

基于深度强化学习的车联网汽车感知任务分配

向朝参^{1),(2)} 李耀宇^{1),(2)} 冯 亮^{1),(2)} 陈 超^{1),(2)} 郭松涛^{1),(2)} 杨盘隆³⁾

¹⁾(重庆大学计算机学院 重庆 400044)

²⁾(信息物理社会可信服务计算教育部重点实验室 重庆 400044)

³⁾(中国科学技术大学计算机科学与技术学院 合肥 230026)

摘 要 随着智能/辅助/自动驾驶以及电动汽车技术的飞速发展,当前车联网汽车配备越来越多的传感器,拥有越来越强大的计算、存储和通信能力.作为车联网汽车中重要一类,出租车、滴滴等网约车(Mobility-on-demand vehicles)具有城市覆盖规模大和粒度细,以及空闲时间充足的优点.因此,利用这些车联网汽车现有的移动感知设备能够为城市大规模、细粒度、低成本的感知提供很好的机会.本文以出租车、滴滴等这一类重要的车联网汽车为研究对象,重点研究它们的感知任务分配问题,主要面临两方面挑战:一是车联网汽车载客收益(如出租车载客收入)以及汽车/感知任务分布具有时空动态性,导致车联网汽车的感知成本具有高时空动态性且很难建模和学习.二是感知任务的最优分配问题是NP-hard问题,具有指数级时间复杂度.同时,高移动性的车联网汽车对任务的实时分配要求很高.因此,如何对大规模的车联网汽车实现任务的最优实时分配具有挑战性.为了解决这两个挑战,本文提出基于深度强化学习的车联网汽车感知任务分配方法.通过深度强化学习模型对高时空动态性的汽车感知成本进行精确学习,然后基于学习结果进行感知任务的实时最优分配.具体地,针对挑战一,利用基于双注意力机制的循环神经网络挖掘汽车载客收益的时空相关性,并结合驾驶耗费模型,学习车联网汽车的感知成本.针对挑战二,首先通过问题等价转换和理论分析,证明该任务的最优分配问题具有单调子模目标函数和 q -独立系统约束条件.然后,基于子模优化理论,联合考虑整体收益和边际效益,提出多项式时间复杂度的近似最优分配算法(近似率为 $1/\lceil 2+c_{\max}/c_{\min} \rceil$),其中 $c_{\max}$ 和 $c_{\min}$ 分别表示所有感知成本的最大和最小值.最后,基于两个大规模的车联网汽车数据集(重庆市,约12493辆车;纽约市,约超过1.13亿个行程),对所提方法进行深入全面的评估.实验结果表明,所提方法比7种对比方法平均提高载客收益的预测精度25.1%,提高任务分配的总感知效益37.7%.同时,面向城市道路违规停车监测应用,构建车联网汽车感知原型系统.基于该系统验证了所提方法的可行性和实际应用价值.

关键词 车联网汽车;感知任务分配;深度强化学习;子模优化;循环神经网络

中图法分类号 TP311 **DOI号** 10.11897/SP.J.1016.2022.00918

Near-Optimal Vehicular Crowdsensing Task Allocation Empowered by Deep Reinforcement Learning

XIANG Chao-Can^{1),(2)} LI Yao-Yu^{1),(2)} FENG Liang^{1),(2)} CHEN Chao^{1),(2)}

GUO Song-Tao^{1),(2)} YANG Pan-Long³⁾

¹⁾(College of Computer Science, Chongqing University, Chongqing 400044)

²⁾(Key Laboratory of Dependable Service Computing in Cyber Physical Society (Chongqing University), Ministry of Education, Chongqing 400044)

³⁾(School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026)

Abstract With the rapid development of intelligent vehicular technologies, such as Self-driving systems and Advanced Driver Assistance systems, off-the-shelf intelligent vehicles are equipped

收稿日期:2020-10-30;在线发布日期:2021-09-09. 本课题得到国家自然科学基金项目(62172063,61872447)资助. 向朝参,博士,副教授,博士生导师,中国计算机学会(CCF)会员,主要研究领域为群智感知、人工智能、普适计算. E-mail: xiangchaocan@cqu.edu.cn. 李耀宇,硕士研究生,主要研究方向为车联网汽车感知、深度强化学习. 冯 亮,博士,教授,主要研究领域为强化学习、人工智能. 陈 超,博士,教授,主要研究领域为城市计算、人工智能. 郭松涛(通信作者),博士,教授,主要研究领域为边缘计算、感知任务最优分配. E-mail: guosongtao@cqu.edu.cn. 杨盘隆,博士,教授,主要研究领域为群智感知、无源感知.

with more and more sensors, including GPS, camera, Lidar, etc., thus possessing powerful capabilities of computation and communication along with large-scale storage capacity. As an important kind of the intelligent vehicle, the Mobility-On-Demand (MOD) vehicles (such as Uber, DiDi, and connected taxis) have large-scale, fine-grained coverage in cities along with non-negligible amounts of spare time. Hence, utilizing their available sensors provides promising opportunities in achieving large-scale, fine-grained, and low-cost vehicular crowdsensing for smart cities. As a result, this paper focuses on these MOD vehicles and studies how to optimally allocate the vehicular crowdsensing tasks for the MOD vehicles. It chiefly involves two main challenges: (1) Both the distributions of the MOD vehicles and the sensing tasks have spatial-temporal differences. Also, the pick-up earnings of MOD vehicles vary with the location and time. Hence, it renders the sensing cost highly dynamic in both temporal and spatial dimensions. Even worse, such sensing cost is hard to model because of its highly dynamic nature. (2) The optimal sensing task allocation is a NP-hard problem, which has exponential time complexity. Furthermore, owing to the high mobility of the vehicles, it requires real-time task allocation in vehicular crowdsensing. To address these challenges, we propose a deep reinforcement learning-empowered near-optimal task allocation method for vehicular crowdsensing. We utilize deep reinforcement learning to extract the highly dynamic sensing cost of vehicles, which is fed back to optimally allocate the sensing tasks for each MOD vehicle. Specifically, targeting the first challenge, we deploy the Encoder-Decoder Recurrent Neural Network based on dual attentions (including the spatial attention and the temporal attention) to extract the spatial-temporal correlations of pick-up earnings, which are then used to learn the sensing cost according to the driving cost model. Furthermore, through the equivalent problem transformation, we prove that the task allocation problem has a submodular objective function and a q -dependent constraint. Hence, based on the sub-modularity theory, we propose a near-optimal task allocation algorithm, jointly considering the total utility and marginal utility. It is proved to achieve a $1/\lceil 2+c_{\max}/c_{\min} \rceil$ -approximation ratio in polynomial time, where $c_{\max}$ and $c_{\min}$ represent the maximal and minimal values of the sensing costs for all the vehicles, respectively. Finally, we exploit two large-scale datasets to evaluate the performance of the proposed method. One dataset is about 12493 MOD vehicles in Chongqing City, China, while the other is about 113 million vehicle trips in New York City, America. The results demonstrate that our method averagely improves the prediction accuracy of pick-up earnings and the allocation utility of sensing tasks by 25.1% and 37.7%, respectively, compared with seven baselines. Moreover, we implement a prototype system for on-road illegal parking detection, i. e., leveraging the smartphone sensor (such as camera and GPS) of massive MOD vehicles to detect the on-road illegal parking events when driving on roads. Based on this system, we validate the proposed method is feasible and significant in practical applications.

Keywords vehicular crowdsensing; sensing task allocation; deep reinforcement learning; sub-modularity; recurrent neural network

1 引言

随着智能汽车、辅助/无人驾驶汽车、电动汽车的飞速发展和爆炸式普及,当前汽车集成越来越多的传感器,拥有越来越强大的感知、计算和通信

能力^[1-4]. 例如,根据公安部交管局最新统计,截至2020年9月,全国机动车保有量达3.65亿辆(约占全国总人数的27%)^[5]. 同时,当前普通汽车大约有60~100种传感器^[3]. 此外,面向汽车的无线通信技术迅猛发展、形式多样,如车间通信(V2V)、车与设施通信(V2I)以及OBD-II接口通信等^[6]. 利用众多

普通汽车现有的传感器进行感知,通过已部署的移动网络(如 4G、5G、WiFi 和车联网等)进行数据传输,形成一种新的感知模式,即车联网汽车感知(Vehicular Crowdsensing)^[3,7-10].近年来,涌现很多基于车联网汽车感知的实际应用,如道路路面条件感知^[11]、自动驾驶训练^[12]和城市道路地图实时构建^[13]等.同时,也产生很多实际的车联网汽车感知在线平台,如 Payver^[14]和 Chffr^[15].

作为车联网汽车中重要一类,出租车和滴滴等网约车^[16-18]大量普及和被广泛使用,为车联网汽车感知带来两个独特优势:(1)覆盖面积广、粒度细:大量的出租车、滴滴汽车运行在各个城市^[19-21].根据 2019 年报告^[22],中国四百多个城市已有约 2100 万名滴滴司机.同时,这些汽车覆盖范围广、粒度细.通过分析美国纽约市的出租车数据集^[23](详细介绍见 5.1 节),如图 1 所示,出租车几乎覆盖整个纽约市主要城区的道路;(2)大量的空闲时间:大部分出租车等汽车司机在送达上一个乘客到达后,一般都需要花一定时间等待或者寻找下一个乘客,被称为空闲时间^[18].因此,车联网汽车司机拥有充足的空闲时间执行感知任务.

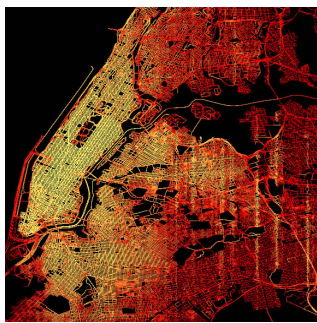


图 1 纽约市车联网汽车(出租车)的道路覆盖热力图

综上,由于这些网约汽车覆盖面积广、粒度细、空闲时间多,因此,利用它们现有的移动感知设备,可以实现城市大规模、细粒度、低成本的车联网汽车感知系统^[9,16,19-21].

本文以出租车、滴滴和 Uber 等这一类重要的车联网汽车^①为例,重点研究车联网汽车感知任务分配.具体地,在车联网汽车感知系统中^[24-26],首先,感知平台(简称平台)发布感知任务;然后,参与的车联网汽车在载客空闲时间内向平台提交感知任务请求和当前状态信息;其次,平台根据车联网汽车状态计算任务的感知成本,并根据成本分配感知任务以最大化平台感知效益;最后,车联网汽车根据任务分配完成感知任务,并得到平台给予的感知报酬(也称

为感知收益).因此,车联网汽车除了载客收益,还能获得额外的感知收益,从而提高司机的总收益,激励他们加入到车联网汽车感知系统中^[7,27-28].但是,实现上述系统的感知任务最优分配需要解决以下两方面主要挑战:

(1)感知成本具有高时空动态性.车联网汽车执行感知任务的成本主要由两方面决定:①基础驾驶耗费,即从区域 A 驾驶到任务区域 B 的耗费,如图 2(a)所示.由于该耗费与驾驶距离和时间相关,因此随着车联网汽车和感知任务的位置时空动态变化而变化;②载客隐性激励,即司机执行感知任务所带来的载客收益变化.例如,如图 2(b)所示,司机从偏远区域 A 到核心区域 B 执行感知任务.由于区域 B 的载客收益比区域 A 高(如载客概率高或单程平均收益高),因此,司机从区域 A 到 B 执行感知任务可能带来载客收益的增加(被称为载客隐性激励),从而降低司机执行感知任务的成本^[7,27].由于各个区域的载客收益具有高度的时空动态变化且难以建模^[17,19],导致感知成本具有复杂的时空动态性.

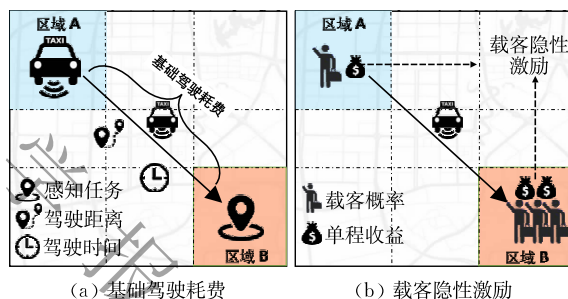


图 2 车联网汽车感知成本示意图

(2)感知任务分配具有高计算复杂性(NP-hard).最优感知任务分配(即在一定感知成本预算内最大化平台的感知效益)在 3.3 节被证明是 NP-hard 问题,具有指数级的计算时间复杂度.进一步地,由于车联网汽车感知系统中汽车和感知任务的数量规模大、任务分配的实时性要求高^[17,19-20],因此,面向大规模车联网汽车进行实时的最优感知任务分配具有挑战性.

为了解决上述两个挑战,本文提出基于深度强化学习的车联网汽车感知任务分配方法,通过深度强化学习模型对高时空动态性的汽车感知成本进行精确学习,然后基于学习结果实现感知任务的实时最优分配.具体包括以下两部分:

① 本文主要研究出租车和滴滴等这一类重要车联网汽车的感知任务分配.为了表述简洁,在文章剩余部分,若没有特别说明,将车联网汽车、出租车、滴滴汽车、汽车等名词混用.

(1) 基于双注意力机制循环神经网络的感知成本学习. 本文利用空间与时间注意力机制的循环神经网络(Recurrent Neural Network, RNN), 分别学习载客收益的空间和时间相关性, 从而准确地预测载客隐性激励, 并基于驾驶耗费模型, 学习车联网汽车的感知成本.

(2) 基于子模优化理论的次优任务分配. 首先通过问题等价转化和理论分析, 证明该分配问题具有单调子模目标函数和 q -独立系统约束条件. 然后, 基于分析结果, 利用子模优化理论(Sub-modularity), 提出联合总体收益和边际效益的任务分配算法, 能够在多项式时间复杂度内获得具有理论下界的近似最优解.

综上所述, 本文主要具有四方面的贡献:

(1) 据作者目前调研, 本文是第一个提出基于深度强化学习的车联网汽车感知任务分配框架. 利用深度学习模型对高时空动态、难建模的汽车感知成本进行学习, 然后基于学习结果对感知任务进行实时最优分配.

(2) 提出基于深度学习的感知成本学习方法. 利用时空双注意力机制的循环神经网络预测载客隐性激励, 并结合驾驶耗费模型学习满足个体理性的任务感知成本.

(3) 提出具有理论下界的近似最优任务分配算法. 通过问题转化和理论分析, 利用子模优化理论, 联合考虑总体收益和边际效益, 在 $O(N^2 M)$ 时间复杂度内实现近似比率为 $1/\lceil 2 + c_{\max}/c_{\min} \rceil$ 的近似最优分配, 其中 N 和 M 分别表示汽车和感知任务个数, $c_{\max}$ 和 $c_{\min}$ 分别表示所有感知成本的最大和最小值.

(4) 基于实际大规模车联网汽车数据集的实验评估和原型系统实现. 利用出租车重庆市数据集(约 12 493 辆车)和纽约市数据集(超过 1.13 亿个行程), 基于 7 种对比方法, 对本文所提算法的性能进行深入全面的评估. 实验结果表明, 本文的载客收益预测精度对比方法提高 25.1%, 任务分配的总感知效益提高 37.7%. 同时, 构建面向城市道路违规停车监测的车联网汽车感知原型系统, 验证了所提系统和方法的实际应用性和可行性.

本文第 2 节介绍和总结国内外相关研究现状; 第 3 节介绍系统模型和问题描述; 第 4 节详细介绍基于深度强化学习的感知任务分配方法; 第 5 节展示实验结果; 最后在第 6 节总结全文.

2 相关工作

近年来关于群智感知任务分配和激励机制已有大量研究, 其中大部分研究关注传统群智感知(Crowdsensing, Crowdsourcing, or Participatory Sensing), 主要考虑静态传感器和人的移动性^[29-31]. 相较于传统群智感知, 车联网汽车感知(Vehicular Crowdsensing)具有更高移动性、更大覆盖范围等优点^[8]. 下面将重点综述车联网汽车感知任务分配相关研究, 其他相关工作可以参考文献^[32].

当前大部分车联网汽车感知任务分配研究都是设计任务分配策略, 以最大化感知质量或最小化平台耗费. 例如, He 等人^[24]联合考虑车辆当前位置和汽车移动模型, 提出在有限预算内最大化时空覆盖的感知任务分配策略. Gao 等人^[25]考虑汽车用户的不可靠性(如虚报感知耗费等), 提出面向不确定性车联网汽车感知的可信任任务分配机制. 另外, Fan 等人^[33]提出一种新颖的车联网汽车感知任务分配系统, 联合考虑感知任务分配的可靠性和汽车移动轨迹的最优调度. Xu 等人^[7]提出一种车联网汽车感知任务分配算法, 在一定感知成本预算内最小化感知数据的时空分布与目标分布的差异性. 以上工作都未考虑 Mobility-On-Demand Vehicles(MOD 汽车)时空动态载客收益对感知任务分配的影响. 与之相反, 本文利用深度学习对载客收益的时空相关性进行学习, 从而实现感知任务的最优分配.

近年来, 也有关于 MOD 汽车载客收益的研究工作, 但是他们都主要研究这些汽车的载客任务分配, 即如何分配汽车与乘客乘车需求, 以最大化平台收益、司机收益和乘客满意度体验. 例如, Tang 等人^[34]利用深度强化学习方法, 提出多司机订单分配算法, 以提高司机收入和乘客的满意度. Xu 等人^[35]考虑订单分配对司机未来收益的影响, 利用强化学习方法, 以最大化平台和司机当前与未来收益的总和. Sühr 等人^[36]提出一种新的载客任务分配机制, 以实现司机和乘客的双边公平. 此类关于 MOD 汽车的工作只关注于汽车的传统载客任务, 忽略了车联网汽车强大的城市感知能力. 与之相反, 本文充分考虑 MOD 车联网汽车的载客任务和感知任务, 巧妙地利用它们之间相互作用所产生的隐性激励, 从而设计既能降低任务感知成本, 又能满足汽车用户个体理性的感知任务分配策略.

总结: 与当前研究工作不同, 本文采用深度强化

学习框架,利用深度学习模型对车联网汽车的载客收益和感知成本进行学习,然后基于学习结果利用子模优化理论实现感知任务的最优分配,从而实现MOD车联网汽车载客与感知之间的很好融合。

3 系统建模和问题描述

3.1 相关概念

定义 1. 城市子区域和时段. 按照一定粒度,将城市均匀划分成 Z 个互不相交的子区域,用 z_i 表示第 i 个子区域. 与之类似,将整个时间区间划分成 T 个长度相同的时段,每个时段用 t 表示.

定义 2. 载客收益. 表示车联网汽车司机通过接送乘客所获得的收益. 该收益由司机所在区域接客概率和单程平均收益决定. 用 p_i^t 和 r_i^t 分别表示第 t 时段在区域 z_i 的接客概率和单程平均收益. 因此,车联网汽车的载客收益可表示为

$$\mathbf{R}_i^t = (p_i^t, r_i^t), i \in \{1, \dots, Z\}, t \in \{1, \dots, T\} \quad (1)$$

定义 3. 载客收益热力图. 由于在同一时段不同区域的载客收益具有差异性. 用载客收益热力图 $\mathbf{H}^t$ 表示第 t 时段所有区域的载客收益,即

$$\mathbf{H}^t = \{\mathbf{R}_i^t | \mathbf{R}_i^t = (p_i^t, r_i^t), \forall i \in \{1, \dots, Z\}\} \quad (2)$$

因此,车联网汽车在 h 个历史时段的载客收益可表示为 h 帧热力图,即 $\{\mathbf{H}^1, \dots, \mathbf{H}^h\}$.

定义 4. 载客隐性激励. 表示第 t 时段将车联网汽车从原区域 z_i 调度到区域 $z_{i'}$ 所产生载客收益的增值期望,可表示为

$$I_{ii'}(t) = E_t[p_{i'}^t \cdot r_{i'}^t - p_i^t \cdot r_i^t] \quad (3)$$

为便于易读,表 1 给出本文主要符号的含义.

表 1 主要符号表说明

符号	符号含义
Z, z_i	城市子区域数量,第 i 个子区域
T, t	时段数量,时段 t
$p, r, \mathbf{R}$	载客概率,单程平均收益,载客收益
$\mathbf{H}, \hat{\mathbf{H}}$	载客收益热力图,预测载客收益热力图
c, c^B, I	感知成本,基础驾驶耗费,载客隐性激励
M, τ_j	感知任务数量,第 j 个感知任务
N, u_k	车联网汽车数量,第 k 辆车联网汽车
l, ρ	感知任务(车联网汽车)位置,完成任务概率
v, v^b, v^e	有效时间区间,开始时间,结束时间
$\mathbf{X}, B$	任务分配策略,预算约束
$\Theta(\cdot), \Phi(\cdot)$	载客收益预测函数,感知成本计算函数
$\Psi(\cdot), U(\cdot)$	驾驶耗费模型函数,平台感知效益函数
$\mathcal{V}, \mathcal{S}, e$	任务分配全集,策略,元素

3.2 基于深度强化学习的系统模型

本文提出基于深度强化学习的车联网汽车感知任务分配系统模型. 如图 3 所示,该模型采用深度强

化学习的基本框架^[37-39],以车联网汽车感知平台作为智能体,基于复杂时空动态的车联网汽车感知环境(1),对感知成本(即某个汽车完成某个感知任务的成本)进行深度学习(2). 然后,基于学习结果,采用强化学习模型进行任务分配(3),从而最大化平台的感知效益. 下面将详述模型的三个主要模块.

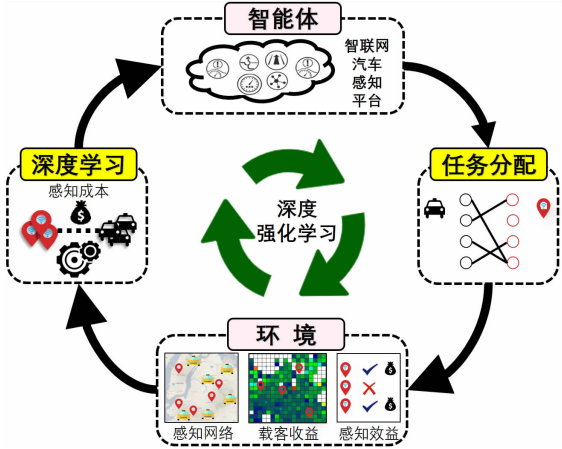


图 3 基于深度强化学习的车联网汽车感知任务分配系统框架

(1) 车联网汽车感知环境. 环境主要包括感知网络、载客收益和感知效益. ① 感知网络. 包括感知任务和车联网汽车的时空分布. 假设在城市不同区域存在 M 个感知任务. τ_j 表示第 j 个感知任务; l_j 和 z_j 分别表示 τ_j 的位置和所在的区域. 由于任务具有时效性,需在一定时间内完成,用 $v_j = (v_j^b, v_j^e)$ 表示任务 τ_j 的有效时间区间,其中 v_j^b 和 v_j^e 分别表示有效区间的开始和结束时间. 假设存在 N 个参与感知任务的车联网汽车,用 u_k 表示第 k 个参与汽车. 用 l_k 和 v_k 分别表示该汽车所在的位置(其所在区域用 z_k 表示)和可参与感知任务的时间区间,即 $v_k = (v_k^b, v_k^e)$. 该时间区间可由汽车司机根据自身经验和实际情况预先提交和预定,类似于滴滴、出租车司机接收载客订单预定^[40],也可基于车联网汽车用户历史空闲数据进行预测^[41]. ② 载客收益. 如定义 3 所示,由于车联网汽车的载客收益具有时空动态性,用载客收益热力图 $\mathbf{H}^t = \{\mathbf{R}_i^t | \mathbf{R}_i^t = (p_i^t, r_i^t), \forall i \in \{1, \dots, Z\}\}$ 来表示. ③ 感知效益. 根据任务分配决策,各个车联网汽车执行各个感知任务所获得的感知总效益即为平台感知效益.

(2) 深度学习. 车联网汽车感知平台对环境进行深度学习,得到各个汽车用户完成感知任务的成本. 具体地,平台基于历史载客收益热力图 $\{\mathbf{H}^1, \dots, \mathbf{H}^h\}$ 学习预测每个区域在未来 T 个时间段的载客收益热力图. 用 $\hat{p}_i^t$ 和 $\hat{r}_i^t$ 分别表示 p_i^t 和 r_i^t 的预测值. 用 $\hat{\mathbf{H}}$

表示未来 T 个时间段的载客收益预测热力图集合, 即 $\hat{\mathbf{H}} = \{\hat{\mathbf{H}}^t | h+1 \leq t \leq h+T\}$, 其中 $\hat{\mathbf{H}}^t$ 表示第 t 时段载客收益的预测热力图. 基于 $\hat{\mathbf{H}}$, 根据汽车用户 u_k 和感知任务 τ_j 的时空位置, 估计用户的载客隐性激励, 从而学习得到 u_k 完成 τ_j 的感知成本 c_{kj} .

(3) 任务分配. 车联网汽车感知平台根据当前学习结果进行感知任务分配. 具体地, 在有限的感知成本预算 B 条件下, 如何将 M 个感知任务分配给 N 个车联网汽车, 以最大化平台感知效益. 用 x_{kj} 表示是否将 τ_j 分配给 u_k . 若分配, 则 $x_{kj} = 1$, 反之 $x_{kj} = 0$. $\mathbf{X} = \{x_{kj} | \forall k \in \{1, \dots, N\}, \forall j \in \{1, \dots, M\}\}$ 表示任务分配策略. 用 ρ_{kj} 表示车联网汽车司机 u_k 完成被分配任务 τ_j 的概率. 该概率取决于司机的可靠性以及车辆移动的不确定性, 可基于大量车联网汽车用户历史数据进行学习^[25]. 说明: 虽然本文以基本的感知任务效益模型作为例子, 即假设每个感知任务的效益相同^[26, 33], 但是通过简单扩展就能适用于更复杂的模型^[42]. 因此, 基于分配策略 $\mathbf{X}$, 平台的感知效益 $U(\mathbf{X})$ 可表示为

$$U(\mathbf{X}) = \sum_{j=1}^M \left[1 - \prod_{k=1}^N (1 - \rho_{kj} x_{kj}) \right] \quad (4)$$

其中式(4)中平台感知效益表示感知任务完成数量的期望. $\rho_{kj} x_{kj}$ 和 $1 - \prod_{k=1}^N (1 - \rho_{kj} x_{kj})$ 分别表示车联网汽车 k 完成感知任务 j 的概率, 以及感知任务 j 被所有车联网汽车的完成概率.

3.3 研究问题

基于以上系统模型, 本文的研究问题主要包括: (1) 基于复杂动态车联网汽车感知环境的感知成本学习: 如何基于历史载客收益热力图, 预测城市各个区域未来 T 个时间段的载客收益热力图 $\hat{\mathbf{H}}$, 从而学习得到每个智联汽车用户对于每个感知任务的感知成本 $\{c_{kj}\}$; (2) 车联网汽车感知任务的最优分配: 基于任务感知成本学习结果, 如何将 M 个感知任务分配给 N 个车联网汽车, 以在感知成本预算 B 内最大化平台感知效益 $U(\mathbf{X})$. 具体地, 可形式化表示如下:

$$\text{Max } U(\mathbf{X}) = \sum_{j=1}^M \left[1 - \prod_{k=1}^N (1 - \rho_{kj} x_{kj}) \right] \quad (5)$$

$$\text{s. t. } \hat{\mathbf{H}} = \Theta(\mathbf{H}^1, \dots, \mathbf{H}^T), \forall t \in \{1, \dots, T\} \quad (6)$$

$$c_{kj} = \Phi(\hat{\mathbf{H}}, l_k, v_k, l_j, v_j), \forall k, j \quad (7)$$

$$\sum_{k=1}^N \sum_{j=1}^M c_{kj} \rho_{kj} x_{kj} \leq B \quad (8)$$

$$\sum_{j=1}^M x_{kj} \leq 1, \forall j \in \{1, \dots, N\} \quad (9)$$

$$v_k^b \leq v_j^e, v_k^e \geq v_j^b, \forall k, j, x_{kj} = 1 \quad (10)$$

其中, 式(5)表示基于任务分配策略 $\mathbf{X}$ 的平台感知效益. 式(6)中 $\Theta(\cdot)$ 表示车联网汽车载客收益的预测函数. 式(7)中 $\Phi(\cdot)$ 表示基于载客收益预测结果的感知成本计算函数. 式(8)表示平台感知成本预算约束. 式(9)表示任务分配限制, 即在一个任务分配周期中(T 个时间段内), 每个车联网汽车至多分配一个感知任务. 由于车联网汽车的主要工作是接送乘客. 因此, 该约束主要是为了最大化地降低执行任务感知对汽车司机正常载客任务的影响. 约束(10)表示车联网汽车需要在任务有效期内完成感知任务, 即汽车空闲时间与感知任务有效期必须相交. 该约束主要用于筛选能够完成感知任务的候选汽车.

通过分析上述问题的难度, 可得定理 1.

定理 1. 车联网汽车感知任务的最优分配问题是 NP-hard 问题.

证明. 该问题可归约为一个经典的 NP-hard 问题, 即 0-1 背包问题. 具体证明过程可参考文献[43], 本文在此省略.

4 系统和算法设计

本文提出基于深度强化学习的车联网汽车感知任务分配系统, 如图 4 所示. 基于车联网汽车的载客收益热力图和感知任务的时空分布图, 通过深度强化学习, 实现感知任务的最优分配. 具体地, 该系统主要包括以下两部分:

(1) 基于双注意力机制 RNN 的任务感知成本学习(4.1 节). 由于车联网汽车载客收益具有高度的时间和空间相关性. 同时, 该时空相关性具有复杂的、难以建模的动态性(即不同时间、区域之间的相关性不同). 由于传统的统计学习方法^[44]很难挖掘学习这种复杂、动态的时空相关性, 因此, 本文提出基于双注意力机制 RNN 的任务感知成本学习方法, 利用时间、空间注意力机制分别学习高动态的时间和空间相关性. 首先, 基于城市区域拓扑结构建立

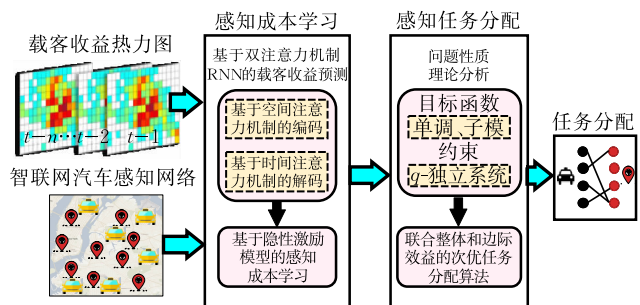


图 4 基于深度强化学习的感知任务最优分配方法架构图

邻接图,并利用本文深度学习预测方法预测车联网汽车的载客收益热力图.然后,基于载客收益的预测结果,通过建立隐性激励模型,学习车联网汽车完成感知任务成本.

(2) 基于子模理论的最优任务分配(4.2节).由于车联网汽车感知任务的最优分配是 NP-hard 问题,具有指数级时间复杂度.因此,为了解决该问题,首先,本文通过问题等价转换对该问题的性质进行理论分析.结果表明该问题具有单调子模目标函数和 q -独立系统约束条件.然后,基于子模优化理论^[43],提出联合总体收益和边际效益的任务分配算法,在多项式时间复杂度内获得 $1/(1+q)$ 性能下界的近似最优解.

4.1 基于双注意力机制 RNN 的感知成本学习

4.1.1 基于双注意力机制 RNN 的载客收益热力图预测

首先,如图 5(a)和图 5(b)所示,基于区域之间的拓扑结构图,构建地图拓扑邻接图,以表示各个子区域之间的相邻关系.用 $\mathcal{G}=(z, g)$ 表示该邻接图,其中每个顶点表示一个子区域 z_i ,每条边 $g=(z_i, z_{i'})$ 表示 z_i 和 $z_{i'}$ 之间的邻接关系.然后,根据区域拓扑邻接图,利用双注意力机制的循环神经网络,基于历史载客收益热力图,精确地预测每个子区域的载客收益,从而得到整个区域的载客收益热力图.

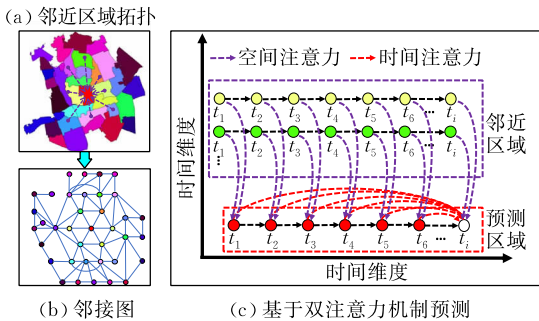


图 5 基于双注意力机制 RNN 的载客收益预测原理

下面以某个子区域 z_i 为例,详细介绍如何基于双注意力机制的循环神经网络预测载客收益 $\mathbf{R}_i'=(p_i', r_i')$, $i \in \{1, \dots, Z\}$, $t \in \{1, \dots, T\}$. 该方法利用基于长短期记忆的编码器-解码器架构(LSTM-based Encoder-Decoder)和时空注意力机制(Temporal-Spatial Attentions),主要包括基于空间注意力机制的编码和基于时间注意力机制的解码两部分.

(1) 基于空间注意力机制的编码:利用空间注意力机制将邻近区域的载客收益热力图编码为特征向量,称为隐藏状态.用 z_i 和 λ_i^i 分别表示第 i 个子区

域和该区域在第 t 时段的空间注意力参数.用 h_t 和 s_t 分别表示第 t 时段的隐藏状态和单元状态.根据第 $t-1$ 时段的隐藏状态 h_{t-1} 和单元状态 s_{t-1} ,可计算 λ_i^i 如下:

$$\lambda_i^i = \frac{\exp(w_i^t)}{\sum_{j=1}^Z \exp(w_j^t)} \quad (11)$$

其中 $w_i^t = d^T \tanh(W_s[h_{t-1}; s_{t-1}] + G_s Y_i + b_s)$, d^T , W_s , G_s 和 b_s 表示循环神经网络编码的学习参数^[45].

(2) 基于时间注意力机制的解码.利用基于 LSTM 的循环神经网络解码每个时段编码生成的隐藏状态,从而预测载客收益.同时,在解码过程中,利用时间注意力机制,从所有时段中挖掘学习得到最相关的隐藏状态.基于解码中第 $t-1$ 时段的隐藏状态 h'_{t-1} 和单元状态 s'_{t-1} ,在第 t 时段中每个编码生成隐藏状态的时间注意力权重参数为

$$\eta_t^k = \frac{\exp(\pi_t^k)}{\sum_{j=1}^T \exp(\pi_j^k)}, \quad k \in \{1, \dots, T\} \quad (12)$$

其中 $\pi_t^k = d^T \tanh(W_s[h_{t-1}; s_{t-1}] + G_t h_k + b_t)$, 且 $h_k = f_e(h_{k-1}, \hat{Y}_k)$, $\hat{Y}_k = (\lambda_k^1 Y_1^k, \dots, \lambda_k^Z Y_Z^k)$. f_e 表示 LSTM 的编码器参数, d^T , W_s , G_t 和 b_t 表示解码器的学习参数.

4.1.2 基于隐性激励模型的感知成本学习

基于 4.1.1 节中预测得到的车联网汽车载客收益热力图,提出基于隐性激励模型的感知成本学习模型.具体地,车联网汽车任务感知成本由完成该任务的驾驶耗费减去完成感知任务所带来的载客隐性激励.用 $c_{kj}(t)$ 、 $c_{kj}^B(t)$ 和 $I_{kj}(t)$ 分别表示第 t 时段 u_k 执行 τ_j 的感知成本、完成任务的驾驶耗费以及载客隐性激励.因此,车联网汽车用户 u_k 在第 t 时段完成感知任务 τ_j 的成本可表示为

$$c_{kj}(t) = c_{kj}^B(t) - I_{kj}(t) \quad (13)$$

其中驾驶耗费 $c_{kj}^B(t)$ 表示 u_k 从其当前位置 l_k 到感知任务位置 l_j 的驾驶成本,主要由它们之间的距离和通勤时间决定^[7,46].用 $d_{kj}(t)$ 和 $\Gamma_{kj}(t)$ 分别表示第 t 时段从 l_k 出发到达 l_j 的驾驶距离和通勤时间.因此,驾驶耗费可表示为 $c_{kj}^B(t) = \Psi(d_{kj}(t), \Gamma_{kj}(t))$, 其中 $\Psi(\cdot)$ 表示驾驶耗费模型函数,主要由车的类型和城市环境等因素决定^[7]. $I_{kj}(t)$ 表示 u_k 在第 t 时段完成感知任务 τ_j 的载客隐性激励,其模型如下:

载客隐性激励模型. 根据定义 4,载客隐性激励 $I_{kj}(t)$ 表示 u_k 从其原始区域 z_k 调度到感知任务区域 z_j 所带来的载客收益增加值的期望.用 $r_{kk}^R(t)$ 表示

未参加感知时 u_k 在原始区域 z_k 的载客收益期望; $r_{kj}^A(t)$ 表示当参加感知时 u_k 在感知任务所在区域 z_j 的载客收益期望. 因此, 载客隐性激励 $I_{kj}(t)$ 可表示为

$$I_{kj}(t) = r_{kj}^A(t) - r_{kk}^R(t) \quad (14)$$

$$r_{kk}^R(t) = \sum_{t=1}^L (\bar{p}_k^t p_k^t r_k^t) \quad (15)$$

$$r_{kj}^A(t) = \sum_{t=t_{kj}+1}^{t_{kj}+L} (\bar{p}_k^t p_j^t r_j^t) \quad (16)$$

$$\bar{p}_k^t = 1 + \text{sgn}(t-1) \left(\prod_{m=1}^{t-1} (1 - p_k^{t-m}) - 1 \right) \quad (17)$$

其中 $t_{kj} = \left\lfloor \frac{\Gamma_{kj}(t)}{t_0} \right\rfloor$. p_k^t 表示司机第 t 时段在区域 z_k 的载客概率, $\bar{p}_k^t$ 表示司机在第 t 时段前未载客的概. r_k^t 表示司机第 t 时段在区域 z_k 的载客平均收益. t_0 表示时段的单位长度, L 表示计算载客隐性激励时间周期的长度.

4.2 基于子模理论的最优任务分配算法

本文提出基于子模理论的联合整体和边际效益的最优任务分配算法. 首先基于问题等价转化, 通过理论分析证明该任务分配问题具有非负、单调和子模目标函数, 以及 q -独立系统的约束条件. 然后, 基于问题性质分析结果, 利用子模优化理论, 提出联合整体和边际效益的任务分配算法, 在多项式时间复杂度内得到具有理论下界的近似最优解.

4.2.1 问题性质分析

为了便于问题性质分析, 首先将该任务最优分配问题转化为集合函数最优化问题. 具体地, 全集为 $\mathcal{V} := \{(k, j) \mid \forall k \in \{1, \dots, N\}, j \in \{1, \dots, M\}\}$. $\mathcal{V}$ 表示任务分配变量 x_{kj} 到集合元素 $e = (k, j) \in \mathcal{V}$ 的映射集合. $x_{kj} = 1$ 表示从 $\mathcal{V}$ 中选择元素 (k, j) ; 反之, $x_{kj} = 0$ 表示从 $\mathcal{V}$ 中不选择该元素. $\mathcal{S}$ 表示已选定的车联网汽车和感知任务之间的配对集合, 即 $\mathcal{S} \subseteq \mathcal{V}$, 且 $\forall (k, j) \in \mathcal{S}, x_{kj} = 1$. 令 $\Omega(\mathcal{S}) := U(\mathbf{X})$. 式(5)、(8)~(10)中的车联网汽车感知任务的最优分配问题可等价转换为

$$\max_{\mathcal{S} \subseteq \mathcal{V}} \Omega(\mathcal{S}) \quad (18)$$

$$\text{s. t.} \quad \sum_{j: (k, j) \in \mathcal{S}} \mathbb{I}_{(k, j) \in \mathcal{S}} \leq 1, \forall k \in \{1, \dots, N\} \quad (19)$$

$$\sum_{(k, j) \in \mathcal{S}} c_{kj} \rho_{kj} \leq B \quad (20)$$

$$v_k^b \leq v_j^e, v_k^e \geq v_j^b, \forall (k, j) \in \mathcal{S} \quad (21)$$

其中, $\mathbb{I}$ 表示指示函数.

基于式(18)~(21)中的集合函数最优化问题, 下面将分析目标函数 $\Omega(\mathcal{S})$ 和约束(19)、(20)的性质.

(1) 目标函数分析

定义 5. 非负性、单调性和子模性^[43]. 基于任意有限全集 $\mathcal{V}$ 的集合函数 f 具有(1)非负性. 当且仅当 $\forall \mathcal{S} \subseteq \mathcal{V}, f(\mathcal{S}) \geq 0$ 且 $f(\emptyset) = 0$; (2)单调性. 当且仅当 $\forall \mathcal{S} \subseteq \mathcal{V}, \forall e \in \mathcal{V}/\mathcal{S}, f(\mathcal{S} \cup \{e\}) \geq f(\mathcal{S})$; (3)子模性. 当且仅当 $\forall \mathcal{S}_1 \subseteq \mathcal{S}_2 \subseteq \mathcal{V}, \forall e \in \mathcal{V}/\mathcal{S}_2, f(\mathcal{S}_1 \cup \{e\}) - f(\mathcal{S}_1) \geq f(\mathcal{S}_2 \cup \{e\}) - f(\mathcal{S}_2)$.

基于定义 5, 通过分析函数 $\Omega(\mathcal{S})$ 的性质, 可得引理 1.

引理 1. 式(17)中的目标函数 $\Omega(\mathcal{S})$ 是非负、单调和子模的.

证明. (1)非负性证明. $x_{kj} \in \{0, 1\}$ 且 $p_{kj} \in [0, 1]$, 根据式(5), 显然 $\Omega(\mathcal{S})$ 是非负的; (2)单调性证明. 假设 $e = \{k_e, j_e\}$, 将集合 $\mathcal{S}$ 拆分成 $\mathcal{S}_e$ 和 $\mathcal{S}/\mathcal{S}_e$, 其中 $\mathcal{S}_e = \{(k_1, j_e), \dots, (k_n, j_e)\}, n = |\mathcal{S}_e|$. 由式(5)可得 $\Omega(\mathcal{S}) = \Omega(\mathcal{S}_e) + \Omega(\mathcal{S}/\mathcal{S}_e)$ 且 $\Omega(\mathcal{S} \cup \{e\}) = \Omega(\mathcal{S}_e \cup \{e\}) + \Omega(\mathcal{S}/\mathcal{S}_e)$. 因此, $\Omega(\mathcal{S} \cup \{e\}) - \Omega(\mathcal{S}) = \Omega(\mathcal{S}_e \cup \{e\}) - \Omega(\mathcal{S}_e) = \rho_{k_e j_e} \prod_{(k, j) \in \mathcal{S}_e} (1 - \rho_{kj}) \geq 0$. 所以, $\Omega(\mathcal{S})$ 是单调的; (3)子模性证明. 与单调性证明类似, 将 $\mathcal{S}_1$ 和 $\mathcal{S}_2$ 分别拆分成 $\mathcal{S}_e$ 和 $\mathcal{S}_1 \setminus \mathcal{S}_e, \mathcal{S}'_e$ 和 $\mathcal{S}_2 \setminus \mathcal{S}'_e$. 由于 $\mathcal{S}_1 \subseteq \mathcal{S}_2, \mathcal{S}_e \subseteq \mathcal{S}'_e$, 根据单调性证明可得 $\Omega(\mathcal{S}_1 \cup \{e\}) - \Omega(\mathcal{S}_1) = \rho_{k_e j_e} \prod_{(k, j) \in \mathcal{S}_e} (1 - \rho_{kj})$, 且 $\Omega(\mathcal{S}_2 \cup \{e\}) - \Omega(\mathcal{S}_2) = \rho_{k_e j_e} \prod_{(k, j) \in \mathcal{S}'_e} (1 - \rho_{kj})$. 因此, $\prod_{(k, j) \in \mathcal{S}_e} (1 - \rho_{kj}) \geq \prod_{(k, j) \in \mathcal{S}'_e} (1 - \rho_{kj})$, $\Omega(\mathcal{S})$ 是子模的. 所以, $\Omega(\mathcal{S})$ 是非负、单调和子模的.

(2) 约束条件分析

定义 6. q -独立系统^[43]. 给定一个独立系统 $(\mathcal{A}, \mathcal{V})$, $\mathcal{A}$ 为所有可行解的集合. $\mathcal{S}_1, \mathcal{S}_2 \in \mathcal{A}$, 设 $\mathcal{S}_1$ 和 $\mathcal{S}_2$ 分别为基数 (即元素个数) 最小和最大的最大可行解 (即满足全部约束且无法再加入任何元素的解), $r(\mathcal{S}_1)$ 和 $r(\mathcal{S}_2)$ 分别表示 $\mathcal{S}_1$ 和 $\mathcal{S}_2$ 的基数. 若存在 q 满足 $\max_{\mathcal{S}_1, \mathcal{S}_2 \in \mathcal{A}} \frac{r(\mathcal{S}_2)}{r(\mathcal{S}_1)} \leq q$, 则称该独立系统为 q -独立系统.

引理 2. 约束条件(19)、(20)构成 q -独立系统, 且 $q = \lceil 1 + c_{\max}/c_{\min} \rceil$, 其中 $c_{\max}$ 和 $c_{\min}$ 分别表示所有感知成本的最大和最小值, 即 $c_{\max} = \max_{\forall k, j} c_{kj}, c_{\min} = \min_{\forall k, j} c_{kj}$.

证明. 由于所有满足问题约束条件的可行解子集也必定满足该约束条件, 因此所有可行解的子集也为可行解. 对于任意两个最大可行解 $\mathcal{S}_1$ 和 $\mathcal{S}_2$, 若向 $\mathcal{S}_1$ 中添加一个元素 $e \in \mathcal{S}_2/\mathcal{S}_1$, 需从 $\mathcal{S}_1$ 中取出集合 $\mathcal{S}'_1$, 使得 $(\mathcal{S}_1/\mathcal{S}'_1) \cup \{e\}$ 满足约束(19)、(20). 为满

足约束(19)(即一个车联网汽车最多分配一个感知任务),只需最多取出 1 个与 e 冲突的元素. 为满足约束(20),最多需从 S_1 中取出 $\lceil c_{\max}/c_{\min} \rceil$ 个元素. 循环上述操作,直到 $S_1 = S_2$. 因此,当最差情况时,换入 $|S_2|$ 个元素需至少取出 $q = \lceil 1 + c_{\max}/c_{\min} \rceil$ 倍数量的元素,即 S_2 基数为 S_1 的 q 倍. 由定义 6 可得,约束(19)、(20)构成 q -独立系统,且 $q = \lceil 1 + c_{\max}/c_{\min} \rceil$.

4.2.2 基于联合整体和边际效益的任务最优分配算法

根据引理 1 和 2 可得,任务最优分配问题是具有 q -独立系统约束的单调子模函数优化问题. 因此,为解决该 NP-hard 问题,可通过贪心迭代逐步得到近似最优解,主要包括两种策略:

(1) 基于整体收益的分配策略. 即每一步贪心地选择目标函数的增量最大的分配方案(即 $\arg\max_{e \in S} (\Omega(S \cup \{e\}) - \Omega(S))$),直至不满足预算. 由于该策略未考虑任务分配所对应的耗费,可能得到性价比很低的分配方案. 例如,预算约束为 100 元, A 司机完成 a 和 b 任务的概率分别为 90% 和 85%,成本分别为 100 元和 10 元; B 司机完成 a 任务的概率为 80%,成本为 20 元. 基于整体收益的分配方案是: 仅仅 A 司机执行 a 任务. 但是,最优任务分配方案是: A 司机执行 b 任务; B 司机执行 a 任务. 由此可见,忽视感知成本而单纯考虑整体收益的方法可能产生收益虽高但耗费相对更高的分配方案,从而获得较差的感知任务分配性能.

(2) 基于边际效益的分配策略. 即每一步贪心地选择边际效益最大的分配方案(即 $\arg\max_{e \in S} (\Omega(S \cup \{e\}) - \Omega(S))/c_e$). 虽然该策略在大部分情况下的性能都优于基于整体收益的分配策略,但是在某些情况下性能有可能更差. 例如,当预算约束为 100 元时, A 车完成 a 和 b 任务的概率分别为 40% 和 50%,耗费分别为 20 元和 90 元. 虽然分配 A 车完成 a 任务的边际效益远高于完成 b 任务,但分配 A 车完成 b 任务却能获得更好的性能. 这是由于基于边际效益的分配策略产生了 80 元的剩余预算. 因此,当预算较为充足时,基于边际效益的贪心分配可能比基于整体收益的性能更差.

根据上述分析,基于整体收益的分配策略和基于边际效益的分配策略各有不同的适用条件. 同时,在不同的场景条件下,它们的性能优劣各不相同. 因此,为了提升算法的鲁棒性和适用性,本文提出联合整体和边际效益的任务最优分配算法,如算法 1 所示. 该算法结合整体和边际效益,利用子模优化理论,能够在多项式时间复杂度内得到具有理论下界

的近似最优解. 具体地,算法 1 的 2~5 行基于整体收益的分配策略,贪心地迭代选择最大化整体收益且满足约束条件(19)~(21)的解. 然后,6~9 行基于边际效益的分配策略,贪心地迭代选择最大化边际效益且满足约束条件(19)~(21)的解. 最后,10~11 行选择两种策略的最优解作为最终分配方案返回.

算法 1. 联合整体和边际效益的任务最优分配.

输入: 感知成本 $\{c_{kj} \mid \forall k, j\}$, 预算 B

输出: 分配方案 $\mathbf{X} = \{x_{kj} \mid \forall k, j\}$

1. 初始化 $S_1 \leftarrow \emptyset, S_2 \leftarrow \emptyset, \mathbf{X} \leftarrow \{0\}$;
2. WHILE $\exists e \in (N \times M) / S_1, S_1 \cup \{e\}$ 满足 (19) ~ (21), DO
3. $e^* = \arg\max_{e: S_1 \cup \{e\} \text{ 满足 (19) ~ (21)}} \Omega(S_1 \cup \{e\})$;
4. $S_1 \leftarrow S_1 \cup \{e^*\}$;
5. END WHILE
6. WHILE $\exists e \in (N \times M) / S_2, S_2 \cup \{e\}$ 满足 (19) ~ (21), DO
7. $e^* = \arg\max_{e: S_2 \cup \{e\} \text{ 满足 (19) ~ (21)}} \frac{\Omega(S_2 \cup \{e\}) - \Omega(S_2)}{c_e}$;
8. $S_2 \leftarrow S_2 \cup \{e^*\}$;
9. END WHILE
10. $S = \arg\max_{S \in \{S_1, S_2\}} \Omega(S)$;
11. 令 $x_{kj} = 1, \forall (k, j) \in S$;
12. 返回 $\mathbf{X}$.

最后,分析算法 1 的时间复杂度和最优性,得到定理 2.

定理 2. 算法 1 能够在 $O(N^2 M)$ 时间复杂度内得到近似率为 $1/\lceil 2 + c_{\max}/c_{\min} \rceil$ 的近似最优解,其中 N 和 M 分别表示汽车和感知任务个数; $c_{\max}$ 和 $c_{\min}$ 分别表示所有感知成本的最大和最小值.

证明. 算法 1 最多需要 $O(N)$ 次迭代. 由于每次迭代需要从 $O(NM)$ 个分配方案中选择最优分配方案,因此它的时间复杂度为 $O(N^2 M)$. 此外,根据文献[43]可知,对于具有单调子模目标函数和 q -独立系统约束的最大化问题,基于整体收益的贪心分配算法可达到 $1/(1+q)$ 的近似率. 由于算法 1 联合考虑整体和边际效益,因而性能不低于基于整体收益的贪心分配算法. 所以,根据引理 2,算法 1 的近似率为 $1/\lceil 2 + c_{\max}/c_{\min} \rceil$.

5 基于大规模数据集的实验评估

5.1 数据集和实验方法介绍

5.1.1 大规模车联网汽车数据集

为了更加全面准确地评估所提算法的性能,利

用如以下两个大规模数据集进行实验评估：

(1) 重庆市车联网汽车数据集^[17]. 该数据集包括大量车联网汽车一个月的轨迹数据(大约 12 493 辆出租车, 2017/03/01/~2017/03/31), 覆盖了重庆市约 25 472 m² 的城市面积. 其中每条记录主要包含汽车 ID、时间、GPS 位置、速度以及载客/空载状态. 该数据集的采样频率约 15 s, 总存储大小约为 92 GB.

(2) 美国纽约市车联网汽车数据集^[23]. 该数据集由纽约市黄色出租车一年(即 2017 年)的轨迹行程数据组成, 共包含超过 1.13 亿条行程数据. 每条数据包括行程 ID、载客/下客的时间/位置、行程距离以及载客收益等.

5.1.2 实验方法和参数设置

利用重庆市和纽约市车联网汽车数据集进行实验评估, 方法如下: (1) 对于重庆数据集, 选择重庆市江北区作为实验区域, 将它均匀划分成 125 个子区域(每个子区域大小约为 2 km²). 对于每一个子区域中每一个时段, 用该时段子区域内成功载客车辆数与总空闲车辆数之比作为载客概率. 对于成功载客的车辆, 基于行程轨迹, 通过计算轨迹中采样点之间的球面距离, 得到行程距离, 并基于重庆市出租车计价规则得到载客收益. 最后, 计算该时段子区域内所有成功载客车辆的平均载客收益, 作为单程收益期望; (2) 对于纽约数据集, 按照其街区拓扑划分成 265 个子区域. 其它实验方法和数据集处理方法类似于重庆数据集.

在实验评估中, 平台发布 M 个感知任务随机分布于城市各个区域. 随机选择 N 个空闲车联网汽车作为候选感知用户, 将他们的实际空闲时间区间作为可参与感知任务的时间区间. 平台根据所提算法将感知任务分配给候选车联网汽车用户. 车联网汽车司机接到任务后以概率 ρ 完成该任务. ρ 服从均匀随机分布 $[0, 1]$. 感知任务有效时间区间设为 8:00~9:00, 计算载客隐性激励的时间周期长度设为 10 min, 感知成本是基于实际车联网汽车数据集并根据式(13)计算得到.

5.1.3 性能评价指标和实验对比方法

本文使用 4 种性能评价指标, 包括: (1) 预测平均误差(MAE). 指深度学习预测值与真实值之间的平均绝对误差, 用于评估基于深度学习的载客收益预测的准确性; (2) 感知效益. 指车联网汽车感知平台所获得感知任务的总效益, 如式(5)所示; (3) 任务分配时间耗费. 指感知平台完成感知任务分配的时间耗费, 用于评估任务分配算法的时间复杂度;

(4) 用户收入增长率. 指车联网汽车司机参加感知后所获的总收入相对于原收入(即只有载客收益)的增长率. 该指标被用来评估所提算法是否满足用户的个体理性, 即司机参加感知后是否提高收入. 在经济学中, 假设用户是个体理性的, 只有当用户执行感知任务能给他们带来收入的增加, 用户才会自愿参加^[47].

本文采用 7 种对比方法分别对所提方法的两个主要模块进行评估.

(1) 感知成本学习. 为了评估基于双注意力机制循环神经网络的感知成本学习预测方法, 利用 3 种深度学习方法进行对比: ① LSTM(Long Short-Term Memory Network)^[48]. 即长短期记忆网络深度学习算法, 是递归神经网络的变体, 可以解决传统神经网络训练过程中梯度消失和下降问题; ② Seq2Seq(Sequence to Sequence Model)^[49]. 即基于递归神经网络的序列预测模型, 利用编码器-解码器结构, 主要包括编码器、上下文向量和解码器三个主要模块; ③ TAtt(Temporal Attention)^[50]. 基于时间注意力机制的循环神经网络.

(2) 任务最优分配. 为了评估所提任务最优分配算法的性能, 采用 4 种对比方法: ① OPT(Optimal Allocation). 通过穷举搜索获得最优任务分配方案. 由于任务分配问题是 NP-hard, 因此, OPT 方法具有指数级时间复杂度; ② RD(Random Allocation). 将感知任务随机地分配给车联网汽车司机; ③ OR(Overall Revenue)^[51]. 基于整体收益的贪心分配算法, 即贪心地选择整体收益最大的分配方案; ④ ILS(iLOCuS)^[7]. 利用文献[7]的感知成本计算方法和本文任务分配策略.

5.2 实验结果

下面分别对所提方法的两个主要模块进行评估, 包括感知成本学习和任务最优分配.

5.2.1 感知成本学习

本实验随机选取美国纽约市 141 号街区作为目标预测区域, 以评估感知成本学习中载客收益预测方法精度, 包括载客概率预测和单程收益预测. 由于两者的结果相似, 本文主要展示单程收益预测的实验结果.

首先, 分别基于重庆与纽约数据集对比本文方法与 3 种深度学习方法、4 种非深度学习方法的载客收益预测精度. 非深度学习对比方法包括经典机器学习算法(如 SVR^[52]、Adaboost^[53] 和 KNN^[54]) 和经典统计分析预测算法(如 MLR^[44]). 如表 2 所示, 实验结果表明本文方法比 SVR、Adaboost、

KNN 和 MLR 在纽约数据集中分别平均提升预测精度 75.9%、56.8%、50.0% 和 47.0%，在重庆数据集中分别平均提升 69.1%、42.9%、35.5% 和 38.0%。

表 2 区域载客总收益预测精度

算法类别	算法名称	MAE	
		纽约数据集	重庆数据集
深度学习算法	本文方法	116.37	121.18
	TAtt	192.42	172.54
	Seq2Seq	216.16	188.05
	LSTM	184.59	188.81
非深度学习算法	SVR	483.88	392.56
	Adaboost	269.33	212.13
	KNN	232.87	187.84
	MLR	219.55	195.38

然后,评估不同输入和输出长度对深度学习算法性能的影响.其中输入和输出长度分别表示深度学习算法输入数据的长度和输出数据的长度.例如,本文使用过去 h 个时间段历史数据来预测未来 T 个时间段数据,则输入和输出长度分别为 h 和 T .一方面,变化输入长度从 12 到 120,输出长度设为 3,如图 6 所示.实验结果表明,随着输入长度的增加,本文方法预测精度逐渐提高.当输入长度为 96 时,预测精度最高,即 $MAE = 8.33$.同时,相比于 LSTM、Seq2Seq 和 TAtt,本文方法分别平均提高预测精度 31.2%、27.1% 和 25.1%.另一方面,变化输出长度从 2 到 10,输入长度设为 96.如图 7 所示,实验结果表明,随着输出长度的增加,本文方法与 TAtt 性能都比较稳定.同时,本文方法的预测精度比 Seq2Seq 和 TAtt 分别平均提高 55.5% 和 13.5%.图 8 展示本文方法和三种对比方法在不同输入长度下载客概率的预测精度.实验结果表明本文方法的预测精度比 LSTM、Seq2Seq 和 TAtt 分别平均提高预测 18.6%、20.4% 和 14.7%。

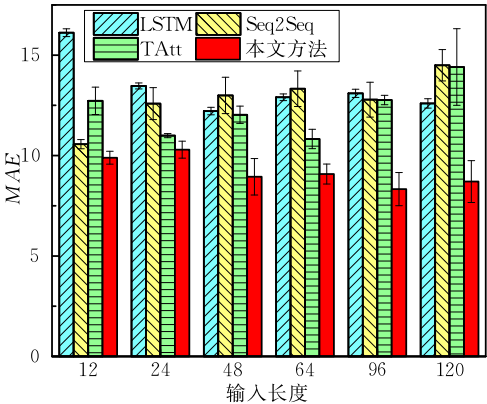


图 6 不同输入长度的载客收益预测精度

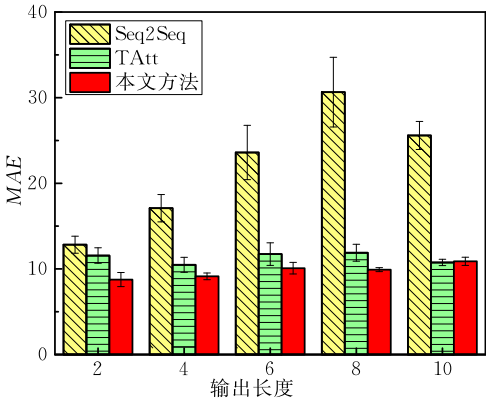


图 7 不同输出长度的载客收益预测精度

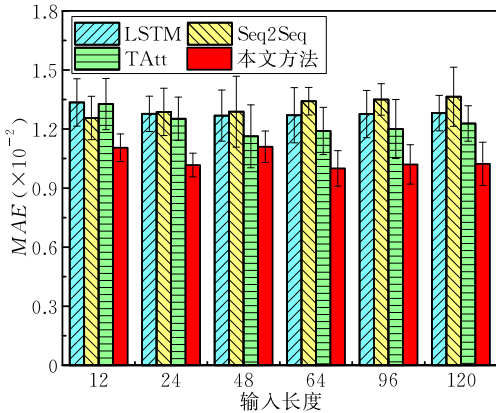


图 8 不同输入长度的载客概率预测精度

其次,由于本文方法采用时空双注意力机制,因此评估邻近区域输入数量对算法的影响.如图 9 所示,随着输入区域数量的增加,本文方法的预测精度逐步提高.当输入的邻近区域数量为 10 时,预测误差可以达到 $MAE = 10.02$.同时,图 9 显示不同邻近区域的加入对预测结果的影响也各不相同.主要原因如下:对于目标区域的预测,加入部分相关性较强的邻近区域可提高它的预测性能;反之,加入相关性弱的区域则会给模型学习带来更多的噪声,从而降低预测精度.同时,图 10 展示不同迭代次数的本文预测方法损失值,即预测值与真实值的差值绝对值.输入和输出长度分别设定为 96 和 3.共做 15 次实验,每次实验训练 100 次,每次分别计算训练集和测试集的损失值.实验结果显示,本文方法在 20 次、60 次和 100 次训练后,在训练集中的平均损失分别为 0.0167、0.0040 和 0.0028,在测试集中的平均损失分别为 0.0167、0.0064 和 0.0061.因此,实验结果表明所提算法只需 60 次训练就能获得非常稳定的预测性能,且测试集上的性能接近训练集的性能。

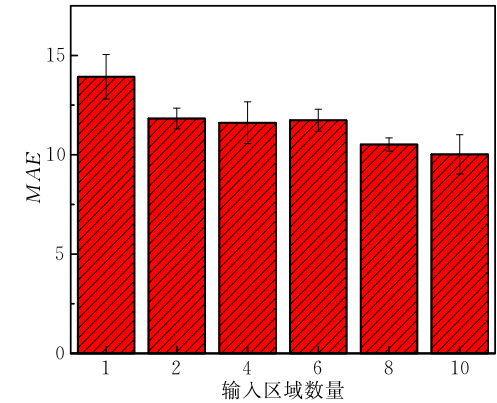


图 9 不同输入区域数量的载客收益预测精度

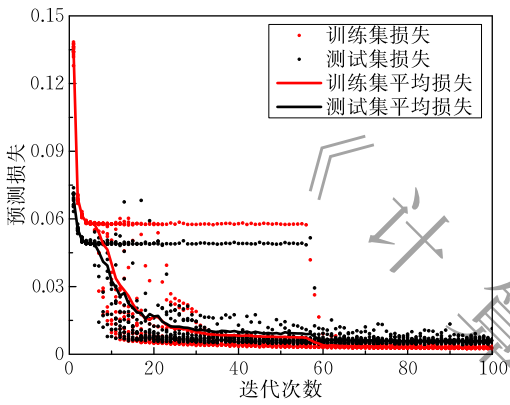


图 10 不同迭代次数的预测损失值

最后,评估本文方法所学习的车联网汽车感知成本是否满足用户的个体理性,即是否能够保证所有用户的收益增加.如图 11 所示,除对比 ILS 方法,将本文方法的变种 wo/p(即无隐性激励的成本计算方法)作为对比方法.通过统计 14 916 个车联网汽车用户的收入增长率,结果显示本文方法的用户收入增长率都大于 0,满足用户的个体理性.然而,ILS 和 wo/p 都分别存在 7%和 6.9%的用户收入增长率小于 0.因此,这两种方法不满足用户个体理性,因而不能很好地激励用户参与车联网汽车感知任务.

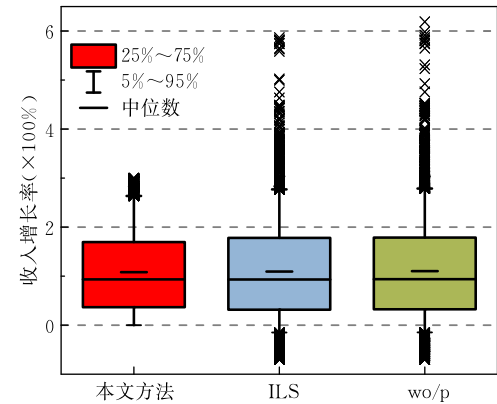


图 11 算法的个体理性评估

5.2.2 任务最优分配

本小节分别评估候选车联网汽车数量、感知任务数量、感知成本预算、车联网汽车可靠性对所提任务分配算法的性能影响,以及与最优解之间的性能比较.

首先,变化候选车联网汽车数量从 10 到 140,如图 12 所示.实验结果表明本文方法的性能都优于 ILS、OR 和 RD 的性能.同时,相比于 3 种对比方法,本文方法的感知效益分别平均提高 44.9%、291.0%和 496.6%.此外,图 12 显示当候选车联网汽车数量较少时,本文方法和 ILS 性能接近.其主要原因是 ILS 方法计算感知成本虽然也考虑了目的地潜在载客需求,但它的计算方法不准确,如不满足个体理性.因此,当候选车联网汽车数量较少时,因为感知成本总预算充足,感知成本的学习精度对算法性能影响很小,因此,本文所提算法与 ILS 方法很接近.同时,该解释也与图 14 的实验结果相一致,即随着感知成本总预算增加和充足,本文算法与 ILS 算法性能逐渐接近.另外,变化感知任务数量从 10 到 140,如图 13 所示.在不同感知任务数量下,本文方法的感知效益比 ILS、OR 和 RD 分别平均提高 50.5%、297.6%和 649.4%.

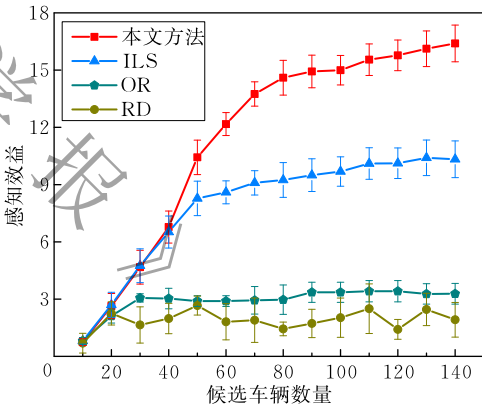


图 12 不同候选车辆数量的任务分配性能比较

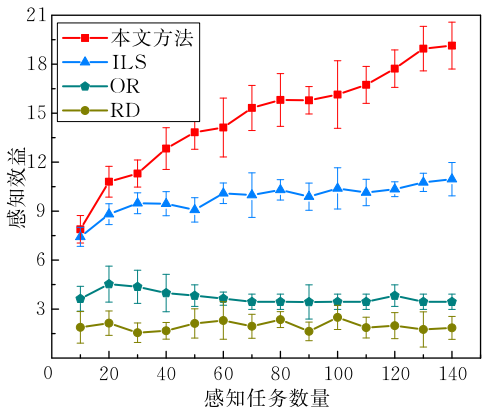


图 13 不同感知任务数量的任务分配性能比较

然后,变化感知成本预算从 20 元到 280 元. 如图 14 所示,在不同预算下,本文方法的感知效益比 ILS、OR 和 RD 平均提高 37.7%、266.9% 和 480.8%. 同时,变化车联网汽车完成任务概率从 10%到 100%. 图 15 表明感知效益随着车联网汽车可靠性的增加而增加. 同时,在不同车联网汽车可靠性下,本文方法性能分别比 ILS、OR、RD 和 wo/p 平均提高 56.8%、356.7%、307.9%和 119.2%.

最后,比较本文方法与 OPT 的任务分配性能和 时间耗费. 由于 OPT 具有指数级时间复杂度,因此,本实验基于小规模场景进行评估(即候选汽车和感知任务数量分别为 10 和 6). 如图 16 所示,本文

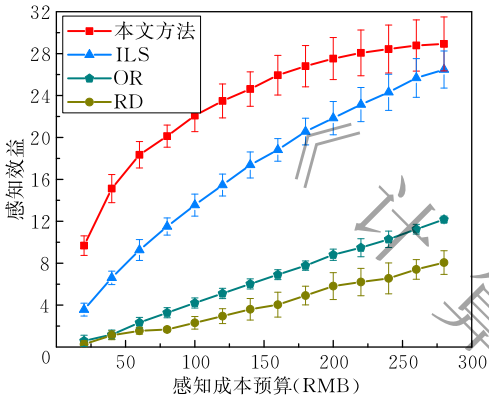


图 14 不同感知成本预算的任务分配性能比较

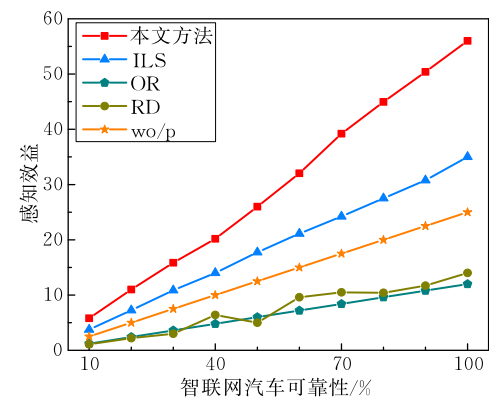


图 15 不同车联网汽车可靠性的任务分配性能比较

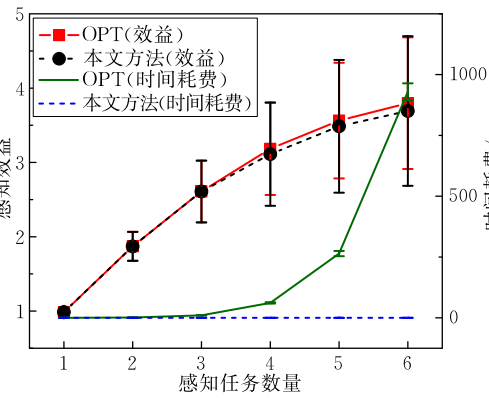


图 16 与 OPT 方法的任务分配性能和时间耗费比较

方法以极低的时间耗费获得接近于 OPT 的感知效益. 例如,当感知任务数量为 6 时,本文方法和 OPT 的平均感知效益分别为 3.69 和 3.80,时间耗费分别为 0.023 s 和 929.540 s. 因此,本文方法仅耗费 0.0025%的时间就能获得 97.1%的最优感知效益. 同时,该实验结果与定理 2 的理论分析结果相一致.

5.3 车联网汽车感知原型系统实现

为了验证所提车联网汽车感知的实际应用性、可行性和作用性,下面构建面向城市道路违规停车监测的车联网汽车感知原型系统.

系统意义. 随着当前城市化的飞速发展,城市机动车使用数量日益庞大^[5],停车位的严重紧缺导致城市违规停车现象越来越普遍,从而对城市管理造成严重的危害,如引发交通堵塞、阻碍消防救援、造成交通事故等^[55]. 因此,构建大规模城市道路违规停车监测系统对于城市管理非常重要. 传统方法主要基于部署摄像头等感知系统和交通警察人工巡检两种方式,但在大规模城市应用中具有成本高、覆盖率低、效率低等缺点^[56]. 因此,为了解决上述问题,本文提出基于车联网汽车的城市道路违规停车监测系统,利用大量车联网汽车(如出租车和滴滴汽车)中现有的移动感知资源,实现大规模、细粒度、低成本、高效率的城市违规停车事件感知和监测.

系统实现和结果. 该原型系统主要包括以下四步: (1) 平台将需要监测的道路作为感知任务发布出去,并按照本文所提方法分配给车联网汽车司机去感知和执行; (2) 司机开车前往指定路段,利用行车记录仪和车载智能手机采集该路段的视频数据与其他感知数据,如 GPS、加速度和陀螺仪等; (3) 司机通过手机 4G/5G 或者 WiFi 网络将感知数据上传至平台并获得一定的报酬奖励; (4) 平台利用 Faster RCNN^[57]深度学习算法,基于视频数据识别汽车是否停止,并联合汽车 GPS 位置估计该停止汽车的精确位置,判断该汽车是否处于禁停区域,从而准确地识别道路违规停车事件.

为了验证该系统的可行性,以重庆市沙坪坝区为例,招募出租车和滴滴网约车实现道路违规停车监测. 如图 17 所示,在 15km 的城市道路中,共监测到违停汽车约 74 辆,分布在 16 个路段上. 同时,实验结果显示绝大部分违停事件具有明显的聚集性,即集中在部分无人监管或无监控摄像头的路段,如路段 A、B 和 C 上分别有违停汽车 4 辆、6 辆和 11 辆.

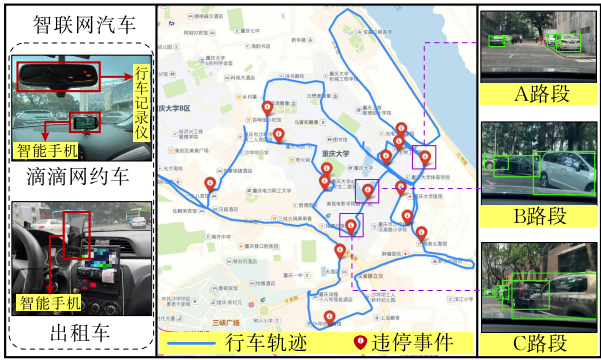


图 17 面向城市道路违规停车监测的车联网汽车感知原型系统

6 总结与展望

6.1 工作总结

为了充分地利用现有众多车联网汽车上感知、计算和通信资源进行大规模城市感知,本文提出了基于深度强化学习的车联网汽车感知任务分配架构和算法.具体地,通过城市区域邻接图建模,利用基于双注意力机制的循环神经网络,对高时空动态的车联网汽车载客收益进行预测,从而学习车联网汽车执行感知任务的成本.其次,基于问题等价转化和理论分析,利用子模优化理论,提出了基于整体和边际效益的近似最优任务分配算法,在多项式时间复杂度内获得具有理论下界的近似最优解.然后,基于两个实际大规模车联网汽车数据集进行实验评估.结果表明所提算法的预测学习精度和任务分配算法性能都高于对比方法.最后,构建了面向城市道路违规停车监测的车联网汽车感知原型系统,验证了所提方法的实际应用性和可行性.

6.2 研究展望

本文未来工作拟从以下两个方面进行改进和完善:

(1)任务分配模型的通用性.为了提高所提任务分配模型的通用性,拟从以下两种场景进行扩展.①针对车联网汽车用户在非空闲时间段(即载客途中)执行感知任务的场景,拟采用文献[7]的方法进行扩展;②针对多任务分配场景(即一个汽车用户在一个空闲周期内可以被分配多个感知任务)^[58],可以研究以下两部分:一是面向多任务分配的感知成本学习模型.由于一个用户可以执行多个任务,因此,多个任务的感知成本不是各个感知任务成本的简单线性相加,而是一个复杂的组合模型,需要考虑该多任务分配导致用户行驶轨迹的长度改变等.二

是面向多任务分配的最优路径规划.由于用户执行多任务,因此,需要进行路径规划,这是一个经典的最短旅行商 NP-hard 问题.同时,路径规划影响任务感知成本,从而影响任务最优分配.因此,需要联合考虑路径规划和任务分配,解决两个紧耦合的 NP-hard 问题.

(2)车联网汽车感知应用实现.本文只构建和实现面向城市道路违规停车监测的车联网汽车感知原型系统,初步验证车联网汽车感知的应用性、可行性和作用性.下一步将进一步完善和构建更大规模的违规停车感知系统.同时,研究和实现基于车联网汽车感知的其它实际应用,如①道路路面条件感知^[11],即通过车联网汽车司机的手机内置加速度传感器感知道路坑洼和破损状况,以便于安全驾驶预警和城市道路及时修复;②城市道路地图实时构建^[13],即通过车联网汽车 GPS 和行车记录仪等感知设备采集道路视频等感知数据,然后提取道路语义信息,从而构建面向自动驾驶的高清地图.

致 谢 感谢重庆大学计算机学院研究生周彦霖、程文辉、张乾元、陈瑞在车联网汽车感知原型系统实现中的贡献,同时感谢编辑部和审稿专家给予本文的宝贵意见!

参 考 文 献

[1] Zhao Dong, Ma Dong-Hua. Development and challenges of crowd-sensing networks. Information and Communications Technology, 2014, 8(5): 66-70(in Chinese)
(赵东, 马东华. 群智感知网络的发展及挑战. 信息通信技术, 2014, 8(5): 66-70)

[2] Liu Yun-Hao. Crowd-sensing computing. Communication of the CCF, 2012, 8(10): 38-41(in Chinese)
(刘云浩. 群智感知计算. 中国计算机学会通讯, 2012, 8(10): 38-41)

[3] Li Jing-Lin, Yuan Quan, Yang Fang-Chun. Vehicles network based crowd-sensing and service. ZTE Communications, 2015, 21(6): 6-9(in Chinese)
(李静林, 袁泉, 杨放春. 车联网群智感知与服务. 中兴通讯技术, 2015, 21(6): 6-9)

[4] Xiang C, et al. Calibrate without calibrating: An iterative approach in participatory sensing network. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(2): 351-361

[5] The number of motor vehicles in the country reaches 365 million. http://www.gov.cn/shuju/2020-10/20/content_5552831.htm, 2020(in Chinese)

- (全国机动车保有量达 3.65 亿辆. http://www.gov.cn/shuju/2020-10/20/content_5552831.htm, 2020)
- [6] Wahlström J, Skog I, Händel P. Smartphone-based vehicle telematics: A ten-year anniversary. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 18(10): 2802-2825
 - [7] Xu S, et al. iLOCuS: Incentivizing vehicle mobility to optimize sensing distribution in crowd sensing. *IEEE Transactions on Mobile Computing*, 2020, 19(8): 1831-1847
 - [8] Xu T, et al. Taxi driving behavior analysis in latent vehicle-to-vehicle networks: A social influence perspective//*Proceedings of the 22nd ACM SIGKDD*. San Francisco, USA, 2016: 1285-1294
 - [9] Zhang L, et al. A taxi order dispatch model based on combinatorial optimization//*Proceedings of the 23rd ACM SIGKDD*. Halifax, Canada, 2017: 2151-2159
 - [10] Xu J, et al. Online incentive mechanism for mobile crowdsourcing based on two-tiered social crowdsourcing architecture//*Proceedings of the 15th SECON*. Hong Kong, China, 2018: 1-9
 - [11] Eriksson J, et al. The pothole patrol: Using a mobile sensor network for road surface monitoring//*Proceedings of the 6th MobiSys*. Breckenridge, USA, 2008: 29-39
 - [12] Liu L, et al. BigRoad: Scaling road data acquisition for dependable self-driving//*Proceedings of the 15th MobiSys*. Niagara Falls, USA, 2017: 371-384
 - [13] Chen D, Shin K G. TurnsMap: Enhancing driving safety at intersections with mobile crowdsensing and deep learning//*Proceedings of the ACM UbiComp*. London, UK, 2019: 1-22
 - [14] Payver. <https://www.lazymoneyguy.com/payver/>, 2019
 - [15] Chffr. <https://comma.ai/>, 2019
 - [16] Tang X, et al. A deep value-network based approach for multi-driver order dispatching//*Proceedings of the 25th ACM SIGKDD*. Anchorage, USA, 2019: 1780-1790
 - [17] Chen C, et al. TrajCompressor: An online map-matching-based trajectory compression framework leveraging vehicle heading direction and change. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 21(5): 2012-2028
 - [18] What the Sharing Economy Really Delivers: Entitlement. <https://www.nytimes.com/2018/01/26/nyregion/what-the-sharing-economy-really-delivers-entitlement.html>, 2018
 - [19] Tong Y, et al. The simpler the better: A unified approach to predicting original taxi demands based on large-scale online platforms//*Proceedings of the 23rd ACM SIGKDD*. Halifax, Canada, 2017: 1653-1662
 - [20] He S, Shin K G. Spatio-temporal adaptive pricing for balancing mobility-on-demand networks. *ACM Transactions on Intelligent Systems and Technology*, 2019, 10(4): 1-28
 - [21] He S, Shin K G. Spatio-temporal capsule-based reinforcement learning for mobility-on-demand network coordination//*Proceedings of the World Wide Web Conference*. San Francisco, USA, 2019: 2806-2813
 - [22] 25 Amazing Didi Facts and Statistics. <https://expandedramblings.com/index.php/didi-chuxing-facts-statistics/>, 2019
 - [23] TLC Trip Record Data. <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>, 2019
 - [24] He Z, Cao J, Liu X. High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility//*Proceedings of the IEEE INFOCOM*. Hong Kong, China, 2015: 2542-2550
 - [25] Gao G, et al. Truthful incentive mechanism for nondeterministic crowdsensing with vehicles. *IEEE Transactions on Mobile Computing*, 2018, 17(12): 2982-2997
 - [26] Zhang X, et al. Truthful incentive mechanisms for crowdsourcing//*Proceedings of the IEEE INFOCOM*. Hong Kong, China, 2015: 2830-2838
 - [27] Chen X, et al. PAS: Prediction-based actuation system for city-scale ridesharing vehicular mobile crowdsensing. *IEEE Internet of Things Journal*, 2020, 7(5): 3719-3734
 - [28] Xiang C, et al. Incentivizing platform—User interactions for crowdsensing. *IEEE Internet of Things Journal*, 2020, 8(10): 8314-8327
 - [29] Fan X, et al. BuildSenSys: Reusing building sensing data for traffic prediction with cross-domain learning. *IEEE Transactions on Mobile Computing*, 2020, 20(6): 2154-2171
 - [30] Xiang C, et al. Edge computing-empowered large-scale traffic data recovery leveraging low-rank theory. *IEEE Transactions on Network Science and Engineering*, 2020, 7(4): 2205-2218
 - [31] Xiang C, et al. CARM: Crowd-sensing accurate outdoor RSS maps with error-prone smartphone measurements. *IEEE Transactions on Mobile Computing*, 2016, 15(11): 2669-2681
 - [32] Liu Y, Kong L, Chen G. Data-oriented mobile crowdsensing: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 2019, 21(3): 2849-2885
 - [33] Fan G, et al. Joint scheduling and incentive mechanism for spatio-temporal vehicular crowd sensing. *IEEE Transactions on Mobile Computing*, 2019, 20(4): 1449-1464
 - [34] Tang X, et al. A deep value-network based approach for multi-driver order dispatching//*Proceedings of the 25th ACM SIGKDD*. USA, 2019: 1780-1790
 - [35] Xu Z, et al. Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach//*Proceedings of the 24th ACM SIGKDD*. London, UK, 2018: 905-913
 - [36] Sühr T, et al. Two-sided fairness for repeated matchings in two-sided markets: A case study of a ride-hailing platform//*Proceedings of the 25th ACM SIGKDD*. Anchorage, USA, 2019: 3082-3092
 - [37] Zhu Fei, et al. A dual deep network based secure deep reinforcement learning method. *Chinese Journal of Computers*, 2019, 42(8): 1812-1826(in Chinese)

(朱斐等. 基于双深度网络的安全深度强化学习方法. 计算机学报, 2019, 42(8): 1812-1826)

[38] Liao Xiao-Min, et al. Cellular network resource allocation algorithm based on deep reinforcement learning. Journal on Communications, 2019, 40(2): 11-18(in Chinese)
(廖晓闽等. 基于深度强化学习的蜂窝网资源分配算法. 通信学报, 2019, 40(2): 11-18)

[39] Liu Quan, et al. A survey on deep reinforcement learning. Chinese Journal of Computers, 2018, 41(1): 1-27(in Chinese)
(刘全等. 深度强化学习综述. 计算机学报, 2018, 41(1): 1-27)

[40] Liao Z. Real-time taxi dispatching using global positioning systems. Communications of the ACM, 2003, 46(5): 81-83

[41] Wang J, et al. Crowd-powered sensing and actuation in smart cities: Current issues and future directions. IEEE Wireless Communications, 2019, 26(2): 86-92

[42] Cao X, et al. Trajectory penetration characterization for efficient vehicle selection in HD map crowdsourcing. IEEE Internet of Things Journal, 2020, 8(6): 4526-4539

[43] Balinski M L, Hoffman A J. Polyhedral Combinatorics; Dedicated to the Memory of D.R. Fulkerson. New York, USA: Elsevier, 1978

[44] Andrews D F. A robust method for multiple linear regression. Technometrics, 1974, 16(4): 523-531

[45] Qin Y, et al. A dual-stage attention-based recurrent neural network for time series prediction//Proceedings of the IJCAI. Melbourne, Australia, 2017: 2627-2633

[46] Abdelhamid S, Hassanein H S, Takahara G. Reputation-aware, trajectory-based recruitment of smart vehicles for public sensing. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(5): 1387-1400

[47] Liu Yuan-Ni, et al. Mobile crowd-sensing network incentive mechanism based on auction model. Journal on Communications, 2019, 40(7): 208-222(in Chinese)
(刘媛妮等. 基于拍卖模型的移动群智感知网络激励机制. 通信学报, 2019, 40(7): 208-222)

[48] Xie Zhao, Wu Yi, Wu Ke-Wei, Zhang Shun-Ran. Activity recognition based on spatial-temporal attention LSTM. Chinese Journal of Computers, 2021, 44(2): 261-274 (in Chinese)
(谢昭, 周义, 吴克伟, 张顺然. 基于时空关注度 LSTM 的行为识别. 计算机学报, 2021, 44(2): 261-274)

[49] Liu B, et al. A sequence-to-sequence air quality predictor based on the n-step recurrent prediction. IEEE Access, 2019, 7: 43331-43345

[50] Fan C, et al. Multi-horizon time series forecasting with temporal attention learning//Proceedings of the 25th ACM SIGKDD. Anchorage, USA, 2019: 2527-2535

[51] Farhadi V, et al. Service placement and request scheduling for data-intensive applications in edge clouds//Proceedings of the IEEE INFOCOM. Paris, France, 2019: 1279-1287

[52] Hearst M A, et al. Support vector machines. IEEE Intelligent Systems and their Applications, 1998, 13(4): 18-28

[53] Solomatine D P, Shrestha D L. AdaBoost.RT: A boosting algorithm for regression problems//Proceedings of the IEEE IJCNN. Budapest, Hungary, 2004, 2: 1163-1168

[54] Zhang S, et al. Learning k for kNN classification. ACM Transactions on Intelligent Systems and Technology, 2017, 8(3): 1-19

[55] How can the phenomenon of the fire channel being blocked by private cars be managed. https://www.sohu.com/a/452087869_582914, 2020(in Chinese)
(消防通道被私家车堵现象如何才能得以治理. https://www.sohu.com/a/452087869_582914, 2020)

[56] He Tian-Fu, et al. Detecting vehicle illegal parking events using sharing bikes' trajectories//Proceedings of the 24th ACM SIGKDD. London, UK, 2018: 340-349

[57] Cheung H, et al. Accurate distance estimation using camera orientation compensation technique for vehicle driver assistance system//Proceedings of the IEEE ICCE. Las Vegas, USA, 2012: 227-228

[58] Xu J, et al. Incentive mechanism for multiple cooperative tasks with compatible users in mobile crowd sensing via online communities. IEEE Transactions on Mobile Computing, 2020, 19(7): 1618-1633



LI Yao-Yu, M.S. candidate. His current research interests include vehicular crowd-sensing networks and deep

XIANG Chao-Can, Ph. D. , associate professor, Ph. D. supervisor. His current research interests include crowd-sensing networks, AI, and pervasive computing.

reinforcement learning.

FENG Liang, Ph. D. , professor. His research interests include reinforcement learning and artificial Intelligence.

CHEN Chao, Ph. D. , professor. His research interests include urban computing and AI.

GUO Song-Tao, Ph. D. , professor. His research interests include edge computing and task optimal allocation.

YANG Pan-Long, Ph. D. , professor. His research interests include crowd-sensing networks and sensor-less sensing.

Background

The CNN News in 2017 reported that “the car’s sensing data will be more valuable than the car itself”. It is because the ubiquitous vehicles are equipped with plenty of on-board sensors, as well as powerful computation and communication capabilities, that numerous vehicular sensing data are obtainable for many significant applications, such as self-driving training, Pothole Patrol detection, and air pollution monitoring. Moreover, with the increasing number of Mobility-on-Demand (MOD) vehicles (also called for-hire vehicles, such as Uber, DiDi, and connected taxis), these MOD vehicles can not only convey passengers but also conduct vehicular sensing in cities. For example, a recent report by CNN News shows that “Uber and Lyft drivers are using their dashboard cameras in urban sensing for self-driving map construction”.

Given above, leveraging such massive MOD vehicles with equipped sensors to execute the sensing tasks during their unoccupied time, is referred to as the MOD Vehicular CrowdSensing, which gains much attention owing to large-scale coverage and powerful sensing ability of MOD vehicles. Its critical research problem is sensing task allocation, i.e., how to allocate vehicles to execute sensing tasks so as to maximize sensing utility within a limited budget. Most of the existing works neglect the highly spatial-temporal dynamics of vehicular cost, which is dependent on not only the driving cost (such as driving distance and time), but also the pick-up earnings of MOD vehicles.

In this work, we propose a vehicular sensing task allocation method empowered by deep reinforcement learning. It learns the vehicular cost via a deep learning scheme, which is fed back to optimally allocate sensing tasks in polynomial time. In specific, we exploit the multi-attention-based Recurrent Neural Network (RNN) to extract the spatial-temporal correlations of pick-up revenues, which are used to learn the sensing cost based on the driving cost model. Furthermore, through the equivalent problem transformation, we prove that the problem of task allocation has a submodular objective function and a q -dependent constraint. Hence, based on the sub-modularity theory, we propose a sub-optimal task allocation algorithm, jointly considering the whole utilization and the marginal one. The proofs demonstrate that it achieves a constant-factor approximation ratio in polynomial time.

The authors have made considerable researches on vehicular crowdsensing and reinforcement learning, publishing more than 40 high-quality papers, such as ACM MobiCom, IEEE INFOCOM, ACM UbiComp, IEEE TMC, IEEE ITS, IEEE TPDS, IEEE TVT, and IEEE EC. This work is supported by the National Natural Science Foundations of China (No. 62172063 and No. 61872447). The opinions, findings, and conclusions presented in this paper are those of the authors and do not necessarily reflect the view of the funding agencies or the government.