

基于图神经网络多智能体强化学习的电力-交通融合网协同优化运行

江昌旭, 卢玥君, 邵振国, 林俊杰

(福州大学电气工程与自动化学院(福建省电器智能化工程技术研究中心), 福州 350108)

摘要: 针对多重不确定性因素下配电系统、交通系统和电动汽车动态交互的电力-交通融合网序贯协同优化问题, 提出一种基于图神经网络多智能体强化学习的电力-交通融合网协同优化方法。首先, 基于图理论方法将电动汽车间的相互影响关系转换为一种动态网络图模型, 采用一种基于注意力机制的图神经网络多智能体强化学习算法求解电动汽车充电引导策略, 探讨电动汽车多智能体间的相互影响作用。然后, 在含可再生能源出力的主动配电网中采用二阶锥优化及对偶优化理论对配电网最优潮流进行求解, 得到配电网节点边际电价, 研究电力和交通系统的动态交互特性。最后, 在某区域 108 节点交通网络和 IEEE 33 节点电力系统上验证所提方法有效性。

关键词: 电力-交通融合网; 图神经网络; 多智能体强化学习; 电动汽车; 节点边际电价

Collaborative Optimization Operation of Integrated Electric Power and Traffic Network Based on Graph Neural Network Multi-agent Reinforcement Learning

JIANG Changxu, LU Yuejun, SHAO Zhenguo, LIN Junjie

(College of Electrical Engineering and Automation, Fuzhou University (Fujian Smart Electrical Engineering Technology Research Center), Fuzhou 350108, China)

Abstract: Aiming at the sequential collaborative optimization problem of the electric power and traffic network integration network (IETN) with dynamic interaction of power distribution system, transportation system and electric vehicle under multiple uncertain factors, a collaborative optimization method for IETN based on graph neural network multi-agent reinforcement learning is proposed to solve the above problems. Firstly, the interaction relationship among electric vehicles is converted into a dynamic network graph model based on the graph theory. A graph neural network multi-agent reinforcement learning algorithm based on attention mechanism is proposed to solve the charging navigation strategy of electric vehicles, so as to explore the mutual influence among electric vehicles. Then, the optimal power flow of distribution network considering the renewable energy output in an active distribution network is solved, and the marginal price of distribution network node is obtained by the second-order cone optimization and dual optimization theory, so as to study the dynamic interaction characteristics of power and transportation systems. Finally, the simulation verification is carried out on a 108-node traffic network and the IEEE 33-node power system in a certain area.

Key words: integrated electric power and traffic network; graph neural network; multi-agent reinforcement learning; electric vehicle; locational marginal price

0 引言

近年来, 在全球能源危机日益严重和环境问题不断恶化的背景下, 以电动汽车为代表的新能源汽车因具有高效、节能、低噪声和零排放等优势被认为是解决化石能源危机与城市污染的重要举措之一^[1-2]。预计 2025 年我国将实现新能源汽车新车销

售量占汽车新车销售总量的 20%, 2035 年纯电动汽车成为新销售车辆的主流, 公共领域用车全面电动化的目标^[3-4]。

随着电动汽车规模化运行, 电力系统和交通系统将会存在许多的交互融合, 形成电力-交通融合网 (integrated electric power and traffic network, IETN)。目前, 国内外学者对 IETN 协同运行和优化等相关问题进行了广泛研究。文献[5-6]基于充电站充电服务定价模型对电动汽车进行充电引导以实现 IETN

基金资助项目: 福建省自然科学基金(2022J05125; 2021J05134)。
Project supported by Science Foundation of Fujian Province (2022J05125, 2021J05134).

系统综合社会效益最大化。文献[5]提出了一种状态电价策略引导电动汽车在高速公共服务区进行快速充电。但是该方法将电价简单地分为闲电价和忙电价, 并不能较准确地反映充电电价与电动汽车排队长度和充电负荷之间的关系。针对该问题, 文献[6]根据电动汽车快充负荷和充电站负荷阈值建立电动汽车充电站充电电价函数, 以用户充电总成本最小为目标引导用户选择快速充电站。文献[7-10]基于 Dist-Flow 潮流模型和(混合)用户均衡模型分别对电力系统最优潮流和交通流分布进行建模, 以此构建 IETN 系统协同优化模型, 并采用混合整数二阶锥优化、最优响应分解算法等对 IETN 优化模型进行求解。文献[11]基于多代理系统构建了一种考虑电力和交通系统交互的电动汽车充电引导策略, 该充电引导策略以电动汽车时间消耗和充电总成本最小为目标, 采用分布式最小/最大一致性和偏差最小一致性算法进行求解。

然而, 以上针对 IETN 的研究属于确定性需求下的静态优化问题, 没有考虑到 IETN 多种主体、多重不确定性因素在连续时间尺度上的耦合关系, 因此其获得的优化决策存在较大的片面性。实际上, IETN 涉及电动汽车、电力系统和交通系统等多个主体, 包含了多种随机不确定因素。多个主体相互作用、多重随机不确定因素及其耦合关系使得厘清电力和交通系统的交互影响机理以及解决 IETN 协同优化变得更加困难^[12-13]。为了更好地考虑电动汽车充电时间、充电需求、可再生能源间歇性和充电电价等不确定性因素, 有学者采用强化学习(reinforcement learning, RL)和深度强化学习(deep reinforcement learning, DRL)方法对电力、交通系统及其耦合系统进行优化求解^[14-17]。RL 和 DRL 均属于免模型算法, 智能体通过“试错”的方式根据环境反馈的奖惩信号对行为策略进行更新, 旨在获得最大的长期累计回报。相比于传统的数学优化方法存在大量假设和计算效率低的问题, 该类方法非常适用于解决未知环境状态下连续时间尺度的复杂随机决策问题。文献[14]在考虑抵达充电站时间不确定性和充电电价不确定性情况下, 提出一种免模型的线性强化学习方法对电动汽车充电功率和充电站充电电价进行优化。文献[15]在考虑交通状况、电动汽车充电等待时间和充电电价等多重不确定因素下, 构建了以电动汽车的行驶时间和充电成本最小的双层随机优化模型, 提出一种基于 DRL 的电动

汽车充电引导优化方法对该模型进行求解。然而, 文献[14-15]仅针对电网侧的定价策略或者电动汽车侧充电引导问题进行了研究, 并没有考虑电力-交通融合网的动态交互以及协同优化, 同时也没有考虑电动汽车间的交互影响, 缺乏详细的电动汽车交互影响模型。文献[16-17]在电力-交通融合网背景下构建了车-站-路-网的多目标模型, 提出一种图神经网络强化学习方法对电动汽车充电站进行决策, 但是该方法基于最短路径算法计算前往充电目的地时间成本, 没有考虑到道路交通情况对充电路径的影响。此外, 其仅对电动汽车充电最优位置进行了优化, 没有对电力系统的最优经济调度进行优化求解, 即没有考虑到电动汽车充电负荷变化和可再生能源出力对电力系统经济调度的影响及其相互作用。

为了解决以上问题, 本文提出了一种基于图神经网络多智能体强化学习的 IETN 协同优化方法。首先, 基于图理论方法将电动汽车间的相互影响关系转换为一种动态网络图模型, 采用一种基于注意力机制的图神经网络多智能体强化学习来处理不规则非欧式结构图模型, 并制定电动汽车充电引导策略, 以此研究电动汽车多智能体间的相互影响作用。然后, 在考虑可再生能源出力的主动配电网基础上, 通过二阶锥优化及对偶优化理论对配电网最优潮流进行求解, 并得到配电网节点边际电价, 研究电力和交通系统的动态交互特性以及 IETN 协同优化。最后, 在某区域 108 节点交通网络和 IEEE 33 节点电力系统中进行了仿真验证。研究结果表明所提出的方法在考虑 IETN 多种不确定性因素的情况下, 能够有效地降低电动汽车充电总成本, 实现电动汽车以及电力系统协同优化调度。

1 电动汽车充电引导数学模型

电动汽车充电引导行为涉及交通、电力和电动汽车等多个主体, 包含了大量的不确定性因素, 如电动汽车初始充电剩余电量(state of charge, SOC)不确定性, 交通路况的不确定性和充电排队时间不确定性等, 这些不确定因素造成电动汽车充电行为也具有较强的不确定性。电动汽车需要进行充电时, 其根据当前车辆 SOC、交通系统和充电站状况确定最优的充电路线和充电站进行充电, 目的是使得电动汽车充电花费的时间成本和充电成本最小。电动汽车充电引导数学模型^[18]可以构建为:

$$C_{i,c,t} = w_i T_{i,c,t} + F_{i,c,t} \quad (1)$$

式中: w_i 表示第 i 辆电动汽车用户的单位时间价值; $T_{i,c,t}$ 和 $F_{i,c,t}$ 分别表示 t 时刻第 i 辆电动汽车前往充电站 c 所花费的时间成本和充电成本。

1.1 时间成本模型

电动汽车 i 在 t 时刻前往充电站 c 进行充电时所需的时间成本为:

$$T_{i,c,t} = T_{i,c,t}^{\text{road}} + T_{i,c,t}^{\text{wait}} + T_{i,c,t}^{\text{charging}} \quad (2)$$

式中: $T_{i,c,t}^{\text{road}}$ 表示 t 时刻第 i 辆电动汽车前往充电站 c 路途上所需时间; $T_{i,c,t}^{\text{wait}}$ 和 $T_{i,c,t}^{\text{charging}}$ 分别表示 t 时刻第 i 辆电动汽车在充电站 c 所需的充电等待时间和充满电所需的充电时间。

根据美国联邦公路局(Bureau of Public Roads, BPR)的研究, 电动汽车在路段 a 的通行时间 $t_{a,t}$ 与车流量之间的关系^[19]可以表示为:

$$t_{a,t} = t_{a,t}^0 \left(1 + 0.15 \left(\frac{n_{a,t}}{c_a} \right)^4 \right), \quad \forall a \in E_a \quad (3)$$

式中: $n_{a,t}$ 表示 t 时刻路段 a 上的电动汽车数量; c_a 和 $t_{a,t}^0$ 分别表示路段 a 的容量上限和 t 时刻电动汽车自由通行时间; E_a 表示交通路段集合。由此可以得到第 i 辆电动汽车在通过路径 Z 前往充电站 c 所需时间 $T_{i,c,t}^{\text{road}}$, 即:

$$T_{i,c,t}^{\text{road}} = \sum_{a \in Z} t_{a,t} \quad (4)$$

第 i 辆电动汽车的充电等待时间为:

$$T_{i,c,t}^{\text{wait}} = \begin{cases} 0, & n_{c,t}^{\text{wait}} \leq 0 \\ T_{j,c,t}^{\text{re}} + T_{\text{avg}} \lfloor n_{c,t}^{\text{wait}} / C_c \rfloor, & n_{c,t}^{\text{wait}} > 0 \end{cases} \quad (5)$$

式中: $n_{c,t}^{\text{wait}}$ 表示电动汽车到达充电站 c 时站内等待充电的电动汽车数量; T_{avg} 表示电动汽车平均充电时长; C_c 表示充电站 c 中配置的充电桩数量; $\lfloor \cdot \rfloor$ 表示向下取整操作; $\lfloor n_{c,t}^{\text{wait}} / C_c \rfloor$ 表示电动汽车等待充电的轮数; $T_{j,c,t}^{\text{re}}$ 表示在第 c 个充电站内位于第 j 位的电动汽车剩余充电时间(从小往大排序), 其中 $j = \text{mod}(n_{c,t}^{\text{wait}}, C_c) + 1$, $\text{mod}(\cdot)$ 表示取余操作。式(5)表示在充电站 c 中电动汽车 i 的充电等待时间根据站内排在第 j 位的电动汽车剩余充电时间和充电轮数进行估计。

第 i 辆电动汽车的充电时间为:

$$T_{i,c,t}^{\text{charging}} = \frac{(1 - S_{i,c,t}) E_i^{\text{cap}}}{\eta P^{\text{charging}}} \quad (6)$$

式中: $S_{i,c,t}$ 表示 t 时刻第 i 辆电动汽车达到充电站 c 时的剩余电量 SOC; η 表示充电功率因素; E_i^{cap} 表示第 i 辆电动汽车电池容量; P^{charging} 表示电动汽车充电的额定功率。

1.2 充电成本模型

电动汽车 i 在 t 时刻前往充电站 c 进行充电时所产生的充电成本为:

$$F_{i,c,t} = \lambda_{c,t} (1 - S_{i,c,t}) E_i^{\text{cap}} / \eta \quad (7)$$

式中: $\lambda_{c,t}$ 表示 t 时刻充电站 c 的充电电价, 该值由第 3 章的配电网节点边际电价模型通过最优潮流模型得到。

2 基于图神经网络多智能体强化学习的电动汽车充电引导策略

本文的电动汽车充电引导策略主要对电动汽车充电路径进行优化, 如文献[20-21]。由于电动汽车充电行为涉及交通状况和充电电价等多重不确定性因素, 为了降低各种不确定性因素对预期目标的影响, 人们通常会注重短期的利益, 即主要考虑就近时间、就近距离等各种因素。因此, 本文假设邻近电动汽车其充电路径和充电目的地选择相比于距离较远的电动汽车对充电引导策略的影响更大。

为了更好地描述这种关系, 本文将研究区域中的电动汽车看作智能体, 并将其视为节点 $n \in N$; 将 2 辆在同一条道路上或者距离在某一半径内的电动汽车视为一条边 $e \in E$ 。通过图理论的方式构造图神经网络模型 $G_T = (N, E)$, 以此研究电动汽车用户间的相互影响, 制定更贴合实际和合理的电动汽车充电引导策略。

随着电动汽车的运行, 电动汽车间的连接关系也在不断发生变化, 即图网络 G_T 在不断变化, 为了处理这种图结构数据及其不断变化的连接关系, 提出了一种基于图神经网络的多智能体强化学习算法(graph multi-agent reinforcement learning, GMARL)对电力-交通融合网进行协同优化运行研究。其中, 图神经网络算法采用图注意力网络(graph attention network, GAT)^[22], 用于特征提取和处理这种不规则非欧式结构数据; 多智能体强化学习算法采用深度 Q 网络(deep Q network, DQN)算法^[23-24], 用于对电动汽车充电引导策略进行优化决策。

2.1 局部观测马尔科夫决策过程

本文将电动汽车充电引导问题构造为一个部分可观测马尔科夫决策过程(partially observable Markov decision process, POMDP)。POMDP 仅考虑有限辆电动汽车的信息, 在构建神经网络时能够有效降低神经网络规模, 从而加快算法的计算速度和收敛速度。通常, POMDP 问题被定义为一个六元

组 $\langle \mathcal{S}, X, \mathcal{A}, \mathcal{R}, \mathcal{T}, \gamma \rangle$, 包括: 环境状态 $s \in \mathcal{S}$, 电动汽车局部观测空间 $\mathbf{x}_i \in X_{i:N}$, 行为空间 $\mathbf{a}_i \in \mathcal{A}_{i:N}$, 立即奖励函数 $r_i(\mathbf{x}, \mathbf{a}) \in \mathcal{R}_{i:N}$, 状态转移函数 $\mathcal{T}(s, \mathbf{a}_{1:n}, \mathbf{w})$, 其中 \mathbf{w} 表示环境中的多种不确定性因素, $\gamma \in [0, 1]$ 表示折扣因子。为了更加及时地对电动汽车的充电引导策略进行优化决策, 时间间隔设置为 $\Delta t = 30$ s。

POMDP 过程如下所示。首先, t 时刻电动汽车智能体 i 通过观察环境得到观测状态 $\mathbf{x}_{i,t}$, 然后根据策略 $\pi(\mathbf{x}_{i,t})$ 选择一个行为动作 $\mathbf{a}_{i,t}$, 该动作作用于系统环境后, 根据状态转移函数 \mathcal{T} 移动到下一状态, 同时电动汽车智能体 i 获得立即奖励函数 $r_{i,t}$ 和一个新的局部观测 $\mathbf{x}_{i,t+1}$ 。以上过程不断持续, 每个电动汽车智能体 i 都会产生观测状态 $\mathbf{x}_{i,t}$ 、行为 $\mathbf{a}_{i,t}$ 和奖励函数 $r_{i,t}$: $\tau_i = \mathbf{x}_{i,1}, \mathbf{a}_{i,1}, r_{i,1}, \mathbf{x}_{i,2}, \dots, r_{i,T}$, 并且每个电动汽车智能体 i 的最大化累计期望回报 $R_i = \sum_{t=0}^T \gamma^t r_{i,t}$, 时间 $T=24$ h。下面对 POMDP 中观测状态 $\mathbf{x}_{i,t}$ 、行为 $\mathbf{a}_{i,t}$ 和奖励函数 $r_{i,t}$ 进行详细介绍。

2.1.1 观测状态

电动汽车智能体 i 在不同的位置根据其自身的信息 $\mathcal{E}_{i,t}^{\text{EV}}$ 、接收到的交通信息 $\mathcal{E}_{i,t}^{\text{RO}}$ 、近邻电动汽车信息 $\mathcal{E}_{i,t}^{\text{NE}}$ 以及从充电站接收到的信息 $\mathcal{E}_{i,t}^{\text{CS}}$ 制定合适的电动汽车充电引导策略。POMDP 的状态 $\mathbf{x}_{i,t}$ 可以表示为:

$$\mathbf{x}_{i,t} = [\mathcal{E}_{i,t}^{\text{EV}}, \mathcal{E}_{i,t}^{\text{RO}}, \mathcal{E}_{i,t}^{\text{NE}}, \mathcal{E}_{i,t}^{\text{CS}}] \quad (8)$$

$$\mathcal{E}_{i,t}^{\text{EV}} = [h_{i,t}^{\text{next}}, I_{i,t}^{\text{road}}, v_{i,t}, S_{i,t}] \quad (9)$$

$$\mathcal{E}_{i,t}^{\text{RO}} = [h_{i,j,t}^{\text{start}}, h_{i,j,t}^{\text{end}}, l_{i,j,t}^{\text{road}}, n_{i,j,t}^{\text{EV}}], j \in J \quad (10)$$

$$\mathcal{E}_{i,t}^{\text{NE}} = [h_{i,k,t}^{\text{next}}, I_{i,k,t}^{\text{road}}, v_{i,k,t}, S_{i,k,t}], k \in K \quad (11)$$

$$\mathcal{E}_{i,t}^{\text{CS}} = [\lambda_{c,t}, n_{c,t}^{\text{EV}}]_{c \in \text{CS}} \quad (12)$$

式中: 第 i 辆电动汽车状态 $\mathcal{E}_{i,t}^{\text{EV}}$ 包括电动汽车前往充电站充电时通过该路径中的下一节点 $h_{i,t}^{\text{next}}$ 、道路编号 $I_{i,t}^{\text{road}}$ 、电动汽车行驶速度 $v_{i,t}$ 和剩余电量 $S_{i,t}$, 其中电动汽车行驶速度 $v_{i,t}$ 和开始充电时剩余电量 $S_{i,t}$ 为随机变量。近邻交通道路信息状态 $\mathcal{E}_{i,t}^{\text{RO}}$ 包括与电动汽车 i 所在路段下一节点 $h_{i,t}^{\text{next}}$ 相连道路 j 的起始节点 $h_{i,j,t}^{\text{start}}$ 、末节点 $h_{i,j,t}^{\text{end}}$ 、道路长度 $l_{i,j,t}^{\text{road}}$ 以及道路上的电动车数量 $n_{i,j,t}^{\text{EV}}$, 其中道路上的电动车数量 $n_{i,j,t}^{\text{EV}}$ 代表交通流量信息, 为随机变量; J 表示与电动汽车 i 所在路段下一节点 $h_{i,t}^{\text{next}}$ 相连的节点集合。近邻电动汽车状态 $\mathcal{E}_{i,t}^{\text{NE}}$ 包括各近邻电动汽车 k 的状态, 如与第 i 辆电动汽车邻近的第 k 辆电动汽车下一节点 $h_{i,k,t}^{\text{next}}$, 其所在的道路编号 $I_{i,k,t}^{\text{road}}$ 、电动汽车行驶速度

$v_{i,k,t}$ 和剩余电量 $S_{i,k,t}$; K 表示近邻电动汽车集合。充电站信息 $\mathcal{E}_{i,t}^{\text{CS}}$ 包括各充电站的充电电价 $\lambda_{c,t}$ 和充电站中电动汽车数量 $n_{c,t}^{\text{EV}}$, 由于电动汽车充电行为具有一定的随机性, 同时加上风电和光伏出力的随机性, 从而使得充电站的充电电价 $\lambda_{c,t}$ 和充电站中电动汽车数量 $n_{c,t}^{\text{EV}}$ 具有一定的随机性, 因此这 2 个变量为随机变量。

综上所述, 本文涉及的不确定性因素主要包括道路交通流量、电动汽车行驶速度、电动汽车开始充电时剩余电量和充电电价等。

2.1.2 动作行为

电动汽车充电引导策略由一系列动作组成, 即电动汽车在当前状态下选择不同的道路前往充电站。不同路径其行驶时间、距离以及车速各异, 会影响电动汽车达到充电站时间, 进而造成电动汽车接入充电站的时间、充电等待时间等不同。本文电动汽车的行为 $\mathbf{a}_{i,t}$ 可以表示为:

$$\mathbf{a}_{i,t} = \begin{cases} 0, & h_{i,t}^{\text{cur}} \neq h_{i,t}^{\text{next}} \\ l_{n_c,j}, & h_{i,t}^{\text{cur}} = h_{i,t}^{\text{next}}, j \in M_{n_c} \end{cases} \quad (13)$$

式中: $h_{i,t}^{\text{cur}}$ 和 $h_{i,t}^{\text{next}}$ 分别表示 t 时刻电动汽车 i 当前位置和电动汽车下一节点位置; $l_{n_c,j}$ 表示电动汽车在节点 n_c 时选择第 j 条路径前往充电站; M_{n_c} 表示与节点 n_c 相连的路径集合。式(13)表示当电动汽车 i 没有到达下一个节点 $h_{i,t}^{\text{next}}$, 即 $h_{i,t}^{\text{cur}} \neq h_{i,t}^{\text{next}}$ 时, 电动汽车不进行决策优化, 即保持上一个行为动作; 只有当电动汽车达到下一个节点 $h_{i,t}^{\text{next}}$, 即 $h_{i,t}^{\text{cur}} = h_{i,t}^{\text{next}}$ 时, 电动汽车智能体 i 在 M_{n_c} 条路径中选择其中一条, 如第 j 条路径, 用 $l_{n_c,j}$ 表示。

2.1.3 奖励

当电动汽车未到达充电站, 即 $h_{i,t}^{\text{cur}} \neq h_{i,t}^{\text{ar}}$ 时, 其奖励函数设置为 0。一旦电动汽车到达充电站, 即 $h_{i,t}^{\text{cur}} = h_{i,t}^{\text{ar}}$ 时, 其奖励函数根据充电引导数学模型式(1)可以定义为:

$$r_{i,t} = w_i (T_{i,c,t}^{\text{road}} + T_{i,c,t}^{\text{wait}} + T_{i,c,t}^{\text{charging}}) + \lambda_{c,t} E_i^{\text{cap}} (1 - S_{i,c,t}) / \eta \quad (14)$$

2.2 图神经网络多智能体强化学习算法

2.2.1 图神经网络算法

通过以上分析可知, 为了更好地考虑电动汽车间的相互作用对充电行为的影响, 本文针对区域内的电动汽车及其交互关系构建了一种动态图模型结构, 并采用一种图注意力网络 GAT^[22] 对该非欧式结构的图模型数据进行特征提取, 提取的特征以便进一步被 DQN 利用制定最优的充电引导策略。与传统的图卷积神经网络平等对待其相邻节点不同, 图

注意力网络 GAT 可以为每个邻居分配不同的注意力权重,从而识别出对研究对象影响力较大的邻居,并赋予其一个较大的权重,以此改善模型的准确性。

假设各智能体状态 x_i 的特征维度均为 F , 则第 i 个智能体的特征向量可以表示为 $x_i \in \mathbf{R}^F$, \mathbf{R} 表示实数集。图网络模型 $G_T=(N, E)$ 中所有智能体的特性向量集合可以表示为 $\mathbf{X}=\{x_i \in \mathbf{R}^F | i=1,2,\dots,N\}$ 。图注意力网络的目的是输出智能体的新特征集合,即 $\mathbf{X}'=\{x'_i \in \mathbf{R}^{F'} | i=1,2,\dots,N\}$, 其特征维度为 F' 。

电动汽车智能体 j 对智能体 i 的重要性可以采用注意力机制公式进行计算, 即:

$$\alpha_{ij} = \frac{\exp(\delta(\mathbf{V}^T[\mathbf{W}x_i \| \mathbf{W}x_j]))}{\sum_{k \in \mathcal{N}_i} \exp(\delta(\mathbf{V}^T[\mathbf{W}x_i \| \mathbf{W}x_k]))} \quad (15)$$

式中: α_{ij} 表示智能体 j 对智能体 i 的注意力系数; $\delta(\cdot)$ 表示 LeakyReLU 激活函数; \mathcal{N}_i 表示智能体 i 的邻居集合; $\mathbf{W} \in \mathbf{R}^{F \times F}$ 表示在每个智能体上进行线性变换的权重矩阵; $\mathbf{V} \in \mathbf{R}^{2F'}$ 表示权重向量, 其将输入向量映射到实数 \mathbf{R} 空间; $\|$ 表示合并操作。

得到归一化注意力系数 α_{ij} 后, 计算其与相邻智能体 \mathcal{N}_i 对应特征的线性组合, 最终通过非线性激活函数 σ 得到每个智能体的输出特征向量为:

$$x'_i = \sigma \left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W}x_j \right) \quad (16)$$

为了稳定图注意力网络的学习过程, 对式(15)、(16)调用 K 组相互独立的注意力机制层, 并将 K 组注意力的输出进行平均操作, 可以得到最终注意力机制层的输出, 即:

$$x'_i = \sigma \left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in \mathcal{N}_i} \alpha_{ij}^k \mathbf{W}^k x_j \right) \quad (17)$$

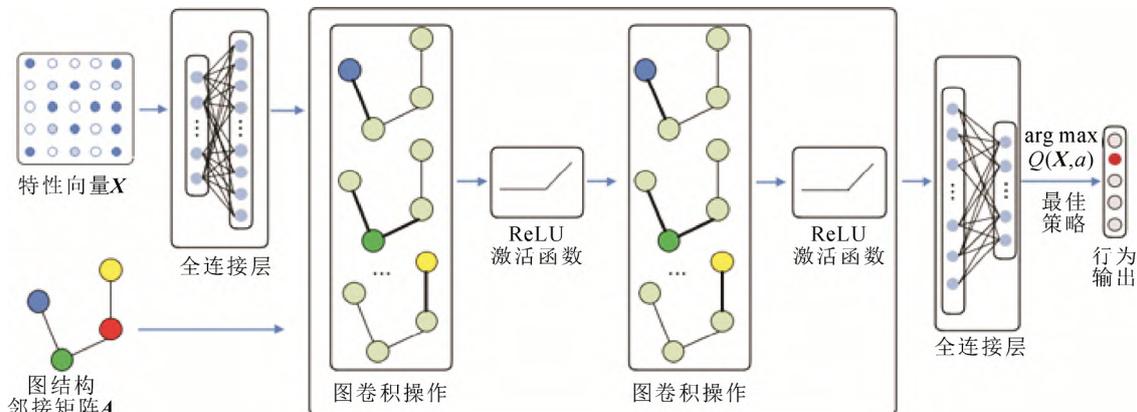


图1 图神经网络多智能体强化学习算法的神经网络结构设计

Fig.1 Neural network structure design of graph multi-agent reinforcement learning algorithm

2.2.2 神经网络结构设计

图神经网络多智能体强化学习算法的神经网络结构如图1所示。从图1中可以得知,神经网络的输入包括电动汽车相关特征向量 \mathbf{X} 以及由电动汽车关系构造的邻接矩阵 \mathbf{A} 。电动汽车相关特征向量 \mathbf{X} 先经过全连接层进行特性处理后, 与邻接矩阵 \mathbf{A} 一起输入到图卷积神经网络中进行处理。本文采用两层图卷积网络对输入的特性进行提取, 并全部采用 ReLU 非线性激活函数进行处理, 最后通过全连接层对动作进行输出, 其输出的动作行为根据动作值函数的最大值进行求取。图1中的 Q 表示状态-动作值函数。

2.2.3 图神经网络多智能体强化学习训练过程

强化学习算法通过智能体与环境不断交互, 根据环境的反馈智能体不断调整自身行为策略, 形成一种从状态到动作的映射, 使得智能体从环境中获得的平均累积奖励值 $R_{i,t}$ 最大。多个智能体在环境中协同合作, 通过环境的反馈调整行为策略, 其目的是使得智能体奖励之和最大, 即:

$$R_i = \sum_{t=1}^N (r_{i,1} + \gamma r_{i,2} + \dots + \gamma^T r_{i,T}) = \sum_{i=1}^N \sum_{t=0}^T r_{i,t} \quad (18)$$

本文提出的图神经网络多智能体强化学习算法结合了 GAT 和 DQN 的特性, 能够有效地解决区域内多智能体交互影响的问题。DQN 是强化学习中一种重要的方法, 它结合了深度学习强大的非线性处理能力和特征提取能力以及强化学习的序贯决策能力^[23-24]。深度 Q 网络算法通过神经网络对真实的状态-动作值函数 $Q(x,a;\theta)$ 进行拟合估计, 根据环境反馈奖励值采用梯度下降的方法不断调整神经网络参数 θ 以寻找满足最大回报的最优策略。图神经

网络多智能体强化学习损失函数 L 可以构造为:

$$\begin{cases} L(\theta_t) = E_{\mathbf{x}, a, r, \mathbf{x}' \sim D} ((y_t^{\text{GMARL}} - Q(\mathbf{x}, a; \theta_t))^2) \\ y_t^{\text{GMARL}} = r + \gamma \max_{a' \in \mathcal{A}} Q(\mathbf{x}', a'; \theta_t) \end{cases} \quad (19)$$

式中: \mathbf{x} 、 a 、 \mathbf{x}' 和 a' 分别为当前智能体所处状态、动作行为、下一步状态和下一步动作行为; γ ($0 \leq \gamma \leq 1$) 表示折扣因子, 其反映未来 Q 值对当前动作的影响; r 表示智能体在状态 \mathbf{x} 下执行动作 a 后得到的奖励信号; D 表示记忆单元; $E_{\mathbf{x}, a, r, \mathbf{x}' \sim D}(\cdot)$ 表示从记忆单元 D 中随机采样一批 \mathbf{x} 、 a 、 r 和 \mathbf{x}' , 并进行期望操作; y_t^{GMARL} 表示在目标网络 θ_t' 下的状态-动作值函数, 其参数每经过一定的步数根据当前网络 θ_t 进行更新。

当 GMARL 采用随机梯度下降方法对图神经网络参数进行更新时, 要求样本服从独立同分布条件, 以保证样本所含噪声相互抵消, 从而可以在一定程度上使得算法收敛得更快。实际上, 智能体与环境交互得到的训练样本并不满足独立同分布假设, 为了解决这一问题, GMARL 借鉴 DQN 的方式引入了经验回放机制: 将多次交互产生的累积经验样本 $(\mathbf{x}, a, r, \mathbf{x}')$ 存放在记忆单元 D 中, 在训练时, 从记忆单元 D 中随机抽取一定数量的样本进行训练, 以降低样本数据的相关性及训练数据的方差, 提高算法的收敛效率。GMARL 算法的训练过程如附录 A 图 A1 所示。

3 配电网最优潮流模型

电动汽车与配电网交互通过充电站完成, 电动汽车充电站连接在电力系统配电网某一个节点上。电力系统配电网模型可以描述为 $G_E = (E_N, E_L)$, 其中 E_N 和 E_L 分别表示配电网节点和线路集合。传统的配电网呈辐射状形式, 其意味着 1 个节点(除松弛节点外)仅有一个上级节点, 节点数和线路数相差 1^[25]。根据文献[26], 基于支路潮流模型的最优潮流(branch flow model-optimal power flow, BFM-OPF)和二阶锥松弛(second order conic relaxation, SOCR)转换, 并令 $\tilde{I}_{ij} = I_{ij}^2$ 以及 $\tilde{U}_{ij} = U_{ij}^2$ 可以得到:

$$\min \sum_{j \in E_N} (a_j (p_j^g)^2 + b_j p_j^g) + \rho P_{0j} \quad (20)$$

其约束条件为:

$$P_{ij} + p_j^g + p_j^{\text{wind}} - r_{ij} \tilde{I}_{ij} = \sum_{k \in \pi(j)} P_{jk} + p_j^d, \forall (i, j) \in E_L \quad (21)$$

$$Q_{ij} + q_j^g + q_j^{\text{wind}} - x_{ij} \tilde{I}_{ij} = \sum_{s \in \pi(j)} Q_{js} + q_j^d, \forall (i, j) \in E_L \quad (22)$$

$$\tilde{U}_j = \tilde{U}_i - 2(r_{ij} P_{ij} + x_{ij} Q_{ij}) + (z_{ij})^2 \tilde{I}_{ij}, \forall (i, j) \in E_L \quad (23)$$

$$\left\| \begin{matrix} 2P_{ij} \\ 2Q_{ij} \\ \tilde{I}_{ij} - \tilde{U}_j \end{matrix} \right\|_2 \leq \tilde{I}_{ij} + \tilde{U}_j, \forall (i, j) \in E_L \quad (24)$$

$$\tilde{I}_{ij} \leq \bar{I}_{ij}^2, \forall (i, j) \in E_L \quad (25)$$

$$\underline{U}_j \leq \tilde{U}_j \leq \bar{U}_j^2, \forall j \in E_N \quad (26)$$

$$0 \leq p_j^{\text{wind}} \leq \bar{p}_j^{\text{wind}}, \forall j \in E_N \quad (27)$$

$$q_j^{\text{wind}} = \tan(\arccos(\phi_j^{\text{wind}})) p_j^{\text{wind}}, \forall j \in E_N \quad (28)$$

$$p_j^d = p_j^{\text{dc}} + p_j^{\text{EV}}, \forall j \in E_N \quad (29)$$

式中: a_j 、 b_j 分别表示与节点 j 相连的发电机的二次煤耗系数和一次煤耗系数; ρ 表示向主网购电的电价; P_{0j} 表示传输到与松弛节点相连线路的有功功率^[7]; P_{ij} 、 Q_{ij} 分别表示支路 $i-j$ 的有功潮流和无功潮流; p_j^g 、 q_j^g 分别表示发电机注入到节点 j 的有功功率和无功功率; p_j^{wind} 、 q_j^{wind} 分别表示风机注入到节点 j 的有功功率和无功功率; $z_{ij} = r_{ij} + x_{ij}$ 表示支路 $i-j$ 的阻抗; I_{ij} 、 \underline{I}_{ij} 和 \bar{I}_{ij} 分别表示支路 $i-j$ 的电流、下限值和上限值; U_j 、 \underline{U}_j 和 \bar{U}_j 分别表示节点 j 电压、下限值和上限值; $\pi(j)$ 表示与节点 j 相连支路的节点的集合; ϕ_j^{wind} 表示连接节点 j 上的风机的功率因素; \bar{p}_j^{wind} 表示连接节点 j 上的风机出力最大值; p_j^d 表示配电网节点 j 的负荷, 其由 2 部分组成: 基础负荷 p_j^{dc} 和电动汽车充电负荷 p_j^{EV} 。

式(20)一式(29)构成了松弛后的静态最优潮流基本形式。文献[27-28]证明了在目标函数为凸函数和严格递增函数情况下对大部分配电网进行 SOCR 是严格准确的。在这样严格准确性证明下, 在 t 时刻下的节点边际电价 $\lambda_{c,t}$ (locational marginal price, LMP) 可以通过求解式(21)的对偶变量获得。求解对偶变量 $\lambda_{c,t}$ 的方法有 2 种, 其中一种方法是构造问题(式(29))的对偶问题, 然后通过该二阶锥规划对偶问题的最优解得到节点边际电价 $\lambda_{c,t}$ 。另一种方法是采用现有的商业求解器, 如 Mosek、Gurobi 等, 在求解原二阶锥规划问题(20)一式(29)的时候, 同时得到原、对偶变量最优解, 以此获得节点边际电价 $\lambda_{c,t}$ 。

基于图神经网络多智能体强化学习的 IETN 协同优化运行研究计算流程如附录 B 图 B1 所示。

4 仿真算例

4.1 仿真设置

本文在某市 108 节点区域交通网络和 IEEE 33 节点电力系统上进行仿真验证, 其拓扑结构图分别如附录 C 图 C1、C2 所示^[29-31]。配电网典型日负载率曲线如图 C3 所示, 风力发电机典型出力曲线如图 C4 所

示^[25,30], 其中风力发电的渗透率约为 20%。本文经过多次仿真试验对图神经网络多智能体强化学习参数进行优化, 其网络结构和超参数设置见附录 C 表 C1。

4.2 学习速率 α 和折扣因子 γ 选择

学习速率 α 和折扣因子 γ 会对 GMARL 学习效果产生较大影响, 可以通过比较奖励函数和损失函数值的大小以及收敛速度判断 GMARL 算法的好坏。不同学习速率 α 下 GMARL 算法的奖励函数和损失函数值如附录 C 图 C5。从图 C5(a)可以得知, 在学习速率 $\alpha = 1 \times 10^{-2}$ 、 1×10^{-3} 时, GMARL 算法获得的奖励函数较低, 约为学习速率 $\alpha = 1 \times 10^{-4}$ 时的 20%。当学习速率 $\alpha = 1 \times 10^{-4}$ 时, 其奖励函数值快速上升, 然后进入一个平缓上升阶段, 最后 GMARL 算法的奖励函数值趋于稳定。从图 C5(b)可以得知, 3 种学习速率下其损失函数值的变化趋势比较一致, 先经历一个小幅的上升, 然后逐步下降, 最后趋于稳定。结合 GMARL 算法的奖励函数和损失函数值, 最终选择学习速率 $\alpha = 1 \times 10^{-4}$ 。

不同折扣因子 γ 下的 GMARL 算法的奖励函数和损失函数值如附录 C 图 C6 所示。图 C6(a)的变化趋势和图 C5(a)基本一致, 但是当折扣因子 $\gamma = 0.99$ 时, 其奖励函数值低于其他折扣因子, 并且在后期经历了较大的波动, 这主要是由于 GMARL 算法在某些情况下采取了较差的行为, 影响了最优策略的获取, 产生了过拟合。从图 C6(b)可以得知, 折扣因子 $\gamma = 0.96$ 时可以获得较低的损失函数。综合奖励函数和损失函数值, 最终选择折扣因子 $\gamma = 0.96$ 。

4.3 仿真结果及比较分析

假设该区域中汽车保有量为 1 万辆, 电动汽车渗透率和需要充电汽车比例均为 10%, 可以得到需要充电的电动汽车数量 $N=100$ 。采用图神经网络多智能体强化学习算法和最短路径(shortest distance, SD)算法进行充电引导, 通过统计得到奖励函数、平均行驶速度、平均充电排队时间、平均充电行驶距离和峰谷差最大值比例如表 1 所示。

从表 1 中可以得知, 相比于 SD 算法, 所提出的 GMARL 算法能够有效地考虑电动汽车间相互影响和充电站的边际成本电价, 能够获得更优的策略以获得更大的奖励函数值, 降低电动汽车充电总费用。具体地, 在 GMARL 算法下电动汽车平均充电行驶速度为 1.097×10^{-2} km/s、充电平均的排队时间为 1.0403 min、峰谷差最大值百分比为 69.87%, 其优于 SD 算法下获得的各项指标, 其中平均充电排

队时间降低了 90%。尽管在 GMARL 算法下单次充电平均行驶距离为 7.9838 km, 高于 SD 算法的 5.3083 km, 但是其获得的单次奖励函数值高于 SD 算法, 表明大部分情况下电动汽车选择更通畅的充电路径而非最短路径前往充电电价更低的充电站进行充电, 以此降低电动汽车的充电总成本。

在 GMARL 和 SD 算法下, 电动汽车典型的充电负荷曲线分别如图 2 和图 3 所示, 其相应的节点边际电价如图 4 和图 5 所示。从图 2 中可以得知, 在 SD 算法下大部分电动汽车会选择离区域中心较近的 107 号充电站进行充电, 因此其充电负荷会远高于其他节点。这将会导致该节点的边际成本电价偏高(如图 4 所示), 进一步会导致电动汽车充电总成本提高, 降低电动汽车充电奖励函数值。相比于 SD 算法, 基于 GMARL 的 IETN 协同优化算法得到的各充电站负荷量基本一致, 表明不会出现大量的电动汽车涌向某一个充电站的情况, 此时不会出现充电电价过高导致增大充电总费用的情况。因此, 基于 GMARL 的 IETN 协同优化算法能够有效地处理电动汽车在运行过程中的各种不确定性因素, 如路网交通流量和电网边际成本电价, 能够有效地考虑电动汽车间相互影响以获得最优电动汽车充电路径, 从而降低电动汽车充电总成本。

表 1 不同算法结果比较

Table 1 Comparison of results of different algorithms

评价指标	图神经网络多智能体强化学习	最短路径算法
单次平均奖励函数	111.27	108.39
平均行驶速度/(km·s ⁻¹)	1.097×10^{-2}	1.072×10^{-2}
平均充电排队时间/min	1.0403	11.4621
平均充电行驶距离/km	7.9838	5.3083
峰谷差最大值/负荷值/%	69.87	70.86

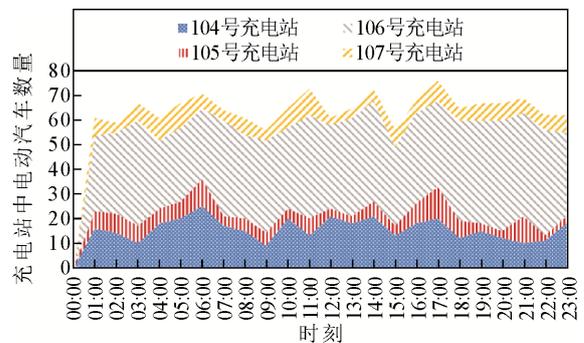


图 2 最短路径算法下的电动汽车负荷曲线

Fig.2 Charging load of electric vehicles under the shortest distance algorithm

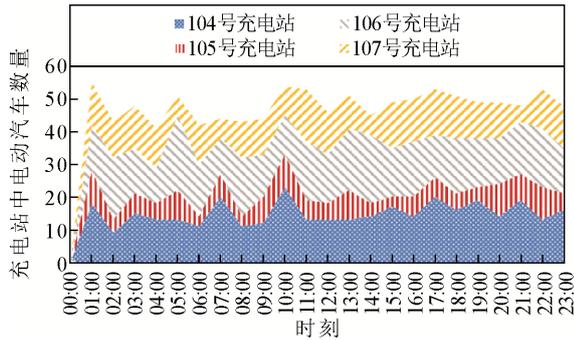


图3 图神经网络多智能体强化学习算法下的电动汽车负荷曲线

Fig.3 Charging load of electric vehicles under the GMARL algorithm

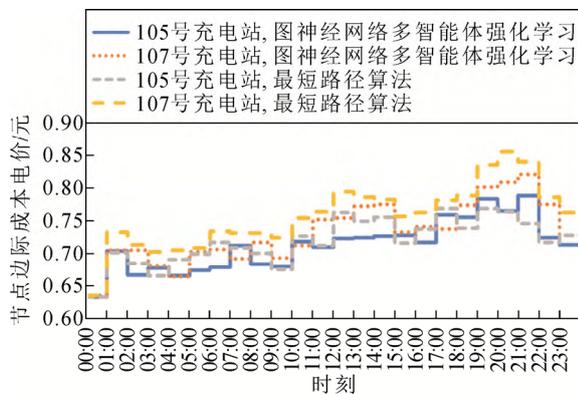


图4 不同算法下105号和107号充电站所在节点边际电价
Fig.4 Distribution local marginal price of nodes where No. 105 and No.107 charging stations are located under different algorithms

为了验证本文提出的 GMARL 的适应性能, 将充电电动汽车数量从 50 逐渐增加到 500, 并与最短路径 SD 算法比较各项指标: 电动汽车平均奖励值、平均行驶速度、平均行驶距离和充电排队时间指标, 如图 6 所示。从图 6 中可以看出, 随着充电电动汽车数量的增加, 电动汽车平均奖励值和平均行驶速度差距不断变大, 表明协同优化效果越来越明显。当电动汽车数量为 100 时, 其平均充电排队时间减少为最短路径算法的 1/10。由于充电站和充电桩规模有限, 导致电动汽车充电排队时间都会随着电动汽车充电数量的增多而增加, 充电排队时间方面的优势不断下降, 此时对奖励函数起主要作用的是路途行驶时间和电动汽车充电费用。

4.4 验证对比分析

1) 电动汽车间协作有效性验证

为了验证电动汽车间相互影响, 将输入的状态

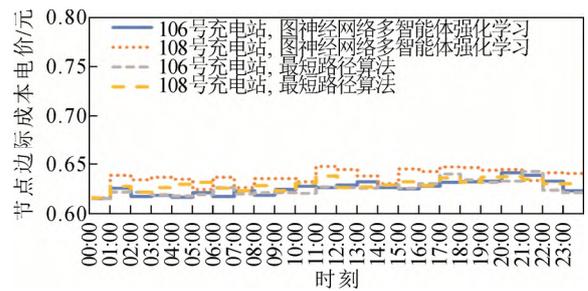
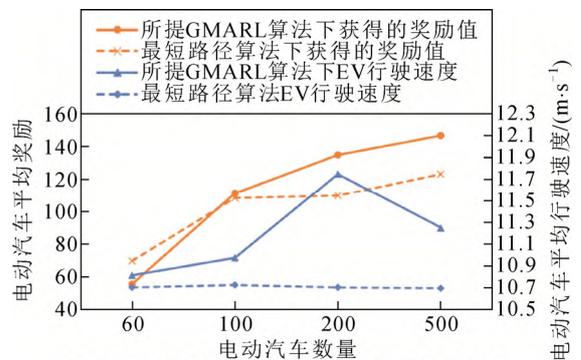
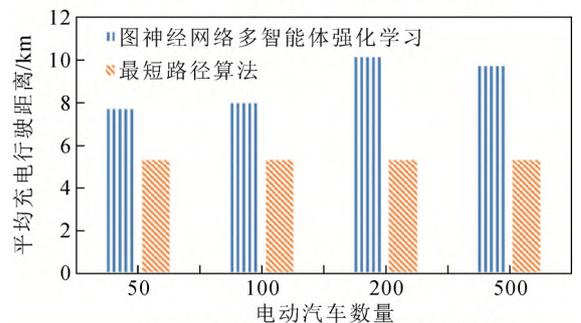


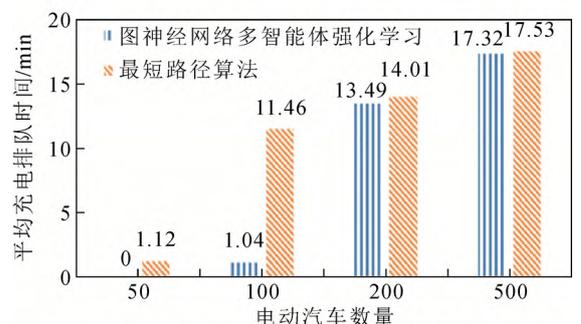
图5 不同算法下106号和108号充电站所在节点边际电价
Fig.5 Distribution local marginal price of nodes where No.106 and No.108 charging stations are located under different algorithms



(a) 电动汽车平均奖励值和平均行驶速度比较



(b) 电动汽车充电平均行驶距离比较



(c) 平均充电排队时间比较

图6 不同电动汽车数量情况下不同算法各项指标比较
Fig.6 Comparison of various indicators of different algorithms for different numbers of electric vehicles

$\mathbf{x}_{i,t}$ 进行相应修改: 删除交通道路信息 $\mathcal{E}_{i,t}^{RO}$ 和近邻电

电动汽车状态 $\mathcal{E}_{i,t}^{NE}$ ，仅保留当前电动汽车状态 $\mathcal{E}_{i,t}^{EV}$ 和充电站信息 $\mathcal{E}_{i,t}^{CS}$ ，即：

$$\mathbf{x}_{i,t} = [\mathcal{E}_{i,t}^{EV}, \mathcal{E}_{i,t}^{CS}] \quad (30)$$

在学习速率 α 和折扣因子 γ 等参数设置保持不变的前提下，不同状态下的 GMARL 算法的奖励函数和损失函数值如附录 C 图 C7 所示。从图 C7(a) 可以得知，在不考虑电动汽车相互影响状态时，GMARL 奖励函数值明显小于全状态下的奖励函数值。此外，在不考虑电动汽车相互影响状态时，在迭代次数为 2×10^5 时损失函数经历了较大的波动(如图 C7(b)所示)，进而导致电动汽车充电策略的奖励函数急剧下降，最终影响了最优策略的获取。

2) 图神经网络多智能体强化学习有效性验证

为了进一步验证 GMARL 算法的效果，本文将图卷积网络层(如图 1 所示)全部替换为全连接层，并且设置隐藏层的神经元个数为 256，以保证训练的参数量与 GMARL 算法参数量基本保持一致。GMARL 算法和基于全连接层深度强化学习算法的奖励函数和损失函数值如附录 C 图 C8 所示。从图 C8(a)可以得知，基于全连接层深度强化学习算法的损失函数值随着迭代次数的增加而不断增大，出现发散的情况，其原因主要在于全连接层神经网络很难有效处理电动汽车的关系，无法对输入的不规则非欧式结构数据进行有效信息提取。GMARL 算法奖励函数值明显高于基于全连接层深度强化学习算法的奖励函数值，并且其波动性更加小。仿真结果表明，GMARL 算法能够较好地对不规则非欧式结构数据进行特性提取，考虑电动汽车的交互，从而实现多智能体的协同优化运行控制。

5 结论

本文结合图神经网络强大的非欧式结构数据特征提取能力和强化学习算法序贯决策能力，提出了一种基于图神经网络多智能体强化学习算法，以解决考虑多种不确定性因素下的 IETN 协同优化运行问题。本方法在某区域 108 节点交通网络和 IEEE 33 节点电力系统上进行了仿真实验，证明了该方法的有效性。通过仿真对比分析，可以得出以下结论：

1) 提出的 IETN 协同优化方法相比于最短路径算法，其奖励函数和平均行驶速度得到了明显改善，其平均充电排队时间缩短了 90%，能够有效地降低电动汽车充电总成本。

2) 提出的 IETN 协同优化方法能够有效地引导

电动汽车充电，在一定程度上避免了大量的电动汽车涌向某一个充电站的情况，在提高充电站设备的利用率的同时，能够有效降低电动汽车充电电价，可以实现电力-交通系统的协同优化。

3) 提出的基于图神经网络多智能体强化学习的 IETN 协同优化方法具有较好的扩展性和适应性，能够适应不同环境下的 IETN 协同优化。

本文初步探索了将多智能体强化学习用于 IETN 的协同优化，比较了不同算法对 IETN 的优化效果及其扩展性、适应性。如何构建统一的算法来高效求解 IETN 协同随机优化模型将成为下一步工作的研究重点。

附录见本刊网络版(<http://hve.epri.sgcc.com.cn>)。

参考文献 References

- [1] 宋晓通, 吕倩楠, 孙 艺, 等. 基于电价引导的电动汽车与综合能源系统交互策略[J]. 高电压技术, 2021, 47(10): 3744-3754. SONG Xiaotong, LÜ Qiannan, SUN Yi, et al. Interactive strategy of electric vehicles and integrated energy system based on electricity price guidance[J]. High Voltage Engineering, 2021, 47(10): 3744-3754.
- [2] 胡 畔, 吴 斌, 陈红坤, 等. 计及时序特性的电动汽车快充充电站谐波分析[J]. 高电压技术, 2019, 45(11): 3645-3655. HU Pan, WU Bin, CHEN Hongkun, et al. Harmonic analysis of electrical vehicle fast-charging station considering sequence behavior of load[J]. High Voltage Engineering, 2019, 45(11): 3645-3655.
- [3] 王海鑫, 袁佳慧, 陈 哲, 等. 智慧城市车-站-网一体化运行关键技术研究综述及展望[J]. 电工技术学报, 2022, 37(1): 112-132. WANG Haixin, YUAN Jiahui, CHEN Zhe, et al. Review and prospect of key techniques for vehicle-station-network integrated operation in smart city[J]. Transactions of China Electrotechnical Society, 2022, 37(1): 112-132.
- [4] 侯 慧, 王逸凡, 吴细秀, 等. 长时间尺度下计及里程焦虑心理效应的电动汽车充放电调度策略[J]. 高电压技术, 2023, 49(1): 85-93. HOU Hui, WANG Yifan, WU Xixiu, et al. Electric vehicle charging and discharging scheduling strategy considering psychological effect of mileage anxiety on a long-time scale[J]. High Voltage Engineering, 2023, 49(1): 85-93.
- [5] 陈立兴. 电价引导机制下多场景电动汽车有序充电策略研究[D]. 南京: 东南大学, 2017. CHEN Lixing. Research on multi-scenario coordinated charging strategy of electric vehicle under electricity price guidance mechanism[D]. Nanjing, China: Southeast University, 2017.
- [6] 邵尹池, 穆云飞, 林佳颖, 等. “车-站-网”多元需求下的电动汽车快速充电引导策略[J]. 电力系统自动化, 2019, 43(18): 60-66, 101. SHAO Yinchi, MU Yunfei, LIN Jiaying, et al. Fast charging guidance strategy for multiple demands of electric vehicle, fast charging station and distribution network[J]. Automation of Electric Power Systems, 2019, 43(18): 60-66, 101.
- [7] WEI W, MEI S W, WU L, et al. Optimal traffic-power flow in urban electrified transportation networks[J]. IEEE Transactions on Smart Grid, 2017, 8(1): 84-95.
- [8] WEI W, WU L, WANG J H, et al. Network equilibrium of coupled transportation and power distribution systems[J]. IEEE Transactions on Smart Grid, 2018, 9(6): 6764-6779.
- [9] ZHOU Z, ZHANG X, GUO Q L, et al. Decomposition approach for the interdependency analysis of integrated power and transportation

- systems[J]. IET Generation, Transmission & Distribution, 2020, 3(6): 825-834.
- [10] 吕 思, 卫志农, 马骏超, 等. 基于多目标优化的电力-交通系统协同运行分析[J]. 电力系统自动化, 2022, 46(12): 98-106.
LÜ Si, WEI Zhinong, MA Junchao, et al. Analysis on coordinated power-transportation system operation based on multi-objective optimization[J]. Automation of Electric Power Systems, 2022, 46(12): 98-106.
- [11] SHI X Y, XU Y L, GUO Q L, et al. A distributed EV navigation strategy considering the interaction between power system and traffic network[J]. IEEE Transactions on Smart Grid, 2020, 11(4): 3545-3557.
- [12] 杨天宇, 郭庆来, 盛裕杰, 等. 系统互联视角下的城域电力-交通融合网络协同[J]. 电力系统自动化, 2020, 44(11): 1-9.
YANG Tianyu, GUO Qinglai, SHENG Yujie, et al. Coordination of urban integrated electric power and traffic network from perspective of system interconnection[J]. Automation of Electric Power Systems, 2020, 44(11): 1-9.
- [13] LU Y, LIANG Y C, DING Z H, et al. Deep reinforcement learning-based charging pricing for autonomous mobility-on-demand system[J]. IEEE Transactions on Smart Grid, 2022, 13(2): 1412-1426.
- [14] WANG S Y, BI S Z, ZHANG Y A. Reinforcement learning for real-time pricing and scheduling control in EV charging stations[J]. IEEE Transactions on Industrial Informatics, 2021, 17(2): 849-859.
- [15] QIAN T, SHAO C C, WANG X L, et al. Deep reinforcement learning for EV charging navigation by coordinating smart grid and intelligent transportation system[J]. IEEE Transactions on Smart Grid, 2020, 11(2): 1714-1723.
- [16] 袁红霞, 张 俊, 许沛东, 等. 基于图强化学习的电力交通耦合网络快速充电需求引导研究[J]. 电网技术, 2021, 45(3): 979-986.
YUAN Hongxia, ZHANG Jun, XU Peidong, et al. Fast charging demand guidance in coupled power-transportation networks based on graph reinforcement learning[J]. Power System Technology, 2021, 45(3): 979-986.
- [17] XU P D, ZHANG J, GAO T L, et al. Real-time fast charging station recommendation for electric vehicles in coupled power-transportation networks: a graph reinforcement learning method[J]. International Journal of Electrical Power & Energy Systems, 2022, 141: 108030.
- [18] 王岩庆, 王 骁, 丛若晨, 等. 考虑配电网运行安全的出行电动汽车充电引导策略[J]. 高电压技术, 2023, 49(5): 2131-2139.
WANG Yanqing, WANG Xiao, CONG Ruochen, et al. Charging guidance strategy of traveling electric vehicle considering the operation safety of distribution network[J]. High Voltage Engineering, 2023, 49(5): 2131-2139.
- [19] CASTILLO E, CALVIÑO A, SÁNCHEZ-CAMBRONERO S, et al. A multiclass user equilibrium model considering overtaking across classes[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(2): 928-942.
- [20] QIU D W, WANG Y, ZHANG T Q, et al. Hybrid multiagent reinforcement learning for electric vehicle resilience control towards a low-carbon transition[J]. IEEE Transactions on Industrial Informatics, 2022, 18(11): 8258-8269.
- [21] ALQAHTANI M, HU M Q. Dynamic energy scheduling and routing of multiple electric vehicles using deep reinforcement learning[J]. Energy, 2022, 244: 122626.
- [22] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks[EB/OL]. arXiv preprint arXiv: 1710.10903, 2017 [2018-02-04]. <https://arxiv.org/abs/1710.10903>.
- [23] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.
- [24] 漆 星, 郑常宝, 张 倩. 基于深度确信策略梯度的电动汽车异步电机参数标定方法[J]. 电工技术学报, 2020, 35(20): 4266-4277.
QI Xing, ZHENG Changbao, ZHANG Qian. An electric vehicle induction motor parameters calibration method based on deep deterministic policy gradient[J]. Transactions of China Electrotechnical Society, 2020, 35(20): 4266-4277.
- [25] 阮贺彬, 高红均, 刘俊勇, 等. 考虑 DG 无功支撑和开关重构的主动配电网分布鲁棒无功优化模型[J]. 中国电机工程学报, 2019, 39(3): 685-695.
RUAN Hebin, GAO Hongjun, LIU Junyong, et al. A distributionally robust reactive power optimization model for active distribution network considering reactive power support of DG and switch reconfiguration[J]. Proceedings of the CSEE, 2019, 39(3): 685-695.
- [26] GAN L W, LI N, TOPCU U, et al. Exact convex relaxation of optimal power flow in radial networks[J]. IEEE Transactions on Automatic Control, 2015, 60(1): 72-87.
- [27] LOW S H. Convex relaxation of optimal power flow-part I: formulations and equivalence[J]. IEEE Transactions on Control of Network Systems, 2014, 1(1): 15-27.
- [28] LOW S H. Convex relaxation of optimal power flow-part II: exactness[J]. IEEE Transactions on Control of Network Systems, 2014, 1(2): 177-189.
- [29] YANG T Y, GUO Q L, XU L, et al. Dynamic pricing for integrated energy-traffic systems from a cyber-physical-human perspective[J]. Renewable and Sustainable Energy Reviews, 2021, 136: 110419.
- [30] JIANG C X, JING Z X, CUI X R, et al. Multiple agents and reinforcement learning for modelling charging loads of electric taxis[J]. Applied Energy, 2018, 222: 158-168.
- [31] 国家统计局. 2021 年城镇单位就业人员平均工资数据[EB/OL]. [2022-05-20]. http://www.stats.gov.cn/xxgk/jd/sjjd2020/202205/t20220520_1857633.html.
National Bureau of Statistics. Average salary data of employed persons in urban units in 2021[EB/OL]. [2022-05-20]. http://www.stats.gov.cn/xxgk/jd/sjjd2020/202205/t20220520_1857633.html.



江昌旭

1990—, 男, 博士, 副教授, 硕导
主要从事人工智能技术在电力系统中的应用、电力-交通融合网方面的研究
E-mail: cxjiang@fzu.edu.cn

JIANG Changxu

Ph.D.

Associate professor



LIN Junjie

Ph.D.

Associate professor

Corresponding author

林俊杰(通信作者)

1992—, 男, 博士, 副教授, 硕导
主要从事电力系统分析与控制技术、状态估计等方面的研究
E-mail: linjunjie@fzu.edu.cn

收稿日期 2022-11-30 修回日期 2023-06-10 编辑 程子丰