

doi:10.3963/j.issn.1671-4431.2022.05.007

基于马尔可夫决策过程的出租车寻客路径优化

程琳,张晨皓,于新莲,杜明洋,任姿蓉
(东南大学交通学院,南京 211189)

摘要: 为提高出租车的长期收益,作者在路网网格化的基础上建立了基于马尔可夫决策过程的路径优化模型。将车辆当前所位于的网格位置定义为状态,将从当前网格选择某一相邻网格出发定义为动作,使用策略迭代法对问题进行求解,并采用高斯-赛德尔迭代进行加速。以深圳市典型工作日全天797辆出租车的GPS数据进行试算和仿真,计算速度是雅可比迭代的1.85倍。将该算法与随机游走和全局热点算法进行比较,结果表明,所提出模型的平均单位距离收益分别提高了22.1%和12.9%,载客里程占比分别提高了18.8%和10.4%,具有较好的优化效果。

关键词: 城市交通; 动态路径优化; 马尔可夫决策过程; 出租车; 策略迭代; 数据挖掘

中图分类号: U 121

文献标识码: A

文章编号: 1671-4431(2022)05-0040-07

Route Optimization of Taxicab Based on Markov Decision Process

CHENG Lin, ZHANG Chen-hao, YU Xin-lian, DU Ming-yang, REN Zi-rong

(School of Transportation, Southeast University, Nanjing 211189, China)

Abstract: A dynamic route optimization model based on Markov decision process is formulated after gridding of network. The state is defined by the grid location of vehicle while the action is defined by the selection of an adjacent grid from the current location. The problem is solved by the policy iteration, Gauss Seidel iteration is used for acceleration. The model is tested by taking the GPS data of 797 taxicabs in a typical working day in Shenzhen, and the speed is 1.85 times that of Jacobi iteration. The model is compared with two heuristic algorithms by simulation, showing that the average profit per unit distance of the optimal route strategy is 22.1% and 12.9% higher than that of random walk and global hotspot algorithm, while the proportion of passenger carrying distance is 18.8% and 10.4% respectively, which has a good effect on route optimization.

Key words: urban traffic; dynamic route optimization; markov decision process; taxicab; policy iteration; data mining

出租车以其灵活的特点、凭借在运营时间和空间上的巨大优势,在现如今高度活跃的城市道路交通中发挥着日益重要的作用,但同时也带来了一系列的问题,最主要的是空驶问题。以深圳市为例,出租车的空驶里程占比达41.6%,空驶时间占比高达49.2%^[1],造成了严重的环境污染和资源浪费。因此优化空驶出租车的路线来降低空载率,提高长期收益,对司机和社会都具有重要意义。

空驶出租车的路线规划研究方法主要分为载客热点推荐和优化建模两类。前者主要通过挖掘车辆GPS历史数据,采用密度分析^[2]、轨迹聚类^[3-4]等方法为空载出租车推荐一个上客地点或多个上客地点集合,此类推荐方法高度依赖数据质量,且没有充分考虑空载时间、热点载客概率等相关因素的影响。而优化类方法在系统分析各类寻客影响因素的基础上,通过构建数学模型,为车辆提供空间移动的最优解。根据车

收稿日期:2022-03-02.

基金项目:国家自然科学基金(52172318,52131203).

作者简介:程琳(1963-),男,博士,教授.E-mail:gist@seu.edu.cn

通讯作者:张晨皓(1998-),男,硕士生.E-mail:1071955807@qq.com

辆路径规划时域的不同, 优化类方法可进一步分为半周期和多周期两类。半周期问题的规划时域是从车辆上一次落客开始, 到下一次接客结束; 此类研究多以最大化下一次的接客质量为目标, 如以最快时间^[1]、最大概率^[5]、接到预期收益最高的乘客^[6]等, 而不考虑乘客目的地的质量以及后续运营状态^[7-8], 但这种短视性会限制最优解的质量, 例如当出租车的落客点是偏远且乘客稀少的区域时, 会增大潜在的返程成本, 无法在更长期的规划时域下获得最优收益。多周期问题的规划时域由若干个“寻客-接客-送客”阶段构成; 此类研究多以最大化长期收益为目标, 将多个周期的收益考虑在内, 具有多阶段和非短视的特点。多周期问题的研究主要体现在使用马尔可夫决策过程的模型框架^[9-12]对问题进行建模, 马尔可夫决策过程可以充分考虑当前决策对未来多个周期的影响, 从而得到实现最大收益的决策序列, 但其对状态空间的大小十分敏感, 在问题规模较大时求解速度严重下降。

于新莲等^[9]利用排队论模型细化了路段乘客和车辆的到达, 将道路结点定义为状态, 提出了基于马尔可夫决策过程的多周期交叉口空驶出租车路径优化模型, 但其中的部分数学模型参数较难确定, 且状态空间规模较大, 使用传统方法求解时速度较慢。部分学者^[10-11]对道路结点进行聚类, 分别对聚类后的潜在接客点或片区建立了基于马尔可夫决策过程的路径优化模型, 但由于状态空间仍然较大, 不可避免地存在数据稀疏性的问题, 导致部分路段数据缺失或存在异常波动。

为缩小状态空间, 减小数据稀疏性对结果带来的影响, 常见的方式是对路网进行网格化抽象处理或者对 GPS 历史数据进行网格聚类, 此类方法的处理时间独立于数据对象, 处理速度较快, 在研究低维数据时具有显著优势, 此外规整的聚类方式还可以实现与马尔可夫决策过程模型的良好结合。关于路网的网格化抽象处理, 高瞻等^[5]在边长 100 m 的网格划分情况下研究了空载出租车的路径问题, 但仅考虑载客概率并未考虑后续周期的运营收益; 杨昌^[12]在边长 200 m 的网格划分情况下, 研究了空载出租车的路径规划问题, 但所有参数均来自 GPS 数据并未使用实际路网信息。关于 GPS 数据的网格聚类, 徐志锋等^[13]和郑运鹏等^[14]在边长 200 m 和约 1 000 m 的网格划分情况下, 分别对深圳市和南京市的交通热区进行了挖掘识别, 但并没有与车辆的路径优化相结合。

作者在前人研究的基础上, 从提高空驶出租车的长期收益出发, 将问题建模为马尔可夫决策过程, 对状态和动作进行了新的定义。结合实际道路信息对路网进行网格化处理, 网格尺寸设置为 500 m, 筛选可达网格以进一步缩小状态空间。使用策略迭代进行求解, 并通过高斯-赛德尔迭代对策略评估进行加速。最后使用大规模 GPS 历史数据和深圳市路网对模型进行计算和仿真, 结果显示模型和算法对在长期收益和求解速度方面具有较好的提升效果。

1 问题建模

1.1 模型假设

在建模之前做如下假设: 1) 研究的对象是单辆出租车; 2) 研究的问题所基于的路网是非时变的; 3) 当车辆接到乘客时, 问题转变为从上车地点所在网格到目的地所在网格的最短路问题, 作者并未对其进行明确研究。

基于此, 作者提出了用于最优路径问题的马尔可夫决策过程模型, 此类问题的网络模型可以描述为 $G = (N, A)$, 其中 N 是结点集合, A 是路段集合, 车辆在动态路径优化的时域内不断搜索和接送乘客。由于研究中所使用的网络模型是基于网格的, 因此需要对路网进行处理。

1.2 路网处理和网格最短路

将研究区域在地理空间上划分为若干边长为 a 米的正方形网格, 所有网格构成集合 S , 第 i 个网格用符号 i 表示, $i \in S$ 。车辆在接到乘客之后, 以最短路运送乘客至目的地, 由于对路网进行了网格化处理, 因此建立基于网格的最短路计算模型如式(1)所示。

$$D_{ii'} = \frac{\sum_{n_i=1}^{N_i} \sum_{n_{i'}=1}^{N_{i'}} d_{n_i n_{i'}}}{N_i N_{i'}} \quad (1)$$

式中, $D_{ii'}$ 表示网格 i 与网格 i' 之间的最短路长度, $d_{n_i n_{i'}}$ 表示网格 i 中的结点 n_i 与网格 i' 中的结点 $n_{i'}$ 之间的最短路长度, 单位均为 km, 道路网络由 OpenStreetMap 获取, 使用 arcGIS 对道路结点进行网络分析, 计算得到 OD 距离矩阵; N_i 和 $N_{i'}$ 分别表示网格 i 与网格 i' 内的实际道路结点数量; n_i 和 $n_{i'}$ 分别指代在地理位置

上位于网格 i 与网格 i' 内的道路结点。

1.3 状态和动作

作者将车辆的状态定义为当前所位于的网格 $i, i \in S$, 且状态仅与网格位置有关; 将动作 a 定义为车辆选择停留在原网格、向上、向右上、向右、向右下、向下、向左下、向左和向左上中的某一种行为, 分别对应数字 0、1、2、3、4、5、6、7、8, 如图 1 所示。

从网格 i 产生的动作集合为 $A(i)$, 对于一个给定的状态 i 和动作 $a, a \in A(i)$, 转移到新状态 i' 时可能产生两种方式: 1) 空驶出租车采取动作 a , 在前往网格 j 的过程中未匹配到乘客, 下一状态是在状态 i 下采取动作 a 所对应的状态 j ; 2) 空驶出租车采取动作 a , 在前往网格 j 的过程中成功匹配到乘客, 下一状态是乘客目的地 i' , 如图 2 所示。

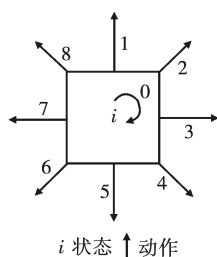


图1 状态与动作的定义

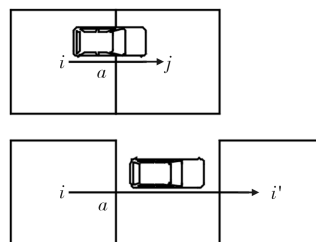


图2 基于网格的两种状态转移方式

1.4 乘客匹配概率和目的地概率

用 N_i^1 和 N_i^0 分别表示空驶车辆在网格 i 内接到和未接到乘客的事件次数, 以 1 h 的固定时间窗进行研究, 挖掘落在不同时间窗内的出租车 GPS 数据, 计算得到前述参数。则可以利用式(2)计算得到每个网格 i 中的乘客匹配概率 $P(i)$, 这表示当空驶车辆穿过网格 i 时接客事件发生的概率, 不同时间窗的不同状态具有不同且固定的乘客匹配概率。

$$P(i) = \frac{N_i^1}{N_i^0 + N_i^1} \quad (2)$$

用 $N_{i \rightarrow i'}$ 表示车辆在网格 i 内接到乘客, 并前往网格 i' 运送乘客的事件次数, 则可以利用式(3)计算各网格 i 在不同时间窗的乘客目的地概率 $P_{i \rightarrow i'}$, 表示乘客在网格内 i 上车, 前往网格 i' 的概率, 不同时间窗具有不同且固定的乘客目的地概率。

$$P_{i \rightarrow i'} = \frac{N_{i \rightarrow i'}}{N_i^1} \quad (3)$$

1.5 状态转移概率

对于在状态 i 采取动作 $a, a \in A(i)$, 下一个状态是 j 的情况, 状态转移概率 $P_{ii'|a}$ 如式(4)所示。

$$P_{ii'|a} = \begin{cases} 1 - P(i) + P(i)P_{i \rightarrow i'} & i' = j \\ P(i)P_{i \rightarrow i'} & i' \neq j \end{cases} \quad (4)$$

对于第 1 种情形, 车辆的当前状态是 i , 下一状态是 j , 则状态由 i 转换为 j 的概率等于在网格 i 内未匹配到乘客的概率, 与在网格 i 内匹配到乘客且目的地是网格 j 的概率之和。对于第 2 种情形, 车辆当前状态是 i , 下一状态不是 j , 则状态由 i 转换为 i' 的概率等于在网格 i 内匹配到乘客且目的地不是网格 j 的概率。车辆选择最短路到达目的地, 落客后继续寻客。

所建立的模型为空驶出租车提供基于网格的决策序列, 以实现最大化长期收益。在匹配到乘客前, 决策序列为车辆巡游路线提供指导作用; 发生匹配后, 乘客的需求决定车辆的送客路线, 后续的决策序列失去作用成为无效序列。过长的无效序列会占用存储空间, 浪费计算资源, 因此需要使用序列中网格的匹配概率的累乘对决策进行约束。设置阈值 $\zeta, 0 < \zeta < 1$, 当进行到第 L 个网格, 若累乘小于 ζ 则序列截止, 此时有 $1 - \zeta$ 的把握认为车辆已经接到乘客, 决策序列长度为 L , 如式(5)所示, 阈值 ζ 设置为 0.05。

$$\prod_{l=1}^L P(i) < \zeta \quad (5)$$

1.6 即时收益

与状态转移概率的公式(4)相对应, 采取动作 a 从状态 i 转移到状态 i' 所获得的即时收益, 分两种情况进

行计算,如式(6)所示。

$$g_{ii'|a} = \begin{cases} \frac{-\alpha D_{ii} [1 - P(i)] + [F(D_{ii'}) - \alpha D_{ii'}] P(i) P_{i \rightarrow i'}}{1 - P(i) + P(i) P_{i \rightarrow i'}} & i' = j \\ \frac{[F(D_{ii'}) - \alpha D_{ii'}] P(i) P_{i \rightarrow i'}}{P(i) P_{i \rightarrow i'}} & i' \neq j \end{cases} \quad (6)$$

α 是车辆单位距离的运营成本,设定为每公里油耗成本。 $F(D_{ii'})$ 是关于行驶距离的票价函数,采用深圳市的分段线性票价收费制度,结构如式(7)所示,不考虑等候时间的收费,距离单位为 km,票价单位为元,其中 $f_0, \beta, \gamma, \delta, D_0, D_1$ 和 D_2 为票价收费制度中的参数,后文会给出具体数值。

$$F(D_{ii'}) = \begin{cases} f_0 & D_{ii'} \leq D_0 \\ f_0 + \beta(D_{ii'} - D_0) & D_0 \leq D_{ii'} \leq D_1 \\ f_0 + \beta(D_1 - D_0) + \gamma(D_{ii'} - D_1) & D_1 \leq D_{ii'} \leq D_2 \\ f_0 + \beta(D_1 - D_0) + \gamma(D_2 - D_1) + \delta(D_{ii'} - D_2) & D_{ii'} \geq D_2 \end{cases} \quad (7)$$

1.7 贝尔曼最优方程

在状态 i 下执行动作 a 产生状态转移,下一状态为 i' ,对应的即时收益为 $g_{ii'|a}$,多次的状态转移组成一条马尔可夫链;定义状态价值函数 $V(i)$ 评价从状态 i 开始的马尔可夫链的质量,用期望即时收益和期望后续收益之和来表示,状态价值函数的期望可由方程(8)给出,即为贝尔曼期望方程。

$$V(i) = E[g_{ii'|a} + \rho V(i')] \quad (8)$$

ρ 是衰减系数, $0 < \rho < 1$, 由于这是一个无限时域问题,因此设置 $\rho = 0.5$, 以避免无限循环。

为实现最优解,司机从状态 i 开始,选择下一步动作 a 来最大化期望收益,得到的最优期望收益通过 $V^*(i)$ 来表示,通过求解贝尔曼方程^[15]可得到基于马尔可夫决策过程的动态路径优化**最优期望收益**如方程(9)所示。

$$V^*(i) = \max_{a \in A(i)} \sum_{i' \in S} \pi(a|i) [g_{ii'|a} + \rho V^*(i')] P_{ii'|a} \quad (9)$$

策略 $\pi(a|i)$ 是在状态 i 的动作分布,在贪婪思想的指导下通过求解方程(10),可得到最优策略 $\pi^*(a|i)$,表示车辆在状态 i 时所选择的下一步具体动作,在每个状态按照对应的最优策略决策即可得到最优状态价值。

$$\pi^*(a|i) = \operatorname{argmax}_{a \in A(i)} \sum_{i' \in S} \pi(a|i) [g_{ii'|a} + \rho V^*(i')] P_{ii'|a} \quad (10)$$

2 求解算法

在状态空间较小时策略迭代算法具有较快的收敛速度,作者在对路网进行网格化处理、筛选可达网络缩小状态空间后,使用策略迭代对问题进行求解,同时在策略评估时使用高斯-赛德尔迭代进行加速,算法的具体流程如下。

Step 1 算法输入 由式(4)和式(6)确定并输入状态转移概率 $P_{ii'|a}$,即时收益 $g_{ii'|a}$ 和收敛准则 ϵ (较小的正数)。

Step 2 算法初始化 初始化状态价值函数 $V(i)$ 、衰减系数 ρ 和初始策略 $\pi = \pi_0(a|i)$ 。

Step 3 **策略评估** 在当前策略 $\pi_{\text{prev}} = \pi$ 下,对每个状态 $i, i \in S$,利用式(8)建立关于 $V(i)$ 的线性方程组,并通过高斯-赛德尔迭代求解,待所有状态的价值函数达到收敛准则 ϵ ,迭代结束并输出 $V(i)$ 。

Step 4 **策略提升** 对每个状态 $i, i \in S$,利用式(9)在贪婪思想的指导下,通过最大化 $V(i)$ 来调整策略 $\pi(a|i)$,更新策略 π 。

Step 5 算法终止判断 判断策略 π 是否与 π_{prev} 相同,若不同则返回 Step 3,若相同则停止算法,输出最优状态价值函数 $V^*(i) = V(i)$,输出最优策略 $\pi^*(a|i) = \pi$ 。

3 计算检验与仿真评价

3.1 实验设置

1) 研究区域和路网 以深圳市路网作为研究对象,道路网络由 OpenStreetMap 获取,使用 arcGIS 根据

式(1)计算得到 OD 距离矩阵,研究区域为东经 $113^{\circ}46'$ 至 $114^{\circ}37'$,北纬 $22^{\circ}27'$ 至 $22^{\circ}52'$ 。道路网络由 14 471 个结点和 23 245 个路段组成,网格边长设置为 500 m,研究区域可划分为 92 行、187 列,共计 17 204 个网格,此处的状态、动作和状态转移都是基于网格的。为了进一步缩小状态集合,在保证路网完整性的基础上剔除不存在 GPS 数据且对车辆不可达的网格,缩小状态后的 5 007 个可达网格如图 3 所示。

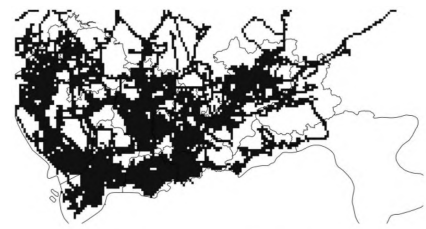


图3 缩小状态集合后的可达网络

2)GPS 数据 数据来源于新泽西州立罗格斯大学 Desheng Zhang 课题组^[16],涵盖 797 辆出租车,平均采样时间间隔约为 15 s,含有载客状态的信息以判断车辆的空驶情况,认为该样本数据集可以推断车辆的一般行驶特点,并对所提出的方法进行评价。完成数据清洗后,选取 2013 年 10 月内不存在引起出行变化特殊事件的一个具有代表性的工作日,分小时对一天的车辆运行数据进行研究。

3)参数设置 作者将整个研究期按小时划分为 24 个时间段,根据数据确定包括状态转移概率在内的相关参数,且默认同一时间段的参数不变。按 92 号汽油每升 7.58 元的价格,设每百公里油耗 7.5 升,则 $\alpha = 0.568 5$ 。根据深圳市出租车收费标准,式(7)中的参数设置如下。

白天(6:00—23:00): $f_0=10, \beta=2.7, \gamma=3.51, \delta=4.32, D_0=2, D_1=20, D_2=35$;

夜晚(23:00—6:00): $f_0=13, \beta=3.51, \gamma=4.56, \delta=5.62, D_0=2, D_1=20, D_2=35$ 。

3.2 计算性能

策略迭代算法通过调用 NumPy 库编写,在 python 3.8 平台上运行,所有计算在六核 3.0 GHz AMD Ryzen 5 4600H 处理器和 16 GB RAM 的计算机上进行,设置衰减系数 $\rho=0.5$,收敛准则 $\epsilon=0.01$,初始策略 $\pi_0(a|i)$ 为均匀随机策略,分 24 个时间段进行计算,序号分别为 1~24。

策略迭代的平均次数为 7 次,单次平均耗时 16.9 min(1 013.6 s),是采用雅克比迭代的 54.1%。实际场景中可根据历史数据提前计算出各时段、各状态所对应的最优路径策略,在之后对应的时段针对不同状态的车辆予以路径规划,计算性能如表 1 所示。

表 1 不同时间段的计算性能

时 段	单次策略迭代耗时/s		时 段	单次策略迭代耗时/s		时 段	单次策略迭代耗时/s	
	高斯-塞德尔迭代	雅克比迭代		高斯-塞德尔迭代	雅克比迭代		高斯-塞德尔迭代	雅克比迭代
1	892.4	1 686.7	9	1 018.5	1 858.0	17	1 072.8	1 955.4
2	846.4	1 581.7	10	1 089.3	2 050.5	18	1 000.7	1 802.1
3	935.6	1 767.4	11	1 134.0	2 047.0	19	1 145.7	2 164.3
4	871.0	1 574.5	12	1 051.6	1 984.2	20	979.7	1 775.0
5	948.0	1 707.9	13	1 146.1	2 150.0	21	982.6	1 798.1
6	924.4	1 748.0	14	1 021.1	1 936.9	22	1 019.8	1 843.4
7	947.6	1 769.7	15	1 025.9	1 869.4	23	1 162.4	2 186.0
8	912.7	1 662.4	16	1 082.9	1 960.7	24	1 115.4	2 083.6

图 4 展示了在时段 6(7:00—8:00)、时段 14(15:00—16:00)和时段 22(23:00—24:00)3 个时段下,价值函数的最大相对差异与迭代次数的关系,在进行到第 3 次策略迭代时,相对差异均下降到 0.01 以下。

3.3 计算结果

图 5 分别展示了各时段的接送客事件总次数,以及各时段的总最优状态价值。两者呈现相似的趋势,计算得到两组数据的相关系数为 0.812,属于高度相关,说明当接送

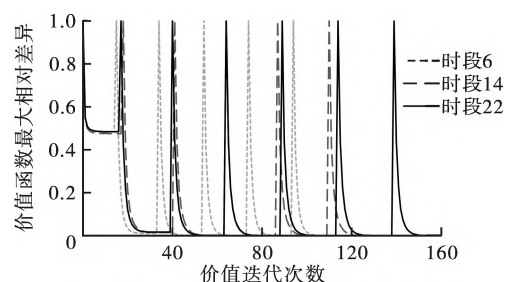


图4 最优状态价值函数的最大相对差异

客事件总数增加时,最优状态价值随之提升的可能性非常大。

使用式(11)标准化各时段的最优状态价值函数的计算结果,输出每个网格标准化后的最优状态价值 $V^*(i)_{\text{std}}$ 。

$$V^*(i)_{\text{std}} = \frac{V^*(i) - V_{\min}^*}{V_{\max}^*} \quad (11)$$

式中, V_{\min}^* 表示所有网格中最优状态价值的最小值; V_{\max}^* 表示所有网格中最优状态价值的最大值,选取时段 6 (7:00—8:00) 的计算结果进行展示如图 6 所示。

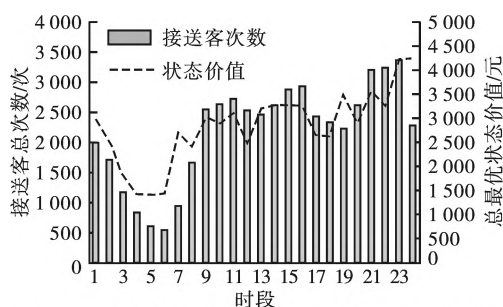


图5 不同时段接送客总次数与总最优状态价值



图6 时段6的最优状态价值标准化计算结果

3.4 仿真评价

为评价模型的最优路径决策效果,将其与随机游走和全局热点两种启发式算法在单位距离收益和载客里程占比两方面进行比较,现将这两种启发式算法具体内容简述如下。

1) 随机游走 车辆从初始位置开始行驶,随机选择周围的可达网格,一旦匹配到乘客,便以最短路运送乘客至目的地,送客后在下客网格继续重复前述过程。

2) 全局热点 车辆从初始位置开始行驶,以最短路朝着匹配概率最高的网格行驶,到达后在区域内部随机巡游,一旦匹配到乘客,便以最短路运送乘客至目的地,送客后在下客网格继续重复前述过程,车辆在前往全局热点的途中也存在匹配到乘客的可能。

分别对作者提出的最优路径策略、随机游走和全局热点 3 种算法在每个时段内进行 100 次仿真,总计 7 200 次。单次仿真中车辆的累计行驶距离若达到 50 km,则停止此次仿真并输出结果,仿真车辆的起点按照历史实际分布情况随机生成。图 7 和图 8 分别描述了 3 种策略在全天的单位距离平均收益和载客里程占比的情况。

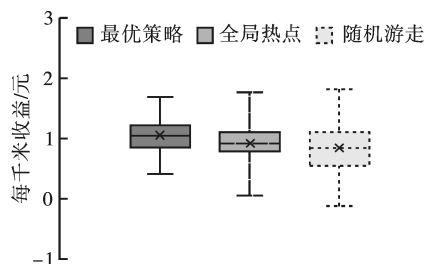


图7 平均单位距离收益

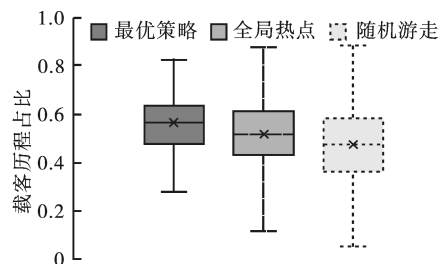


图8 平均载客里程占比

在收益方面,最优策略在全天的平均单位距离收益为 1.05 元/km,随机游走和全局热点分别是 0.86 元/km 和 0.93 元/km,最优策略相对于其他两种算法在平均单位距离收益上分别提高了 22.1% 和 12.9%。在载客里程占比方面,最优策略、随机游走和全局热点在全天的平均载客里程占比分别为 56.2%、47.3% 和 50.9%,最优策略相较于其他两种策略分别提升 18.8% 和 10.4%。同时最优策略仿真结果具有更小的四分位差,表明作者提出的方法对于提高空驶出租车的长期收益和载客率方面具有较好的效果。

4 结 论

作者研究了单辆空驶出租车路径优化问题,旨在提高出租车的长期收益,通过网格化缩小状态空间减少数据波动和稀疏性问题,同时配合策略迭代加速收敛。得到的主要结论如下:

a. 对路网进行网格化处理后,所有网格的最优状态价值之和与对应时段的接送客事件发生的总次数高

度相关,相关系数高达 0.812。

b.使用高斯-赛德尔迭代对策略评估部分进行加速,单次策略迭代的平均耗时是采用雅克比迭代的 54.1%,速度是采用雅可比迭代的 1.85 倍,显示算法求解具有较快的收敛效果。

c.所提出的最优路径策略的平均单位距离收益比随机游走和全局热点算法分别提高了 22.1%和 12.9%,载客距离占比分别提高了 18.8%和 10.4%,同时最优策略仿真结果具有更小的四分位差,显示模型对空驶出租车的长期收益具有较好的提升效果。

最后,在多个方面还可以进行详细补充和深入研究:首先,作者使用非时变道路网络,将研究时域分为 1 h 长的固定时间窗来计算行程时间,未来可考虑在时变网络和滑动时间窗的情境下进行更深入的研究;其次,作者将网格尺寸设置为 500 m,未来可使用如平衡二叉树等启发式算法确定尺寸,或探索更加有效的手段降低数据稀疏性的影响,对不同聚类方式下的路径优化结果进行差异性和敏感性分析;再次,模型中动作的即时奖励只与当前所处的网格位置有关,与上一个驶入网格的位置无关,但对于多个网格构成的序列,之间的转换对应到实际路网中是否可达,还需要根据上一个驶入网格的位置增加状态的维度,并探索在弱马尔可夫性下解决问题的方法;最后,作者在研究车辆路径规划时并没有考虑出租车司机对路径的偏好和对策略的服从性,且考虑的场景为单智能体的决策问题,因此如何在考虑出租车司机路径偏好或策略服从性的多智能体场景下,科学有效地规划车辆的路径,是十分具有价值的研究方向。

参考文献

- [1] Hu X, Gao S, Chiu Y C, et al. Modeling Routing Behavior for Vacant Taxicabs in Urban Traffic Networks[J]. Transportation Research Record Journal of the Transportation Research Board, 2015, 2284(2284): 81-88.
- [2] 龙雪琴, 周 萌, 赵 欢, 等. 基于网络核密度的网约车上下客热点识别[J]. 交通运输系统工程与信息, 2021, 21(3): 86-93.
- [3] 曲昭伟, 王 鑫, 宋现敏, 等. 基于出租车 GPS 大数据的城市热点出行路段识别方法[J]. 交通运输系统工程与信息, 2019, 19(2): 238-246.
- [4] 郑林江, 赵 欣, 蒋朝辉, 等. 基于出租车轨迹数据的城市热点出行区域挖掘[J]. 计算机应用与软件, 2018, 35(1): 1-8.
- [5] 高 瞻, 余 辰, 向郑涛, 等. 基于网格化的出租车空载寻客路径推荐[J]. 计算机应用与软件, 2019, 36(5): 281-288.
- [6] Yuan N J, Zheng Y, Zhang L, et al. T-finder: A Recommender System for Finding Passengers and Vacant Taxis[J]. IEEE Transactions on Knowledge and Data Engineering, 2012, 25(10): 2390-2403.
- [7] Dong H, Zhang X, Dong Y, et al. Recommend a Profitable Cruising Route for Taxi Drivers[C]//17th International IEEE Conference on Intelligent Transportation Systems (ITSC). [S.l.]: IEEE, 2014: 2003-2008.
- [8] Qu M, Zhu H, Liu J, et al. A Cost-effective Recommender System for Taxi Drivers[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: Association for Computing Machinery, 2014: 45-54.
- [9] Yu X, Gao S, Hu X, et al. A Markov Decision Process Approach to Vacant Taxi Routing with E-hailing[J]. Transportation Research Part B: Methodological, 2019, 121: 114-134.
- [10] Qian S, Zhu Y, Li M. Smart Recommendation by Mining Large-scale GPS Traces[C]//2012 IEEE Wireless Communications and Networking Conference (WCNC). [S.l.]: IEEE, 2012: 3267-3272.
- [11] Tang H, Kerber M, Huang Q, et al. Locating Lucrative Passengers for Taxicab Drivers[C]//Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: Association for Computing Machinery, 2013: 504-507.
- [12] 杨 昌. 出租车位置数据的移动性挖掘与推荐方法研究[D]. 长沙: 湖南大学, 2017.
- [13] 徐志锋, 于 欢, 何 兵. 基于 GPS 轨迹数据的深圳市居民出行热点研究[J]. 科技创新与应用, 2020, 10(1): 50-52.
- [14] 郑运鹏, 赵 刚, 刘 健. 基于出租车 GPS 数据的交通热区识别方法[J]. 北京信息科技大学学报: 自然科学版, 2016, 31(1): 43-47.
- [15] Bellman R. A Markovian Decision Process[J]. Indiana University Mathematics Journal, 1957, 6(4): 679-684.
- [16] Zhang D, Zhao J, Zhang F, et al. UrbanCPS: A Cyber-physical System Based on Multi-source Big Infrastructure Data for Heterogeneous Model Integration[C]//Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems. New York: Association for Computing Machinery, 2015: 238-247.