

引用格式:师泽宇,陈阳舟,安树科. 基于交通预测信息的电动汽车充电路径规划[J]. 北京工业大学学报, 2024, 50(8): 974-984.

SHI Z Y, CHEN Y Z, AN S K. Electric vehicle charging routing planning based on the traffic prediction[J]. Journal of Beijing University of Technology, 2024, 50(8): 974-984. (in Chinese)

# 基于交通预测信息的电动汽车充电路径规划

师泽宇<sup>1</sup>, 陈阳舟<sup>2</sup>, 安树科<sup>3</sup>

(1. 北京工业大学交通工程北京市重点实验室, 北京 100124;

2. 北京工业大学信息学部人工智能与自动化学院, 北京 100124;

3. 云南警官学院治安管理学院, 昆明 650223)

**摘要:** 针对电动汽车充电路径规划问题,以行程时间为优化目标,以荷电状态作为硬性约束,基于预测交通状态标定的动态异构交通网及车辆动力学模型,构建最优控制问题。为了在线求解,基于模型预测控制和强化学习方法,构建混合学习优化算法(hybrid learning optimization algorithm, HLOA),依托竞争深度Q网络(Dueling deep Q-networks, Dueling DQN),设计离线在线模块,来实现参数的混合多步训练,并结合预测交通状态,网络拓扑及车辆动力学模型,输出在线行驶策略。采用某市局部路网构建交通仿真实验,分析预测交通信息的价值、HLOA算法可行性及结构必要性。实验结果表明,预测信息存在价值,HLOA在线结果可拟合最优策略且较其他算法精度提高33.65%。该研究可协助电动汽车用户导航,可为自动驾驶电动汽车提供技术支撑。

**关键词:** 智能交通系统; 车辆路径规划; 电动汽车充电路径规划; 强化学习; 在线寻优; 混合学习优化算法

中图分类号: U 492.2

文献标志码: A

文章编号: 0254-0037(2024)08-0974-11

doi: 10.11936/bjutxb2022110037

## Electric Vehicle Charging Routing Planning Based on Traffic Prediction

SHI Zeyu<sup>1</sup>, CHEN Yangzhou<sup>2</sup>, AN Shuke<sup>3</sup>

(1. Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China;

2. College of Artificial Intelligence and Automation, Beijing University of Technology, Beijing 100124, China;

3. School of Public Security Management, Yunnan Police College, Kunming 650223, China)

**Abstract:** Aiming at the electric vehicle charging routing problem, an optimal control model was built based on the dynamic heterogeneous traffic network, vehicle dynamics model, and predicted traffic state. The travel time of the vehicle is the optimized objective and the battery state is the hard constraint in the model. To solve the optimal control model online, a hybrid learning optimization algorithm (HLOA) was constructed based on the spatial-discrete model predictive control theory and reinforcement learning method. Relying on Dueling deep Q-networks (Dueling DQN), the algorithm has designed two solving modules, i. e., offline and online, to realize the mixed multi-step training of its parameters. After the

收稿日期: 2022-11-25; 修回日期: 2023-03-20

基金项目: 国家自然科学基金资助项目(61573030); 云南省教育厅科学研究基金项目(2022J0587)

作者简介: 师泽宇(1992—), 男, 博士研究生, 主要从事交通系统、交通预测、车辆路径规划方面的研究, E-mail: zeyushi@emails.bjut.edu.cn

training, the algorithm combines the predicted traffic state, dynamic heterogeneous traffic network, and vehicle dynamics model to output online driving strategies in real-time. Based on the above methods, the local road network in one city is the experimental scenario. Combined with the traffic simulation, the competitive ratio was used as the evaluation index to analyze the value of the predicted traffic information and the feasibility of HLOA. The experimental results show the value of the prediction information and the feasibility of HLOA. The competitive ratio of HLOA is 33.65% higher than the other algorithms. The algorithm can help electric vehicle users to plan their trips. It can also provide technical support for autonomous vehicles.

**Key words:** intelligent transportation system; vehicle routing planning; electric vehicle charging routing planning; reinforcement learning; online optimization; hybrid learning optimization algorithm

纯电动汽车因无排放、低成本、高能源利用率等优点逐渐成为公路交通工具界的新宠,截止到 2020 年底,我国新能源汽车保有量为 492 万辆<sup>[1]</sup>。作为新兴的交通工具,商业化电动汽车的电池仍存在更换成本高、续航能力有限等问题,同时,电动汽车配套充电设施总量少、分布不均的情况也困扰着急需充电的用户。电动汽车在城市街道中行驶,由于信号灯和交通流状态的实时变化,需要频繁采取制动、换道和加速等动作,这进一步增加了能耗负担,提高了其充电频率,因此,快捷、节能,且可在线计算的充电路径规划方法成为了城市电动汽车用户的重要需求。

电动汽车充电路径规划问题 (electric vehicle charging routing planning, EVCRP) 本质上隶属车辆路径规划问题 (vehicle routing problem, VRP)<sup>[2-3]</sup>, 而许多研究基于此思路将电动汽车充电路径规划问题简化为车辆路径问题来描述,例如:Sun 等<sup>[4]</sup>以动态行程时间和费用作为网络权值,构建 VRP 来描述 EVCRP;Liang 等<sup>[5]</sup>将 EVCRP 看作以能耗最优为目标,随机行程时间为网络权值的 VRP;Perger 等<sup>[6]</sup>在构建 Liang 等<sup>[5]</sup>的 VRP 基础上,还考虑了电池寿命和地形变化作为状态变量,来描述 EVCRP。将 EVCRP 简化为 VRP 忽略车辆对充电站的选择,其默认行驶路径存在唯一终点且仅为已选定的充电站,该类方法的充电站选择基于研究者主观判断,会出现舍弃能耗或时间花费较少的充电站,而选择其他充电站的现象。Yan 等<sup>[7]</sup>用随机用户平衡模型来描述 EVCRP 也存在一定局限性,该模型更适用于车辆调度问题,而不适用于单车充电导航,而其本质还是依赖静态的行驶路径,这并不适合动态交通状态下的在线求解。而将问题描述为最优控制问题,可以摆脱对行驶路径的静态依赖,问题解的形式改

为行驶策略,策略可以动态生成路径,面对在线求解中时变交通流更有优势,采用最优控制描述的文献包括但不限于文献[8-10]及本文。

基于上述对 EVCRP 的描述,对应的求解算法也被相继提出,这些算法可分为最优化解析算法、启发式算法、近似动态规划算法、强化学习算法及混合算法<sup>[11-12]</sup>。其中,最优化解析算法是最早被采用的,常见的算法包括:动态规划算法<sup>[13]</sup>和混合整数规划算法<sup>[14-15]</sup>,但最优化解析算法存在计算复杂度高的特点,并不适用在线计算的需求。近似动态规划算法主要通过构建近似函数获得近似于动态规划算法的解,其常见的函数设计方法为策略函数近似<sup>[16]</sup>和值函数近似<sup>[17-19]</sup>;启发式算法的思路也通过设计规则来近似最优解,且其计算效率高,常见的算法包括:粒子群算法<sup>[20]</sup>、多步采样算法<sup>[21]</sup>、蚁群算法<sup>[22]</sup>等;以上 2 种算法适用于交通特征维度低或交通流变化平缓的路网,当交通流变化剧烈时,其在线解无法保证精度的稳定性。强化学习方法可以从实时数据中获取信息,并通过在线学习提高精度,具有较强的稳定性,成为了热门的在线求解算法,例如文献[23-25]和本文。此外 Pop 等<sup>[26]</sup>和 Gutierrez 等<sup>[27]</sup>尝试多类算法混合构建新算法,但其可解释性不强,存在精度稳定性有待进一步验证的情况,所以并未被学界广泛接受。

本文基于现有的研究,结合了预测交通信息、动态异构交通拓扑结构和车辆动力学模型,采用最优控制理论,首先构建了用户最优为目标的 EVCRP 问题描述。为了使问题便于在线求解,随后采用空间离散模型预测控制理论,将最优控制问题转化为模型预测控制问题。基于以上问题描述,权衡已有求解方法的利弊,本研究其次应用了强化学习中的竞争深度 Q 网络 (Dueling deep Q-networks, Dueling

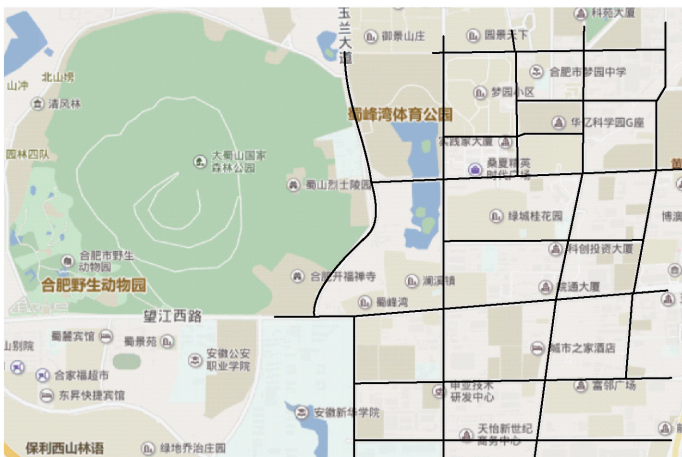
DQN),改进了其网络结构,构建了混合学习优化算法(hybrid learning optimization algorithm, HLOA),并设计了离线在线多步学习策略,来实现问题的在线求解。最后,通过基于真实路网的实验,验证了预测交通信息的作用,证明了 HLOA 的功能和结构可行性。

## 1 问题描述

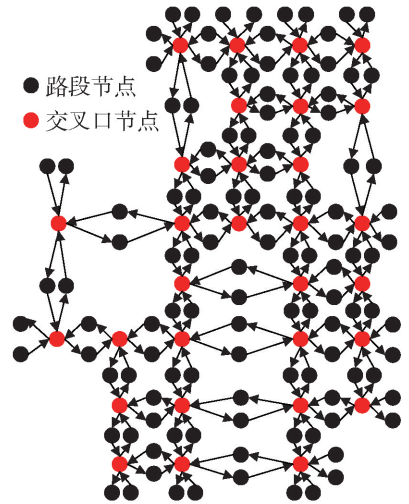
### 1.1 路网描述与交通状态预测

本研究的在线动态电动汽车充电路径规划基于城市街道路网,路网中,街道基本路段和交叉口存在交通流的动态差异,采用异构动态图描述可以准确

反映其实时状态。如图 1 所示,以某市局部路网为例构建了异构图,其中图 1(a)为真实路网地图,图 1(b)为对应的路网异构图,在异构图中赋予动态权值,可获得异构动态图  $\mathcal{S}(V, E, W)$ 。在图  $\mathcal{S}$  中,  $V$  为节点集,  $E$  为边集,  $W$  为节点权值集合,其中,  $V$  可分为路段节点集  $V_L$  和交叉口节点集  $V_I$ , 它们的元素为编号  $i$  的节点;  $E$  中的元素为  $e_{ij}$ , 其表示节点  $i$  与节点  $j$  的交通流连接关系,此外,通过  $E, \mathcal{S}$  的邻接矩阵  $A$  也可以获得;  $W$  的元素为节点上对应的权值,本文中节点的权值为  $\Delta t^l(i)$  (即节点  $i$  在时刻  $t$  的行程时间)和  $l(i)$  (即节点  $i$  的长度)。



(a) 实际路网



(b) 路网异构图

图 1 实例路网及其异构图

Fig. 1 Example network and its heterogeneous graph

由于交通流的时空异质性,实时行程时间标定异构动态图会造成信息延误,而采用动态图卷积递归网络(dynamic graph convolutional recurrent network, DGCRN)<sup>[28]</sup>对异构图进行行程时间预测和

权值再标定,实现了对信息延误的规避。预测网络的体系架构如图 2 所示,该网络结构依托序列到序列体系<sup>[29]</sup>,将动态图生成器和动态图卷积模块内嵌到编码和解码器中实现行程时间预测。该体系的输

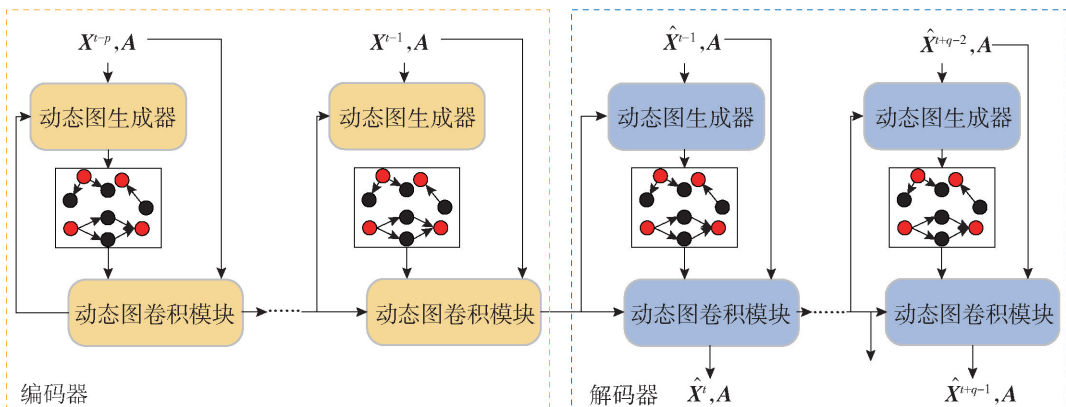


图 2 动态图卷积递归网络(DGCRN)的结构

Fig. 2 Architecture of DGCRN

人为状态张量  $\mathbf{X} \in \mathbb{R}^{p \times n \times d}$  和邻接矩阵  $\mathbf{A}$ ,  $p$  为张量的时间维度即输入的历史周期数,  $n$  为空间维度即输入的节点个数,  $d$  为特征维度即输入特征数。本文  $\mathbf{X}, p=12$  即  $\mathbf{X} = [\mathbf{X}^{t-12}, \mathbf{X}^{t-11}, \dots, \mathbf{X}^{t-1}]$ ,  $n=150$  即  $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{150}]^T$ ,  $d=2$  即  $\mathbf{X} = [\Delta t, \mathbf{t}]^T$ ,  $\Delta t$  为历史行程时间矩阵,  $\mathbf{t}$  为采集时间矩阵; 输出为状态张量  $\hat{\mathbf{X}} \in \mathbb{R}^{12 \times 150 \times 1}$ ,  $\hat{\mathbf{X}} = \Delta \hat{\mathbf{t}}, \Delta \hat{\mathbf{t}}$  为预测的行程时间矩阵。

动态图生成器通过动态分析路网交通流状态和拓扑结构, 来动态地获得交通流的空间关系, 生成动态异构图。动态图生成器的结构如图 3 所示, 由图卷积层、动态滤波器和嵌入节点 3 个模块构成, 图卷积层和动态滤波器对异构交通流的空间关系进行动态分析, 嵌入节点则是将路网固有的 OD (origin-destination) 特征体现在动态图中, 其模块的设计, 本文参照文献[28]。最终生成的动态图将作为动态图卷积模块的输入, 就可解释性来说, 动态图的作用是预处理交叉口的交通状态并反映交叉口的主要通路。

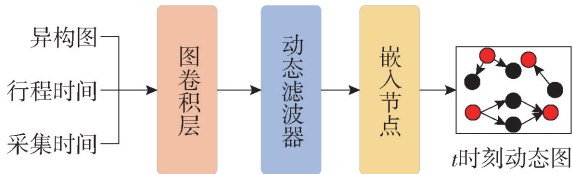


图 3 动态图生成器的结构

Fig. 3 Structure of dynamic graph generator

动态图卷积模块是 DGCRN 是实施预测的关键模块, 其结构如图 4 所示, 由动态图卷积层、静态图卷积层、加权求和层、串联层及全连接层组成。该模块的输入为生成动态图, 状态张量和静态异构图组成, 生成动态图由动态图卷积层进行重构, 其余输入由静态卷积层处理, 处理结果通过加权求和层进行

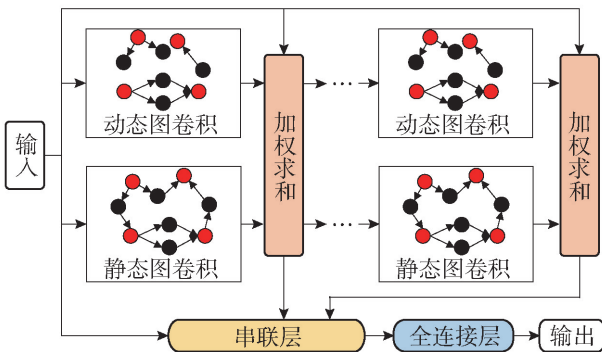


图 4 动态图卷积模块的结构

Fig. 4 Structure of dynamic graph convolution module

结合。本文中的结构配置、参数初始化及网络的学习策略均参考文献[28]。

### 1.2 电动汽车建模

路网建模后, 本文构建电动车的运动学模型和动力学模型。电动车运动学模型主要用于更新电动车位置并影响着动力学模型, 而动力学模型则用于更新电动车的剩余电量。

本文的电动车动力学模型将行驶的电动车看作质点, 则路段  $i$  中电动车在连续时域的状态方程为

$$\begin{cases} \dot{l}_i(t) = v_i(t) \\ \dot{v}_i(t) = a_i(t) \end{cases} \quad (1)$$

式中:  $l_i(t) \in [0, l_i]$  为路段  $i$  中电动汽车行驶距离, 当  $l_i(t) > l_i$ , 电动车所在路段实现切换;  $v_i(t) \in [0, v_{i,max}]$  为路段  $i$  中电动车的速度;  $a_i(t) \in [a_{i,min}, a_{i,max}]$  为加速度, 这两者决定了路段中电动车的能耗。

基于上述变量和动能定理, 路段  $i$  电动车的能耗方程为

$$E^t(i) = \int_i^{t+\Delta t_i} \psi(v_i(t), a_i(t)) dt = \int_i^{t+\Delta t_i} m_j v_i(t) |a_i(t)| dt \quad (2)$$

式中:  $E^t(i)$  为路段  $i$  电动车通过整个路段的能耗;  $\psi(\cdot)$  为瞬时能耗函数;  $m_j$  为电动车质量, 本文  $m_j = 200$  kg, 电动车  $j$  为行驶在路段  $i$  中的电动车。路段  $i$  中电动车离开路段剩余电量 Soc( $i$ ) 为

$$\text{Soc}(i) = \text{Soc}(i-1) - \frac{E^t(i)}{\eta} \quad (3)$$

式中:  $\eta$  为电动车的机械能传动效率, 其影响因素涉及电池、电动机、变速箱和驱动桥的传动功率, 最终  $\eta = 73.75\%$ 。

基于电动汽车位置, 行程时间和剩余电量, 电动汽车通过路段  $k$  对应的状态向量为  $\mathbf{x}(k) = [k, \Delta t^t(k), \text{Soc}(k)]^T$ 。则基于离散空间域的状态方程可描述为

$$\mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k), k) \quad (4)$$

式中:  $\mathbf{x}(k+1)$  为电动车通过路段  $k$  后进入的状态;  $\mathbf{u}(k)$  为电动车的决策变量, 其可以控制电动车在当前状态进行的行驶决策,  $\mathbf{u}(k) = [b_k]$ ,  $b_k \in B$  为电动车经过路段  $k$  后的选择行为, 行为集合  $B$  包含直行, 左转, 右转和调头 4 种行为。

### 1.3 在线电动车路径问题

电动车充电路径规划问题旨在保证其电量充足的情况下, 获得行程时间最短的行驶策略, 并最终到

达充电站。基于电动车状态方程和最优控制理论, 电动车充电路径规划问题表述为

$$\begin{cases} \min \left\{ J(k) = \sum_{k=i_0}^{i_d} g(\mathbf{x}(k), \mathbf{u}(k), k) \right\} \\ \text{s. t. } \mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k), k) \quad k \in V \\ \Delta t'(k) \geq \Delta t_{\min}(k) \\ \text{Soc}_{\min} \leq \text{Soc}(k) \leq \text{Soc}_{\max} \\ b_k \in B \\ i_d \in V_s \end{cases} \quad (5)$$

式中:  $i_0$  为电动车的路径起点;  $i_d$  为充电站节点;  $\Delta t'(k) = g(\mathbf{x}(k), \mathbf{u}(k), k)$ ,  $\Delta t_{\min}(k)$  为路段  $k$  的最短行程时间;  $\text{Soc}_{\min}$  为电动车电池组的最低电量;  $\text{Soc}_{\max}$  为电池组的最高电量;  $V_s$  为充电站节点集合。式(5)的求解过程本质上是求解最优控制策略  $\mathbf{u}^*$ , 即

$$\mathbf{u}^* = \arg \min_{\mathbf{u}(k)} \left\{ J(k) = \sum_{k=i_0}^{i_d} g(\mathbf{x}(k), \mathbf{u}(k), k) \right\} =$$

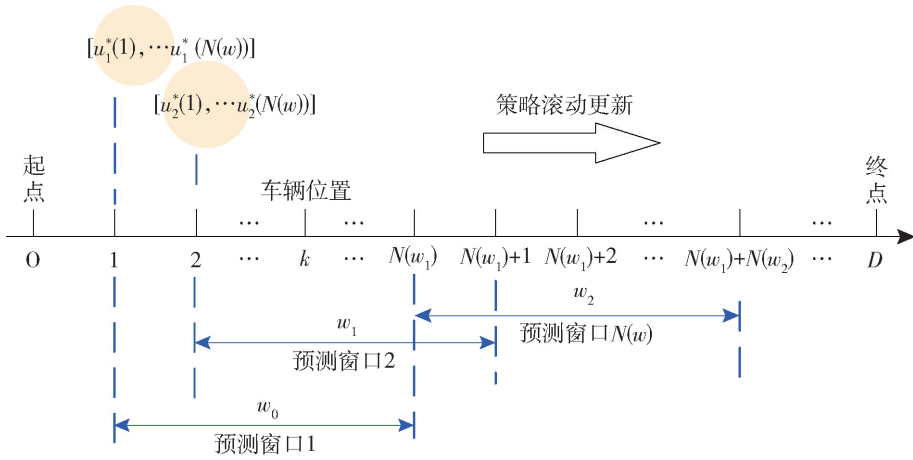


图5 空间域 MPC 滚动优化

Fig. 5 Rolling optimization of the MPC in spatial domain

$$\begin{cases} \min \left\{ J(k) = \sum_{s=1}^{N(w)} g(\mathbf{x}(k+s), \mathbf{u}(k+s), k+s) \right\} \\ \text{s. t. } \mathbf{x}(k+1) = f(\mathbf{x}(k+s), \mathbf{u}(k+s), k) \quad k \in V_r \\ \Delta t'(k) \geq \Delta t_{\min}(k) \\ \text{Soc}_{\min} \leq \text{Soc}(k) \leq \text{Soc}_{\max} \\ b_k \in B \\ i_d \in V_s \end{cases} \quad (7)$$

式中:  $N(w)$  为  $w$  段路径中包含的节点数;  $w$  为切换点间的空间距离也是预测长度;  $V_r$  为路径覆盖的节点集合, 此时模型的解为

$$\{\mathbf{u}(i_0), \dots, \mathbf{u}(k), \dots, \mathbf{u}(i_d-1)\} \quad (6)$$

$\mathbf{u}^*$  是全局最优的策略, 车辆执行该策略可以获得地图中行程时间最短的路径。但在线寻优时, 交通流的实时变化带来路段行程时间的实时变化, 求解的问题是一类能在多项式时间复杂度内归约的问题 (nondeterministic polynomial hard, NP-hard), 全局最优解的计算时间复杂度和空间复杂度大。因此在线规划充电路径需要将全局最优求解转化为滚动局部最优求解, 具体到本问题即将最优控制问题转化为模型预测控制 (model predictive control, MPC) 问题。转化的方式是基于标称路径完成的, 标称路径的获得是根据最短行程时间的时空分布离线获取的多条路径。在转化过程中, 标称路径是空间离散的基础, 多条路径的交叉点作为切换点, 则切换点作为空间离散点, 如图 5 所示, MPC 是以滚动优化对模型进行重构, 则基于 MPC 的电动车在线路径规划模型如下所示。

策略滚动更新

预测窗口  $N(w)$

$$\mathbf{u}_k^* = \begin{cases} [\mathbf{u}_k^*(1), \dots, \mathbf{u}_k^*(N(w))], & 0 < N(w) \leq |D - k| \\ [\mathbf{u}_k^*(1), \dots, \mathbf{u}_k^*(D - k)], & N(w) > |D - k| \end{cases} \quad (8)$$

$w \in [1, |D - 0|]$

式中:  $| \cdot |$  表示两节点间的空间距离, 本文采用其间节点数量来度量。

## 2 算法设计

为求解上述电动汽车充电路径规划问题, 本文提出 HLOA, 其整体基于强化学习, 结合了先验空间约束和离线学习, 实现混合多步求解, 其框架如图 6 所示。HLOA 可分为离线求解和在线求解 2 个部

分, 离线部分为在线部分提供空间约束和在线学习的参数初始化, 而在线部分进行实时求解。

### 2.1 离线求解

离线求解部分根据功能, 可将图 6 中的步骤组合为时空预处理、标称路径获取和学习参数初始化 3 个主要模块, 时空预处理包括划分时段和标定路网 2 个

部分, 根据日交通状态规律, 行程时间数据分为早高峰、平峰、晚高峰及夜间 4 个时间段的数据, 其中早高峰时段为 7:00—9:00, 平峰时段为 9:01—15:59, 晚高峰时段为 16:00—19:00, 夜间时段为 19:01—6:59(次日)。基于时段划分, 动态异构图的行程时间采用对应时刻的历史平均值进行标定。

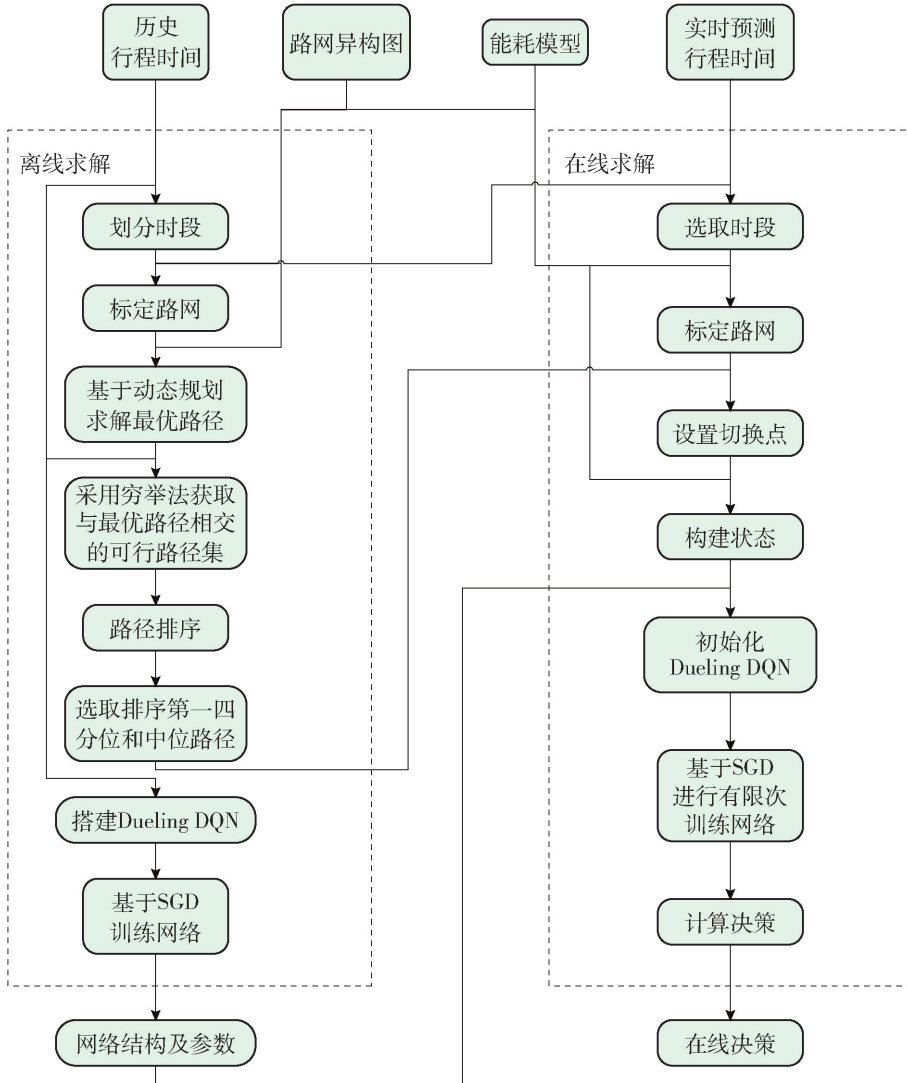


图 6 混合学习优化算法的框架

Fig. 6 Framework of HLOA

标称路径获取包括动态规划求解最优路径、可行路径集的获取、路径排序及路径提取。标称路径的获取基于以下前提:

- 1) 本文所述标称路径的起点覆盖路网中所有的路段节点。
- 2) 每个节点的标称路径数目由充电站的数目决定, 而每个充电站采用 3 条路径作为标称路径。
- 3) 获得最优标称路径的方法为动态规划算法, 算法来自于文献[13], 其求解的问题为式(5)所示

的最优控制问题。

4) 每个节点的标称路径中必包含到充电站的最优路径, 而其他 2 条路径为排序后可行路径集中第一个四分位数和中位数路径, 可行路径需与最优路径有交点, 且满足能量消耗约束。

学习参数初始化是基于强化学习搭建在线求解模型的框架并对参数进行预学习, 该模块包含图 6 中搭建 Deuling DQN 和基于 SGD 训练网络 2 个步骤。在 Deuling DQN 中, 智能体 (Agent) 的

状态  $s_i$  为车辆的状态变量,智能体的动作  $a_i$  为车辆的输入变量,智能体的动作价值函数  $Q(s_i, a_i)$  为

$$Q(s_i, a_i) = E[c_i | s_i, a_i] \quad (9)$$

式中:  $c_i = \sum_{k=i}^{i_d} \lambda_k r_{i+k}$  为节点  $i$  策略的回报,  $\lambda_k = 1$  为折扣系数,  $r_{i+k} = \Delta t^l(i)$  是节点  $i+k$  的惩罚,因此  $Q(s_i, a_i)$  被重构为

$$Q(s_i, a_i) = \frac{1}{N(V_r)} \sum_{k=i_0}^{i_d} g(x(k), u(k), k) \quad (10)$$

式中  $N(V_r)$  为策略路径覆盖的节点个数。

Deuling DQN 是基于值函数的深度强化学习框

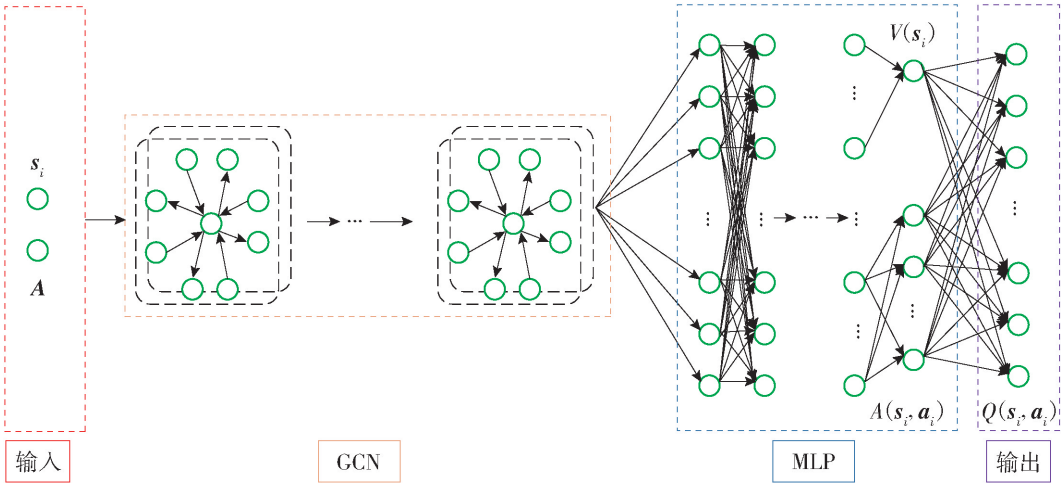


图7 Deuling DQN 的结构

Fig. 7 Structure of Deuling DQN

本文中,GCN 模块包含 5 个 GCN 层,其计算规则为

$$\mathbf{H}^k = \text{Re lu} \left( \sum_{p=0}^P \mathbf{H}^{p-1} \mathbf{O}^{p-1} \right)$$

$$\mathbf{H}^p = \tanh \left( \mathbf{D}^{-\frac{1}{2}} \mathbf{X} \mathbf{D}^{-\frac{1}{2}} \mathbf{\Psi}^{p-1} \mathbf{H}^{p-1} \mathbf{W}^{p-1} \right) \quad (12)$$

式中:  $\mathbf{H}^k$  为 GCN 中  $k$  层的输出;  $\mathbf{H}^p$  为  $k$  层中  $p$ -hop 的输出;  $\mathbf{X} = \mathbf{A} + \mathbf{I}$ ,  $\mathbf{I} \in \mathbb{R}^{150 \times 150}$ ,  $\mathbf{\Psi}^{p-1} \in \mathbb{R}^{150 \times 450}$ ,  $\mathbf{W}^p$ ,  $\mathbf{O}^{p-1} \in \mathbb{R}^{3 \times 3}$  是学习参数,  $\mathbf{D}$  为  $\mathbf{A}$  的度矩阵,  $\text{Re lu}(\cdot)$  是 ReLU 函数,  $\tanh(\cdot)$  为 tanh 函数。

同时,MLP 模块包含 7 个全连接层作为隐藏层和 2 个输出层,每层包含 7 650 个神经元,其计算规则为

$$\mathbf{H}^{(k)} = \tanh \left( \mathbf{H}^{(k)} \mathbf{W}^{(k)} + \mathbf{B}^{(k)} \right) \quad (13)$$

式中:  $\mathbf{H}^{(k)}$  是全连接层的输出;  $\mathbf{W}^{(k)}$  和  $\mathbf{B}^{(k)}$  为学习参数。以上参数的学习基于随机梯度下降法 (stochastic gradient descent, SGD) 进行训练。而输出层包括中间输出层和转化层,中间输出层获得从隐含层计算的

架,其损失函数为

$$\mathcal{L} = \frac{1}{n(n-1)} \sum_{i \in V} (\min(Q(s_i, a_i)) - Q(s_i, a_i^*)) \quad (11)$$

式中  $Q(s_i, a_i^*)$  为最优路径对应的动作价值函数。基于该损失函数,本文构建了如图 7 所示的神经网络结构,其包含图卷积 (graph convolution neural networks, GCN) 模块和多层感知机 (multilayer perceptron, MLP) 模块,网络的输入为  $s_i$  和  $A$ ,输出为  $Q(s_i, a_i)$ ,而  $Q(s_i, a_i)$  由状态价值函数  $V(s_i)$  和动作优势函数  $A(s_i)$  再计算获得,且最小的  $Q(s_i, a_i)$  将作为计算的最优控制问题的解。

$V(s_i)$  和  $A(s_i)$ ,而转化层的形式和隐藏层相同,其功能是利用  $V(s_i)$  和  $A(s_i)$  计算各动作的  $Q(s_i, a_i)$ 。

## 2.2 在线求解

在线求解以离线求解的结果作为空间约束和初始参数,对求解模型进行再训练。其中,训练时间不得超过车辆通过节点的行程时间,当在时间约束中无解时,车辆将采用离线计算结果。由于在线路网的权值是时变的,获得全局最短路径是不现实的,因此,在线学习的损失函数被重构为

$$\mathcal{L}_{new} = \min \left( \frac{1}{N(w)} \sum_{s=1}^{N(w)} g(x(k+s), u(k+s), k+s) + \gamma h(x(k+s), u(k+s), k+s) \right) \quad (14)$$

式中:  $\gamma > 0.1$  为权值参数;  $h(x(k+s), u(k+s), k+s)$  为能耗约束项,其形式为

$$h(x(k), u(k), k) = \begin{cases} \text{Soc}(k) - \text{Soc}_{\min} & \text{Soc}(k) > \text{Soc}_{\min} \\ M & \text{Soc}(k) \leq \text{Soc}_{\min} \end{cases} \quad (15)$$

式中  $M$  表示无穷大。

在线部分中,通过训练后的网络获得了确定的结构和参数。然后,将  $s_i$  和  $A$  逐一输入网络,HLOA 可实时获得如式 8 所示行驶策略序列。车辆执行策略时,在每个节点处,其只执行序列中首个策略,策略为输出的在线决策,最后,基于上述方法,循环往复直至到达充电站。

### 3 实验分析

本部分通过实验,分析预测交通信息的价值和 HLOA 的可行性。实验采用的性能指标为竞争比<sup>[30]</sup>,竞争比是将在线算法表现与离线算法最优表现进行比较的指标,该指标  $\varphi$  在本文中的数学表达为

$$\varphi = \frac{\sum_{i=i_0}^{i_d} Q(s_i, a_i)}{\sum_{i=i_0}^{i_d} \hat{Q}(s_i, a_i)} \quad (16)$$

式中:  $\hat{Q}(s_i, a_i)$  为基于收集的时变信息,逐一采用动态规划法离线求得行驶策略价值函数。

该研究所有实验场景为如图 1 所示的某市局部路网,场景中交通状态数据来源于视频,视频中交通数据的提取方法来自文献[31]。场景共设置充电站 3 座分别位于路网的北侧、东侧和南侧。为了分析方法泛化能力和可视化车辆位置,基于交通仿真软件 simulation of urban mobility (SUMO),构建了仿真路网如图 8 所示。实验分析的时间范围为工作日的 24 h,其中 6:00—15:30 的交通状态为路网真实值,15:31—5:59(次日)的交通状态为路网输入总量对应的各节点仿真值。交通实验目标车辆的起点为随机选择的路网入口,共安排了 150 辆目标车辆均匀地分布在实验时段,最终实验的统计指标为所有策略的平均值  $E(\varphi)$  和标准差  $S(\varphi)$ 。

实验的硬件平台为 Amax 工作站,编程语言为 Python。其中 Python 的版本为 3.7,涉及 PyTorch, Pandas, Numpy 等多个库,与 SUMO 的交互基于 TraCI 接口实现。

#### 3.1 预测信息价值分析

预测信息价值分析主要分析采用预测行程时间和实时行程时间标定路网,有利于在线获得最优策略;此外,还分析了预测周期长度对获取最优路径的影响。实验分别采用预测周期为 10 s 的行程时间、出发时刻的行程时间及更新周期为 10 s 实时行程时间对路网进行标定,计算  $E(\varphi)$  和  $S(\varphi)$ ,而预测

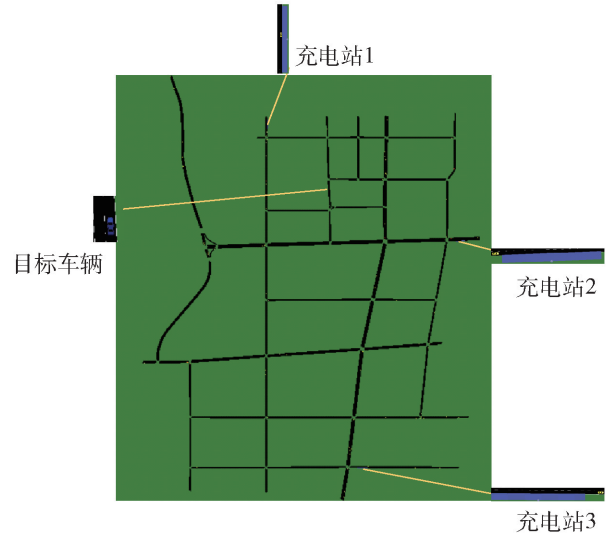


图 8 仿真场景

Fig. 8 Simulation scenario

周期的选取是基于周期比  $\theta$ ,其计算式为

$$\theta = \frac{\delta}{\sum_{k=i_0}^{i_d} \Delta t^k(k)} \quad (17)$$

式中:  $\delta$  为预测周期;实验中  $\theta$  的取值范围为 0 ~ 2,所有取值形成了以 0.05 递增的等差数列。

基于以上实验设置,获得如表 1 所示的实验结果。从表 1 中可发现预测信息可以获得最低的  $E(\varphi)$ ,但其  $S(\varphi)$  并不是三者中最低的,这说明预测信息的自身存在的误差会影响路径规划的准确性。此外快速更新的实时信息也可以降低部分信息的延误,降低算法误差,在路径精确度不是很高的场景中快速更新的交通状态作为路网权值。不更新的实时信息却存在较大偏差,说明电动车充电路径问题的解对交通状态具有较高的时间敏感度。

表 1 预测信息价值分析结果

Table 1 Analysis results of predicted information value

输入	预测行程 时间	出发时刻 行程时间	实时更新 行程时间
$E(\varphi)$	1.32	4.65	1.42
$S(\varphi)$	0.29	1.32	0.25

预测周期长度影响的分析结果如图 9 所示,从中可发现预测或更新周期越小,算法求解的策略越接近最优策略,但就计算的时间和空间成本来说,周期越大成本越低。而当  $\theta = 0.25$  时,预测信息下的算法仍能保证对最优策略的最大近似拟合,相对的实时信息下的算法却出现了误差增加的情况,由此可见,预测周期最可行。此外,从图 9 中可见,当周



期比不断增加时,较预测信息,实时信息下的算法精度下降更明显,究其原因是在实时信息不包含未来信息,低频更新不能满足交通流的时间敏感性,这再次印证采用预测信息更有价值。

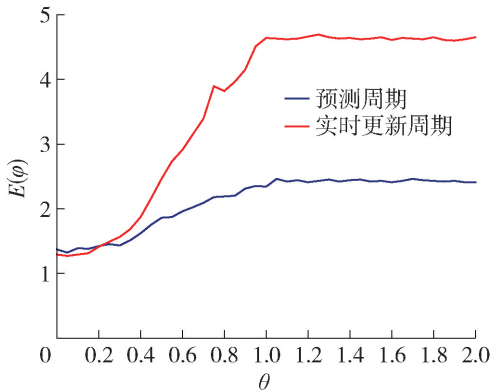


图9 预测周期长度对解的影响

Fig. 9 Influence of prediction period on solution

综上所述,HLOA 采用预测交通状态作为路网标定有助于保证算法的求解精度。

### 3.2 HLOA 可行性分析

HLOA 可行性分析包括求解可行性和结构可行性分析,求解可行性分析比较了 HLOA 算法,价值逼近算法 (value approximation iteration, VAI)<sup>[19]</sup> 和混合元启发式算法 (hybrid metaheuristic algorithm HMA)<sup>[27]</sup> 对应  $\varphi$ 。结构可行性分析比较了 GCN 模块对 DQN 网络中 CNN 模块的替代,对计算结果的影响。

求解可行性从统计参数、节点规模和时间域进行了比较。其中统计参数为  $E(\varphi)$  和  $S(\varphi)$ ;节点规模旨在分析最优决策对应节点的数目变化对各算法精度的影响,其反映了搜索路网范围变化对算法精度的影响;时间域比较统计了每 15 min 进入路网的交通流数据和各算法对应的  $E(\varphi)$ ,反映了时变交通流对算法精度的影响。

统计参数的比较结果如表 2 所示,其中可知,HLOA 在 2 个统计参数上均优于另外 2 个模型,且误差平均降低了 33.65%。这在宏观上,说明 HLOA 的在线求解结果最解近最优策略。

表 2 各算法的统计参数

Table 2 Statistical parameters of the algorithms

算法	HLOA	VAI	HMA
$E(\varphi)$	1.32	2.14	1.86
$S(\varphi)$	0.29	0.84	0.67

节点规模的比较结果如图 10 所示,可知,随着节点数目的增加,各算法精度是波动下降的,但 HLOA 的下降速率相对其他 2 种是较慢的,而且其波动幅度在三者中也是较小的,误差最大增加了 1.29,另外两模型分别为 3.59 和 2.54。因此,HLOA 的解具有可行性,且其受到路网规模的影响也是相对较小的。

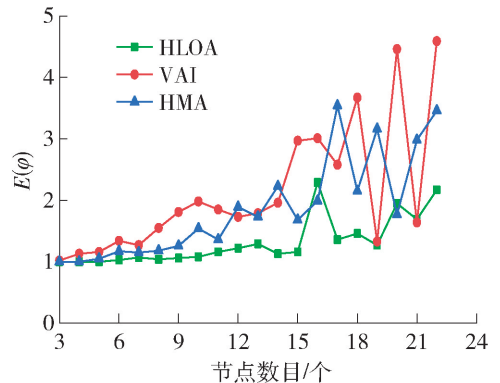
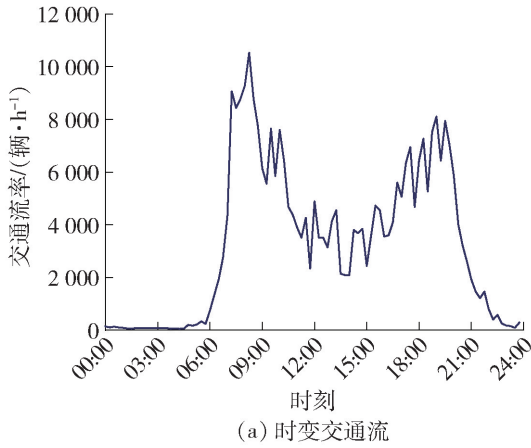


图 10 各算法的空间稳定性

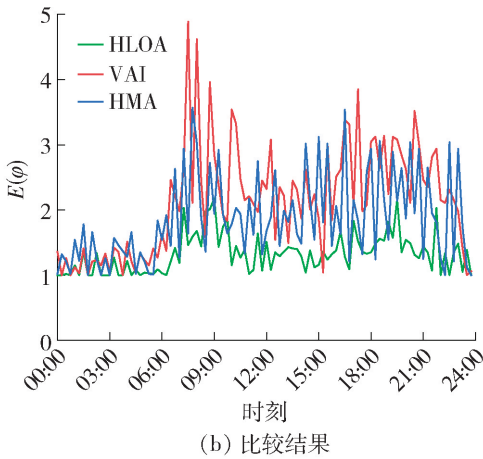
Fig. 10 Spatial stability of the algorithms

时间域实验中路网输入的时变交通流如图 11(a) 所示,其对应的比较结果如图 11(b) 所示。随着交通流变化的剧烈,各算法的精度存在一定程度的波动,但 HLOA 的波动幅度较小,且精度较其他 2 种算法保持在较高水平。HLOA 的精度对高峰时段的密集交通流并不敏感,但会受到交通流剧烈变化的影响,出现误差增加的情况。基于以上讨论,HLOA 在时间域中具有较高的精度和稳定性。

为了处理动态异构图,HLOA 中 DQN 采用 GCN 模块代替了传统 DQN 中的 CNN 模块,这一结构改进需要通过比较实验验证,而验证的指标除了  $E(\varphi)$  和  $S(\varphi)$ ,还包括在线训练的平均次数。HLOA 的比较对象为改用 CNN 模块的 HLOA (CHLOA),与 HLOA 不同 CHLOA 仅能处理栅格化的地理信息数据;则 CHLOA 的输入为栅格化后的图 1(a) 路网,栅格的规模为  $84 \times 84$ ,每个栅格中的数值为预测的行程时间,当栅格中没有道路元素时,其中的值为无穷大,CHLOA 中 CNN 模块的配置来自于文献[32]。HLOA 与 CHLOA 的比较结果如表 3 所示,其中 HLOA 的所有性能指标均优于 CHLOA,可见本文在结构改进方面是可行的。分析 HLOA 性能优越的原因是采用异构动态图结构输入的信息更精练,其可以忽略栅格图中的无用信息和冗余信息。



(a) 时变交通流



(b) 比较结果

图 11 各算法的时间稳定性

Fig. 11 Time stability of each algorithm

表 3 HLOA 与 CHLOA 的比较结果

Table 3 Comparison results between HLOA and CHLOA

算法	HLOA	CHLOA
$E(\varphi)$	1.32	1.56
$S(\varphi)$	0.21	0.42
平均在线训练次数	26	37

综上所述,采用 HLOA 求解在线电动车充电路径规划问题可以近似拟合最优决策,且在不同搜索范围和时变交通流的影响下,具有较高的适应性和稳定性,尤其在较大范围搜索过程和高峰时段中仍能保证较高的求解精度。

## 4 结论

1) 采用了预测的交通状态信息来标定车辆行驶路网,考虑了交通状态的时间敏感度,并通过实验分析了其最优预测周期为周期比 0.25,验证了预测交通信息相较于实时交通信息在电动车充电路径规

划上的优势。

2) 为了在线求解电动车充电路径规划问题,提出了混杂学习优化算法,并通过实验验证了其对问题求解的高精度,总体误差较其他在线求解算法降低了 33.65%,实验也反映出其不同搜索范围和时变交通流的影响下,具备较强的适应性和稳定性。

3) 基于异构动态基本图对路网的描述,采用了图卷积神经网络来搭建 HLOA 的 DQN,并通过与传统采用卷积神经网络的 DQN 进行对比,发现基于 GCN 的改进结构提高了算法精度具有可行性。

## 参考文献:

- [1] 陈莹莹. 纯电动汽车应用于城市物流配送中的路径规划研究[D]. 北京: 华北电力大学, 2021.  
CHEN Y Y. Research on path planning of pure electric vehicles applied in urban logistics distribution [D]. Beijing: North China Electric Power University, 2021. (in Chinese)
- [2] KESKIN M, CATAY B. A Mathuristic method for the electric vehicle routing problem with time windows and fast chargers[J]. Computers & Operations Research, 2018, 100: 172-188.
- [3] PSARAFTIS H N, MIN W, KONTOVAS C A. Dynamic vehicle routing problems: three decades and counting[J]. Networks, 2016, 67(1): 3-31.
- [4] SUN Z, ZHOU X. To save money or to save time: Intelligent routing design for plug-in hybrid electric vehicle [J]. Transportation Research Part D: Transport and Environment, 2016, 43: 238-250.
- [5] LIANG S, HU S, TING W, et al. An energy-efficient reliable path finding algorithm for stochastic road networks with electric vehicles[J]. Transportation Research Part C: Emerging Technologies, 2019, 102: 450-473.
- [6] PERGER T, AUER H. Energy efficient route planning for electric vehicles with special consideration of the topography and battery lifetime [J]. Energy Efficiency, 2020, 13(8): 1705-1726.
- [7] YAN D, GUO J. A multiclass cumulative prospect theory-based stochastic user equilibrium model with path constraints in degradable transport networks [J]. Promet-Traffic &Transportation, 2021, 33(5): 775-787.
- [8] ZHONG J Q, YANG N, ZHANG X H, et al. A fast-charging navigation strategy for electric vehicles considering user time utility differences [J]. Sustainable Energy, Grids and Networks, 2022, 30: 100646.
- [9] KESKIN M, CARAY B. Partial recharge strategies for the electric vehicle routing problem with time windows [J]. Transportation Research Part C: Emerging Technologies,

- 2016, 65: 111-127.
- [10] YI Z, BAUER P H. Optimal stochastic eco-routing solutions for electric vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(12): 3807-3817.
- [11] RIOS B H O, XAVIER E C, MIYAZAWA F K, et al. Recent dynamic vehicle routing problems: a survey[J]. *Computers & Industrial Engineering*, 2021, 160: 107604.
- [12] PRAVEEN V, KEERTHIKA P, SARANKUMAR A, et al. A survey on various optimization algorithms to solve vehicle routing problem [C] // 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS). Piscataway NJ: IEEE, 2019: 134-137.
- [13] PSARAFTIS H N. A dynamic programming solution to the single vehicle many-to-many immediate request dial-a-ride problem [J]. *Transportation Science*, 1980, 14(2): 130-154.
- [14] LEE, J. Dispatching rail-guided vehicles and scheduling jobs in a flexible manufacturing system[J]. *International Journal of Production Research*, 1999, 37(1): 111-123.
- [15] MONROY-LICHT M, AMAYA C A, LANGEVIN A, et al. The rescheduling arc routing problem [J]. *International Transactions in Operational Research*, 2017, 24: 1325-1346.
- [16] ULMER M W, STRENG S. Same-day delivery with pickup stations and autonomous vehicles[J]. *Computers & Operations Research*, 2019, 108: 1-19.
- [17] ULMER M W, MATTFELD D C, KSTER F. Budgeting time for dynamic vehicle routing with stochastic customer requests[J]. *Transportation Science*, 2018, 52: 20-37.
- [18] WU Y, LU X. Capacitated vehicle routing problem on line with unsplittable demands [J]. *Journal of Combinatorial Optimization*, 2022, 44: 1953-1963.
- [19] 张晓楠, 张建雄. 随机需求车辆路径问题的价值逼近在线决策[J]. *控制理论与应用*, 2022, 39(2): 241-254.
- ZHANG X N, ZHANG J X. Value-approximation-based online policy for vehicle routing problem with stochastic demand[J]. *Control Theory & Applications*, 2022, 39(2): 241-254. (in Chinese)
- [20] ILIOPOULOU C, KEPAPSOGLU K. Integrated transit route network design and infrastructure planning for on-line electric vehicles[J]. *Transportation Research Part D: Transport and Environment*, 2019, 77: 178-197.
- [21] MONTROYA A, JOSE-ALEJANDRO, MENDOZA J E, et al. A multi-space sampling heuristic for the green vehicle routing problem [J]. *Transportation Research Part C*, 2016, 70: 113-128.
- [22] 杨立炜, 付丽霞, 郭宁, 等. 多因素改进蚁群算法的路径规划[J]. *计算机集成制造系统*, 2021: 1-18.
- YANG L W, FU L X, GUO N, et al. Path planning with multi-factor improved ant colony algorithm[J]. *Computer Integrated Manufacturing Systems*, 2021: 1-18. (in Chinese)
- [23] LI J, XIN L, CAO Z, et al. Heterogeneous attentions for solving pickup and delivery problem via deep reinforcement learning [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(3): 2306-2315.
- [24] CHEN X, ULMER M W, THOMAS B W. Deep Q-learning for same-day delivery with vehicles and drones [J]. *European Journal of Operational Research*, 2022, 298(3): 939-952.
- [25] LIN B, GHADDAR B, NATHWANI J. Deep reinforcement learning for the electric vehicle routing problem with time windows [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(8): 11528-11538.
- [26] POP C P, MATEI O, SITAR C P. An improved hybrid algorithm for solving the generalized vehicle routing problem[J]. *Neurocomputing*, 2013, 109: 76-83.
- [27] GUTIERREZ A, DIEULLE L, LABADIE N, et al. A hybrid metaheuristic algorithm for the vehicle routing problem with stochastic demands [J]. *Computers and Operations Research*, 2018, 99: 135-147.
- [28] LI F, FENG J, YAN H, et al. Dynamic graph convolutional recurrent network for traffic prediction: benchmark and solution [J]. *ACM Transactions on Knowledge Discovery from Data*, 2023, 17(1): 1-21.
- [29] YOON K. Sequence-to-sequence learning with latent neural grammars [J]. *Advances in Neural Information Processing Systems*, 2021, 31: 26302-26317.
- [30] BORODIN A, EL-YANIV R. On-line computation and competitive analysis [M]. Cambridge: Cambridge University Press, 1998: 30-47.
- [31] SHI Z Y, CHEN Y Z, MA P F. Video data based traffic state prediction at intersection [C] // 2020 IEEE 23rd International Conference on Intelligent Transportation Systems. Piscataway NJ: IEEE, 2020: 1-6.
- [32] FREITAS N D, LANCTOT M, HASSELT H V, et al. Dueling network architectures for deep reinforcement learning [C] // The 33rd International Conference on Machine Learning. Piscataway NJ: IEEE, 2016: 2939-2947.