

基于拍卖和边际效益的自主信号交叉口建模方法

赵伊瑶^{1,2}, 沈震^{1,4}, 张浙鹏³, 熊刚(通信作者)^{4,5}, 朱凤华^{4,5}, 刘陶忠⁶

¹中国科学院自动化研究所, 复杂系统管理与控制国家重点实验室, 100190, 北京

²西安交通大学, 电子与信息工程学院, 710049, 西安

³中国科学院大学, 计算机与控制学院, 100049, 北京

⁴中国科学院云计算中心, 523808, 东莞

⁵中国科学院自动化研究所, 北京市智能化技术与系统工程技术研究中心, 100190, 北京

⁶海南乐导尔标识有限公司, 571924, 澄迈)

摘要: 虽然计算机已经被证实能够很好地驾驶汽车, 但在可见的未来较长时间阶段内, 人们仍将面临车辆总量居高不下的局面, 并且交通拥堵问题将随着智能交通的发展变得愈发突出。本文讨论了如何通过拍卖理论, 多智能体系统和机器学习技术来调整交叉路口的交通相互作用。拍卖过程中使用二级价格密封拍卖原理。近来, Rummery 等人将机器学习技术用于优化智能体的投标策略, 用边际效益控制状态表是否更新, 从而避免了盲目增加无谓的出价, 进一步优化了智能体出价策略。本文旨在对上述策略进行城市规模的仿真实验, 并将结果与仅有自主车辆出价策略的模型运行结果和基于队列长度的交叉路口控制运行结果进行了对比。实验发现该机制可以在一定程度上优化自主车辆的出价策略, 有效地减少车辆通过时间, 从而提高城市交通吞吐量, 缓解交通拥堵问题。

关键词: 智能交通; 拍卖理论; 机器学习; 边际效益; 多智能体

Autonomous Intersections Modeling Based on Auction and Marginal Benefits

Yiyao ZHAO^{1,2}, Zhen SHEN^{1,4}, Xipeng Zhang³, Gang XIONG^{4,5}, Fenghua ZHU^{4,5}, TaoZhong Liu⁶

¹The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, 100190, Beijing

²School of the Electronic and Information Engineering, Xi'an Jiaotong University, 710049, Xi'an

³School of Computer and Control Engineering, University of Chinese Academy of Sciences, 100049, Beijing

⁴Cloud Computing Center, Chinese Academy of Sciences, 523808, Dongguan

⁵Beijing Engineering Research Center of Intelligent Systems and Technology, Institute of Automation, Chinese Academy of Sciences, 100190, Beijing

⁶Hainan LDE Sign Co. Ltd. Chengmai 571924, Hainan)

Abstract: Although the computer has been proven to be able to drive the car well, but in the foreseeable future for a long time, people will still face the situation of high total vehicle traffic, and traffic congestion problems will become more with the development of intelligent traffic prominent. This paper discusses how to adjust traffic

收稿日期:

作者简介: 赵伊瑶, 女, 1996 年 5 月, 西安交通大学学生, 河北人, 西安交通大学电子与信息工程学院, 710049,

E-mail: yiyaozhao@126.com, 联系电话 18292887238

interactions at intersections through auction theory, multi-agent systems, and machine learning techniques. The auction process uses a secondary price seal auction principle. Machine learning techniques are used to optimize the bidding strategy for agents. Recently, there have been some researches on auction strategy using machine learning techniques. Rummery et al use the marginal efficiency to update the status table, thus avoid the blind increase in unnecessary bidding, and further optimize the smart bid strategy. This objective of this research is to test such strategy in large scale traffic simulation, and compare the result of the model with only the independent vehicle bidding strategy and the result of the intersection control based on the queue length. It is found that the bid mechanism can optimized city traffic to a certain extent. The bid strategy of the autonomous vehicle effectively improves the throughput of city road network and reduces the vehicle traveling time and alleviates the traffic congestion problem.

keywords: Intelligent transportation; auction theory; machine learning; marginal benefit; multi-agent

引言

为使社会实现经济繁荣,极大地提高国家的综合国力与全球竞争力,提高物资的运作效率是很重要的一个前提条件。在当前及以后一段时间内,城市地区的主要运输途径都是高速公路。随着社会的发展与科技的进步,公路上的交通工具也将实现从人工驾驶车辆向自动驾驶车辆的过渡。自动驾驶车辆具有减少交通拥堵能力的巨大潜力,能够减少燃料与时间的大量浪费。所以,在城市公路系统中,如何自主地实现车辆之间的信息交互与共同调控,实现自动驾驶车辆的有效调度,使得在更大程度上提高车辆的通行能力,缓解交通压力,成为人工智能的重要研究方向之一。现阶段国内外学者对于交通仿真的研究,大多是通过平行控制等方法,针对于道路交通的进行调整与路线的总体优化^{[1][2][3]}。目前用于控制交通的方法,特别是在交叉路口处车辆的控制,与人力驾驶阶段相比,大多没有充分利用自主车辆的灵敏度和精度高的特点。而交叉路口处灵活的交通管制方案则为自主车辆调控提供了解决交通拥堵的新机会。

现有的交叉路口控制方案,如交通信号灯等,都是适应于人工驾驶的现状的,依赖于人的反应能力与反应时间,效率较低,已不再适用于自动驾驶车辆的高效率调度。而对比之下我们可以发现,与交叉口管理员进行通信的自动驾驶车辆可以进一步提高利用道路的能力,使得在高度拥挤的交叉点处

将时间延迟降低高达两个数量级^[4]。

1 相关工作

传统的交通调度方案多以先到先得(FIFO)的方式分配预留时间或空间,从而实现交叉路口处的有效调度^[5]。利用市场经济的力量来管理拥堵的概念提出可以追溯到 Pigou^[5]和 Beckmann 等人^[6]。他们提出网络边际成本收费可以协调司机选择路线,从而最小化总旅行时间。但是这些早期的模型的建立需要一些具有强烈限制性的假设来支撑,所以在具体应用时的实用性与仿真的效果受到一定影响。之后,相关学者对于这个方向进行了进一步的探索与完善,尝试使用更加复杂的定价模型从而解决因时间价值的变化^{[7][8]}带来的影响以及实现自适应动态定价^{[9][10]}。近年来,有些学者开始尝试以市场经济为基础的定价机制,从而使旅行者做出对自己更有价值的决策^{[11][12]}。在这些文章中^[11],作者提出基于拍卖的自主交叉口管理方案,该方案采用系统投标的方式,反映了旅行者的不同时间价值的变化,同时通过不同钱包的设定从而强调公平性问题。而有些学者^[12]也提出将交叉路口空间视为一组稀缺商品,通过拍卖理论、多智能体系统与机器学习技术来调节交叉路口的信息交互与交通调度问题。但文章中的出价策略主要适用于比较小规模的路口组合,对于大规模系统的应用还存在一些问题。当规模增大时,从总体的旅行时间来看改善效果并不理想。Isukapati I K 等人在之前学者^[11]的模型基础

上进行了进一步改进^[13]。它提出了一个类似博弈的框架，初始费用对所有司机征收，同时通过交叉路口的每个驾驶员影响交通信号的性能。但文章忽略了框架的安全约束，这在现实世界中是不现实的。Levin M W 等人则在拍卖模型中加入优先级的概念以减少延迟，同时研究了交叉口处拍卖模型将如何影响城市网络上的用户平衡(UE)行为^[14]。但是模型对于拍卖中采取极高价格的车辆几乎没有益处。

本文在 Carlino 等人的研究基础^[11]上，加入机器学习 SARSA 算法，对于交叉路口处的车辆智能体的出价策略进行进一步优化，同时在出价决策过程中加入边际效益模型，以避免盲目提高出价对于智能体的后续行驶造成过大影响。

2 拍卖模型

2.1 拍卖建模

拍卖是人类社会特殊的商品交易方式。在经济学中，拍卖有着更深层次的意义。拍卖的理论基础来源于信息经济学和博弈论的核心思想。从这个意义上来看，拍卖本身是一个博弈的过程。

本文中拍卖采用二级价格密封拍卖原理，即最高价得标，次高价付款拍卖法。这种拍卖方法密封投标金额，在拍卖进行时拍卖的出价对于竞标者之间是未知的。在每次拍卖中，每个投标者提交密封的交易价格，出价最高者赢得拍卖的商品，但交易却以所有出价中的第二高价进行。这种拍卖方式可以减少出价的随意性，帮助竞标者优化其出价行为。

在交叉路口的建模过程中，我们将拍卖进行如下应用。假设一次只能有一个司机通过路口。我们在路口处设置一个控制器。将路口的可通过时间离散化，作为每次竞标的商品。在每次出价中，每个车道的头车作为候选人参与竞价，但出价由头车所在车道所有车辆共同决定，即所在车道车辆都为头车进行出价。胜者将可以按照出价的比列分摊支付次高的出价（即二级价格密封拍卖）。

在交叉口处拍卖过程的算法如下：

- 1) 统计确定每个车道的候选者 A, B, C, E, F 与所在车道即将通过路口的所有车辆智能体，找出即将参与拍卖的候选者与出价者的列表；
- 2) 拍卖竞标过程开始，每个车道的车辆进行出价，加和汇总为各车道的头车候选者总出价；
- 3) 路口控制器通过比较决定总出价的最高与次高的候选者，设为 W 与 S；
- 4) 总出价最高的候选者 W 赢得本次拍卖，所在车道车辆按照竞标时的出价比例分摊支付次高出价候选者 S 的出价；
- 5) 赢得本次竞价的车道头车获得此时刻交叉路口的使用权，可以通过路口；
- 6) 本次竞标过程结束，转到步骤 1)，开始下一次竞标过程。

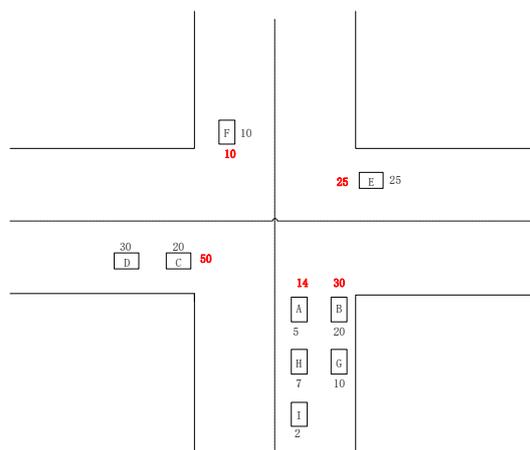


图 1 交叉口处拍卖模型举例

在交叉路口处，每个车道的头车作为该车道的候选者，在拍卖时的出价为所在车道所有车辆的出价和，每轮拍卖出价决定某一个头车是否可以通过路口。举例来说，如图 1 所示，A 所在车道总出价为 14，B 所在车道总出价为 30，C 所在车道总出价为 50，E 所在车道总出价为 25，F 所在车道总出价为 10。则在这个例子下，C 所在车道候选者赢得本次竞价，车辆 C 获得通过此交叉路口的权利，B 所在车道候选者出价第二，C 所在车道内所有车辆 C 和 D 将按照竞标时出价的比列分摊出价费用，此时的出价费用值为

B 所在车道候选者的出价。即最后付款时，C 出价 $\frac{20}{30+20} * 30 = 12$ ，D 出价 $\frac{30}{30+20} * 30 = 18$ 。C 车可以通过交叉路口，本次竞标过程结束。

由于在拍卖过程中，时间作为稀缺商品，是竞标的对象，所以在竞标时不必考虑车辆在路口处的转向是否不同，只需考虑车辆在车道排队的位置即可。

2.2 应用边际效益理论的强化学习算法

以往在交叉路口处的每次出价过程中，一般采用随机出价的方式，即每个车辆智能体的出价是随机的^[14]。但近年来有学者开始考虑使其出价有规律可循。Isukapati 等人尝试利用贝叶斯算法来更新他们提交出价的方法^[13]。Mashayekhi 等人则尝试利用 Q 学习算法来为运动智能体创建最优投标策略^[14]。

强化学习技术是一种机器学习方法，它可以只通过所处环境的状态、当前策略以及不确定的环境奖赏来对行为策略进行优化，从而找出最优行为策略。在交通仿真领域，由于环境因素的不确定性，因此强化学习成为一种优化仿真过程的有力工具。

SARSA^[15]是由 Rummery 和 Niranjan 提出的一种优化算法，它是将预测的瞬时差分学习算法扩展到控制学习中形成的。它仍采用 Q 值迭代的方法。如果状态转移概率 P 和奖赏函数 R 未知，我们可以用 SARSA 算法来直接估计最优动作值函数来确定所要使用的策略。这个算法不用估计环境模型。与 Q 学习方法类似的，它是直接优化一个可迭代计算的 Q 函数，此函数为在状态 s_t 时，执行动作 a_t 后按照最优动作序列执行时的折扣累积强化值。

本文算法中使用 Q 表 $Q(s_t, a_t)$ 是由 t 时刻状态 s_t 和动作 a_t 进行索引的矩阵。假设某次出价时车辆智能体处于状态 s_t ，且它根据已有的 Q 表和其他算法结合选择动作 a_t 与环境交互。然后，智能体从执行动作 a_t 后的结果中接收奖励函数值 r ，并观察新的状态 s_{t+1} 。之后，Q 表通过以下公式进行更新：

$$Q_k(s_t, a_t) = \begin{cases} (1 - \alpha_t)Q_{k-1}(s_t, a_t) + \alpha_t[r_t + \gamma Q_{k-1}(s_{t+1}, a_{t+1}), \frac{s_{t+1} - s_t}{a_{t+1} - a_{t-1}} > m \\ Q_{k-1}(s_t, a_t), \frac{s_{t+1} - s_t}{a_{t+1} - a_{t-1}} \leq m \end{cases} \quad (1)$$

其中， α_t 为学习率， γ 是折现因子，m 为边界效应的临界值，可以通过实验进行设定，从而得到更好的效果。本文实验时取 $m=0.5$ 进行计算。SARAS 算法执行过程如下：

- 1) 确定当前状态 s_t ， s_t 表示 t 时刻时车辆智能体在交叉口处所处的位置信息；
- 2) 选择要执行的操作 a_t ，其中 a_t 表示在 t 时刻车辆智能体的出价策略；
- 3) 从执行动作 a_t 和结果状态 s_{t+1} 接收奖励乘数 r ；
- 4) 使用公式①更新 Q 表状态 $Q(s_{t+1}, a_{t+1})$ ；
- 5) 转到步骤 1)。

其中公式中的奖励值函数 r 用拍卖是否中标来进行判定及更新：当本次投标中标时，设置奖励乘数为 1；否则设为 0。

在仿真过程中，Q 表的初始状态值使用高斯随机数进行建立，得到经验知识 $\langle s_t, a_t, s_{t+1}, r_{t+1} \rangle$ ($t=0$)，在之后的投标迭代中根据式①对函数值进行修正。同时在迭代过程中增加边际效益的判断，但边际效益值大于预期的边界值时，允许 Q 表进行更新，否则表示动作更新时所需付出的成本大于所获得的收益，从经济学角度讲，由于机会成本过高此行为不被允许，此时，Q 表不更新。

SARSA 算法与 Q 学习算法的主要差别在于 Q 学习采用的是值函数的最大值进行迭代，而 SARSA 学习则是使用实际的 Q 值进行迭代。此外，在每次学习时 SARSA 算法中智能体根据当前的 Q 函数值确定下一动作，而 Q 学习则根据修改后的 Q 值确定动作。

3 实验与分析

为了观测 SARSA 拍卖模型是如何影响智能体行驶时间的，我们将 SARSA 拍卖模型应用 AORTA 系统进行了仿真。AORTA 是一个交通微观仿真系统，用 Scala 构建，是一种基于多智能体、简单且直观的微观交通仿真系统^[11]。仿真时采取空间连续，时间分散的方法，将时间设定为固定的步长，在交叉路口处，当拍卖的候选者胜利时，取得

交叉路口在这一时刻的使用权，然后在下一个时刻进行下一轮拍卖，即认为时间是拍卖的物品。

我们将 AORTA 系统在一些城市的地图上进行了仿真，并选取不同模型下车辆的平均行驶时间作为评价指标，仿真结果如表 1 和图 2 所示。

从仿真数据可以看出，拍卖模型的调度效果要好于 FIFO 模型，司机的平均行驶时间大约比 FIFO 模型中的行驶时间减少了约 10%。而加入边际效益理论的 SARSA 拍卖模型的效果要好于改进前，司机平均行驶时间相比于改进前又显著减少。综合对比得出，加入边际效益理论的 SARSA 拍卖模型可以一定程度上减少城市路网车辆的平均行驶时间，缓解交通拥堵。

表 1 不同模型下司机的平均行驶时间(s)

平均行驶时间 (s)	FIFO	拍卖模型	SARSA模型
Austin	4837±92	4647±42	4473±50
Seattle	4713±62	2666±52	2381±67
San Francisco (SF)	1883±31	1596±9	1438±15

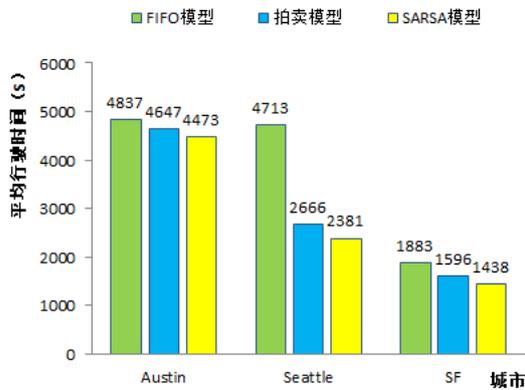


图 2 不同模型下车辆的行驶时间(s)

4 结论

本文通过对已有的拍卖模型进行了讨论，包括竞价拍卖模型、利用多智能体系统和强化学习技术的 SARSA 拍卖模型。拍竞价卖模型中使用二级价格密封拍卖原理。而 SARSA 拍卖模型由于考虑了边际效益，能更有效地调整车辆在交叉路口的相互作用，进而优化了智能体出价策略。本文旨在对于该策略进行城市规模的交通仿真，并将结果与仅有自主车辆出价策略的模型运行结果和基于队列长度的交叉路口控制运行结果进行了对比，发现 SARSA 拍卖机制可以智能地优化车辆的出价策略，在一定程度上减少车辆行驶时间，进而提高城市路网交通的吞吐量，缓解交通拥堵问题。

未来，我们的工作将重点放在以下几个部分：

1. 进一步优化道路交叉口处调度的数学模型，使之更加具体化与现实化；
2. 考虑交叉口的相互影响，实现区域式配合调度；
3. 考虑其他方法进一步解决车辆智能体出价时的公平性问题。

致谢

本文获得中国科学院大学生创新实践训练计划项目“智能车载终端与 GPU 交通实验平台”，国家自然科学基金项目 (61233001、61773381、61773382、71232006、61304201)；广东省科技厅项目 (2014B090902001、2015B010103001、2016B090910001) 等的支持。

参考文献

[1] Xiong G, Zhu F, Liu X, et al. Cyber-physical-social system in intelligent transportation[J]. IEEE/CAA Journal of Automatica Sinica, 2015, 2(3): 320-333.

[2] Li L, Lin Y, Zheng N, et al. Parallel learning: a perspective and a framework [J].

- IEEE/CAA Journal of Automatica Sinica, 2017, 4(3): 389-395.
- [3] Wang F Y, Zhang J, Wei Q, et al. PDP: parallel dynamic programming [J]. IEEE/CAA Journal of Automatica Sinica, 2017, 4(1): 1-5.
- [4] Dresner K, Stone P. A multiagent approach to autonomous intersection management [J]. Journal of artificial intelligence research, 2008, 31: 591-656.
- [5] Pigou A C. The economics of welfare, 1920[J]. McMillan & Co., London, 1932.
- [6] M. J. Beckmann, C. B. McGuire, and C. B. Winston, Studies in the Economics of Transportation. Connecticut: Yale University Press, 1956
- [7] Dial R B. Network-optimized road pricing: Part I: A parable and a model [J]. Operations Research, 1999, 47(1): 54-64.
- [8] Dial R B. Network-optimized road pricing: Part II: Algorithms and examples [J]. Operations Research, 1999, 47(2): 327-336.
- [9] Lou Y, Yin Y, Laval J A. Optimal dynamic pricing strategies for high-occupancy/toll lanes [J]. Transportation Research Part C: Emerging Technologies, 2011, 19(1): 64-74.
- [10] Ozbay K, Morgul E F, Ukkusuri S V, et al. Mesoscopic simulation evaluation of dynamic congestion pricing strategies for New York City crossings[C]. Transportation Research Board 90th Annual Meeting. 2011 (11-3612).
- [11] Carlino D, Boyles S D, Stone P. Auction-based autonomous intersection management[C]. 16th International IEEE Conference on Intelligent Transportation Systems-(ITSC), 2013: 529-534.
- [12] Mashayekhi, M., & List, G. (2015, July). A multi-agent auction-based approach for modeling of signalized intersections, In Second Workshop on Synergies Between Multiagent Systems, Machine Learning and Complex Systems (TRI 2015), Buenos Aires, Argentina.
- [13] Isukapati I K, List G F. Agent based framework for modeling operations at isolated signalized Intersections[C]. 18th International Conference on Intelligent Transportation Systems (ITSC), 2015: 2900-2906.
- [14] Levin M W, Boyles S D. Intersection auctions and reservation-based control in dynamic traffic assignment [J]. Transportation Research Record: Journal of the Transportation Research Board, 2015 (2497): 35-44.
- [15] 李瑞. 强化学习主要算法的研究[J]. 渝西学院学报: 自然科学版, 2004, 3(3): 22-25.