i$  的支付矩阵  $M_i$  作为输入  
 output: 智能体  $i$  的行動的概率分布  $P_i$  作为输出  
 begin  
 $F_i (0, 0, \dots, 0), N 0$

在组成 No-Info 模型的样本概率分布集合

中,对每个概率分布  $P_N$  乘以支付矩阵  $M_i$ , 从中选取产生最大期望效益的行动  $a_i^k$ , 则

$$f_i^k = f_i^k + 1, N = N + 1$$

End for

$$P_i = [f_i^1/N, f_i^2/N, \dots, f_i^N/N]$$

Return

End

动态规划从递归模型结构的第3层开始,应用样本算法(样本密度取为0.1),则A2的可能行动( $a_0, a_1, a_2$ )的概率分布可由计算得出,分别为[0.68, 0.32, 0.00]。这个概率分布表示A1在A2正常情况下对A1采取行动的期望值的认知。然后由下至上在第2层应用样本算法,将上一步计算得到A1的行动概率与递归模型结构中第2层模型的概率分布相乘,并调用样本算法,计算出A2的行动集( $a_0, a_1, a_2$ )的概率分布为[0.22, 0.78, 0.00]。这样,尽管在第3层上模型并不清楚,但是A1的信息也足以确定。如果A2是理性智能体,则它将执行行动  $a_1$ 。

在第1层计算A2行动的概率分布(0.8\*[0.22, 0.78, 0.00] + 0.2\*[1/3, 1/3, 1/3]),结果为[0.24, 0.69, 0.07]。于是便可计算A1行动的期望效用

$$a_0: 0.24 * 45 + 0.69 * 53 + 0.07 * 69 = 52.2$$

$$a_1: 0.24 * 50 + 0.69 * 52 + 0.07 * 80 = 53.68$$

$$a_2: 0.24 * 63 + 0.69 * 75 + 0.07 * 85 = 72.82$$

如果A1是理性智能体,则它将在A2执行行动  $a_1$  时,执行获得最大期望效益(即最小的等待车流)的行动  $a_0$ 。在每个交通信号周期开始时,都执行上述计算过程。

#### 4.4 贝叶斯信念更新

分布式环境中的许多情况都是动态变化的,交通状况中尤其如此。因此,根据环境变化动态调整决策是非常重要的。我们提出采用贝叶斯学习的方法来实现智能体的信念更新,它是在对其它智能体行为观测的基础上,不断改变与其相关的概率分配来实现信念更新。在图2所示的情形中,由递归模型结构可知A1具有A2的两个可能模型。当A1观测到A2执行行动  $a_1$  时,关于A2模型的信念更新过程为:设模型  $M_1^2$  和  $M_2^2$  的初始分配概率分别为0.8和0.2,则有

$$BEL(M_1^2) = P(M_1^2 | a_1) =$$

$$P(a_1 | M_1^2) P(M_1^2) / P(a_1) =$$

$$(0.78 * 0.8) / 0.69 = 0.903$$

$$BEL(M_2^2) = P(M_2^2 | a_1) =$$

$$P(a_1 | M_2^2) P(M_2^2) / P(a_1) =$$

$$(0.33 * 0.2) / 0.69 = 0.097$$

其中,  $M_1^2$  是正常状况下A2的模型,  $M_2^2$  是异常状况下A2的模型。

贝叶斯信念更新在每个信号周期开始时都检测交通网络的变化,并实时调整网络中的信号进行协调控制。

## 5 仿 真

我们用VC++编写了基于上述方法的仿真程序,并与每个交叉口孤立地进行固定信号序列控制以及爬山法优化控制进行了仿真对比。设交通网络中的交通流饱和流量为400veh/h,仿真时间为3\*10<sup>5</sup>个信号周期。控制目标是使网络在交叉口处等候的车流长度最短。实验结果如表1所示,其中M1, M2和M3分别表示多智能体方法、爬山法优化信号方法和孤立交叉口固定信号方法,\*表示等

表1 3种控制方法的每周周期车辆平均等待队长

网络交通 流量(veh/h)	每周周期车辆平均等待队长		
	M1	M2	M3
< 1 008	< 0.05	< 0.05	< 0.05
1 008	0.10	0.12	0.19
1 152	0.15	0.17	0.24
1 296	0.26	0.29	0.37
1 368	0.34	0.48	0.75
1 512	0.47	0.80	1.12
1 656	0.77	1.08	1.95
1 728	0.91	1.67	2.54
1 944	1.23	2.10	3.38
2 088	1.94	2.45	5.80
2 186	2.58	3.87	6.78
2 320	3.56	4.72	8.92
2 548	4.03	6.43	11.2
2 670	4.85	8.77	16.56
2 858	5.60	10.27	*
3 016	6.48	13.89	*
3 120	7.86	15.73	*
3 368	9.22	17.23	*
3 420	10.26	*	*
3 608	11.22	*	*

待车流队长很长, 即已发生交通拥塞。

从仿真结果可以看出, 多智能体方法的控制效果优于其它两种方法。当网络交通流并不拥挤时, 其优势并不明显; 但当网络流量较大时, 本文方法在其它两种方法产生拥塞时仍能有效地控制交通流量。M3 只考虑了单路口情况, 显然不能适应交通流的动态变化, 只适用于网络车流量稀疏的情况。M2 考虑了其它交叉口的优化, 对一般网络车流量情况下能较好地进行网络优化控制, 但实时性较差, 不能适应拥塞时的网络流量优化和处理突发事件。

## 6 结 论

本文在分布式多智能体环境下对城市交通网络控制的建模和协调进行了深入研究, 应用递归建模方法进行网络中消除交通拥塞的高层次决策, 最大程度地减少或消除了交通网络控制中对各节点之间大量通讯的需求。提出采用贝叶斯学习方法与 RMM 框架结构相结合来跟踪交通网络中的动态变化, 各智能体之间实时更新信息, 增强了决策的可靠性。基于 RMM 和贝叶斯学习建立了简单的城市交通多智能体控制系统, 仿真研究取得了较好的效果, 对 ITS 的发展具有重要意义。

## 参考文献:

- [1] 冯蔚东, 贺国光, 刘豹. 交通流理论评述[J]. 系统工程学报, 1998, 3(1): 71-87.
- [2] 王亦兵, 韩曾晋, 贺国光. 城市高速公路交通控制综述[J]. 自动化学报, 1998, 7(1): 485-496.
- [3] B Burmeister, A Haddadi, G Matylis. Applications of multi-agent systems in traffic and transportation [J]. IEE Trans on Software Engineering, 1997, 144(1): 51-60.
- [4] C V Goldman, J S Rosenschein. Mutual supervised learning in multi-agent systems [A]. Distributed AI [M]. 1996. 85-96.
- [5] N V Findler. Distributed control of collaborating and learning expert systems for street traffic signals [A]. IFAC Distributed Intelligence Systems [C]. Pergamon Press, 1991. 125-130.
- [6] Gerhard Weib. Introduction to distributed artificial intelligence[M]. Cambridge: MIT Press, 1998.
- [7] P J Gmytrasiewicz, E H Durfee. A rigorous, operational formalization of recursive modeling[A]. ICMAS-95[C]. 1995. 125-132.
- [8] Martin Anthony, Norman Biggs. Computational learning theory [M]. Cambridge: Cambridge University Press, 1992.

(上接第 286 页)

## 参考文献:

- [1] Smith R P, Eppinger S D. Identifying controlling features of engineering design iteration[J]. Management Science, 1997, 43(3): 276-293.
- [2] Kusiak A, Larson T N, Wang J. Reengineering of design and manufacturing process[J]. Computers and Industrial Engineering, 1994, 26(3): 521-536.
- [3] Luh P B, Liu F, Moser B. Scheduling of design projects with uncertainty number of iterations[J]. European J of Operational Research, 1999, 113(3): 575-592.
- [4] Zhou M C, Di Cesare F, Desrochers A A. A hybrid methodology for synthesis of Petri net models for manufacturing systems [J]. IEEE Trans on Robotics and Automation, 1992, 8(3): 350-361.
- [5] Yan H S, Wang N S, Cui X Y *et al*. Modeling, scheduling and control of flexible manufacturing systems by extended high-level evaluation Petri nets[J]. IIE Trans, 1997, 29(2): 147-158.
- [6] Yan H S, Wang N S, Zhang J G *et al*. Modelling, scheduling and simulation of flexible manufacturing systems using extended stochastic high-level evaluation Petri nets[J]. Robotics and Computer-integrated Manufacturing, 1998, 14(2): 121-140.
- [7] Yan H S, Jiang J. Agile concurrent engineering[J]. Integrated Manufacturing Systems, 1999, 10(2): 103-112.