36th International Electric Vehicle Symposium and Exhibition (EVS36)
Sacramento, California, USA, June 11-14, 2023

Optimal Vehicle Routing and Energy Scheduling in Wireless Electric Vehicle Energy Network via Reinforcement Learning Approach

Chaoran Si, K.T. Chau, Wei Liu

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China, Email: ktchau@eee.hku.hk

Executive Summary

Wireless electric vehicle energy network (WEVEN) uses electric vehicles (EVs) to regulate energy distribution over a large geographical region. EVs can charge and discharge wirelessly at electrified traffic lanes while moving, hence achieving energy transfer. From the EV owner's perspective, efficiently determining the optimal route and energy strategy in the profit-oriented market is an important issue. This paper proposes a reinforcement learning-based approach to solve the above combinatorial optimization problem in the WEVEN. Both the traveling time and energy trading benefits are optimized, while the state-of-charge constraints for EVs are well guaranteed. Simulation results for two cases, including a real traffic system, are given to verify the feasibility of the proposed approach.

Keywords: electric vehicle (EV), energy network, wireless charging, dynamic charging, reinforcement learning

1 Introduction

Transportation electrification is emerging as an important vision in many countries with rising concerns about global warming and fossil energy scarcity [1, 2]. Billions of dollars have recently been committed to funding the development of electric vehicles (EVs) and their auxiliary devices. The advances in EV wireless charging/discharging [3-5] and vehicular communication technologies [6, 7] enable the formation of the Energy Internet built upon the traffic network [8, 9], which is defined as the wireless electric vehicle energy network (WEVEN) [10]. In a WEVEN, EVs travel from the start point to the destination along the routes selected. The traffic lanes are equipped with bidirectional wireless power transfer systems and energy storage devices. According to the electricity price and state of charge (SoC) requirements, EVs can adopt energy management strategies while moving, namely charging, discharging, or idle [11], and thus transfer energy across the network.

Most EV routing or charging problems are formulated as mixed integer programming (MIP) models, which are generally nondeterministic polynomial time (NP)-hard and computationally difficult to solve to optimality. The routing and charging problem of an EV fleet was solved as the pre-trip decision-making in [12]. An efficient distributed algorithm is proposed in [13] to solve the joint optimization problem in EV networks. A method based on constraint logic programming and optimization using a graph-based shortest path algorithm was proposed in [14] to route EVs around mobile charging stations. [15] designed a EV charging system with the bus network and proposed a Route Scheduling Algorithm based on the approximation solution of the Restricted Shortest Path problem. A combined charging–driving navigation model for EVs traveling in the electrified highway network was proposed in [16], which was then solved by dynamic programming and the chronological search algorithm. [17] established a multi-objective optimization model to determine the route and charging strategies, and reduced the model complexity with the Big-M approach. Recently, some machine learning-based strategies, especially reinforcement learning (RL) algorithms, have been applied in this research area to enable fast decision-making [18]. Inspired by [18], [19] proposed an end-to-end deep RL framework to solve the EV routing problem with time windows (EVRPTW). A deep RL-based neural combinatorial optimization strategy was presented in [20] to solve the online vehicle problem transformed to a vehicle tour generation problem. [21] proposed a reinforcement model to solve the dynamic energy scheduling and routing of multiple EVs. A deep RL-based EV charging navigation method was applied in [22], aiming at minimizing the total travel time and the charging cost. As an improvement on [22], [23] used an online shortest-path-based approach to extract the low-dimensional features, which can be used as input to the deep RL algorithm.

This paper presents an action-value-based RL approach to solve the joint route selection and energy scheduling for EVs in the WEVEN. Unlike other RL algorithms, it requires no complete knowledge of the system. Not only the structure and functions of the WEVEN are well considered, but also the SoC requirements for EVs are always satisfied during the process.

The rest of the paper is organized as follows. Section 2 gives an overview of the WEVEN, including its definition, practicality, benefits and related technologies. Section 3 shows the problem statement and mathematical model formulated as mixed integer programming (MIP) optimization problem. Section 4 presents the reformulation of the MIP model and the proposed RL approach to solve it. The experiment results are illustrated in Section 5, and the conclusion is drawn in Section 6.

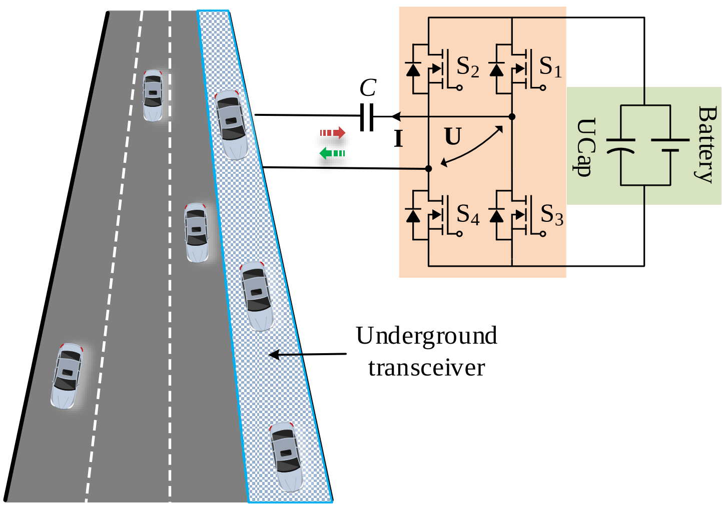


Figure 1: Location and configuration of a dynamic WER in WEVEN

2 Wireless Electric Vehicle Energy Network

WEVEN is built upon the road network and uses EVs to transport energy between multiple nodes along a selected path. A traffic lane can be electrified to become a dynamic wireless energy router (WER) [24], which is a bidirectional WPT system made up of a wireless power transceiver, a bidirectional converter and an energy storage system, as illustrated in Fig. 1. The energy storage system prefers to be a combination of battery and ultracapacitor [25], which can have long cycle life as well as high specific power. EVs can (dis)charge at dynamic WERs without stopping motion, thus saving more time for EVs to finish power exchange. The advantages of EV owners participating in wireless energy trading in the WEVEN include earning discharging income, reducing costs by optimizing charging [26], compressing automotive battery size [27], and alleviating range anxiety. WEVEN is also one of the puzzles of the Energy Internet, which can assist the power grid in peak shaving and power transmission. Moreover, WERs can store intermittent renewable generation, and such energy can be transported from remote areas to urban centers [28].

Information and communication technology (ICT) plays an important role in real-time energy trading in the WEVEN [29]. The fifth-generation communication technology can facilitate participating EVs to initiate their energy requests, communicate with the WERs for reservations and collect traffic information [30, 31]. Regional independent system operators may regulate the electricity price according to energy and traffic conditions and intervene in the profit-oriented activities of EVs. Combined with autonomous driving technology, participating EVs can quickly adjust the routes and (dis)charging strategies based on information such as prices and traffic conditions [32, 33], thereby improving energy conversion efficiency. Moreover, blockchain technology can record and balance the delivery and settlement processes, which can enhance security and privacy, improve transaction efficiency and enable real-time data sharing [31, 34-36].

3 Mathematical Optimization Model

3.1 Problem Description

A directed graph $\mathcal{G}=(\mathcal{N},\mathcal{A})$ can be used to describe the WEVEN, where $\mathcal{N}=\{1,2,...,|\mathcal{N}|-1\}$ is a set of $|\mathcal{N}|$ road junctions and $\mathcal{A}$ is a set of $|\mathcal{A}|$ directed arcs representing electrified roadways functioning as WERs. All EVs participating in wireless energy trading are equipped with hybrid energy storage of battery-ultracapacitor; therefore, they have very high specific power and can achieve fast wireless energy transfer. Suppose a participating EV will travel from node 0 to node $|\mathcal{N}|-1$, aiming to spend less travel time and get more trading profits. Due to the different renewable power generation and load conditions in the region where each WER is located, the electricity price in each region also changes accordingly. The EV can choose one of the available routes and determine the energy management strategy, namely charging, discharging, or idle, as it traverses each WER. During the whole driving process, the EV must ensure that its energy storage state of charge (SoC) is within an acceptable range, that is, overcharging and overdischarging are not allowed.

3.2 Mixed Integer Programming Model

According to the problem described above, a MIP model based on $\mathcal{G}$ can be established. The binary decision variables are introduced as follows:

$$x_{ij}=\begin{cases}1 & \text{if the EV travels through arc } (i,j),\\ 0 & \text{otherwise,}\end{cases}$$

$$xc_{ij}=\begin{cases}1 & \text{if the EV charges at arc } (i,j),\\ 0 & \text{otherwise,}\end{cases}$$

$$xd_{ij}=\begin{cases}1 & \text{if the EV discharges at arc } (i,j),\\ 0 & \text{otherwise.}\end{cases}$$

Then the optimization model can be formulated as:

$$\min \quad (1-\beta)\alpha Time+\beta Cost \tag{1}$$

$$Time=\sum_{(i,j)\in\mathcal{A}}\frac{x_{ij}l_{ij}}{v_{ij}} \tag{2}$$

$$Cost=\sum_{(i,j)\in\mathcal{A}}x_{ij}\left(xC_{ij}\cdot e_{ij}^{c}\cdot\rho_{ij}^{c}-xD_{ij}\cdot e_{ij}^{d}\cdot\rho_{ij}^{d}\right) \tag{3}$$

$$\text{s.t.}\quad \sum_{(i,j)\in\mathcal{A}}x_{ij}-\sum_{(j,k)\in\mathcal{A}}x_{jk}=\begin{cases}-1, & j=Start\\ 0, & \forall j\in\{\mathcal{N}\setminus Start,End\}\\ 1, & j=End\end{cases} \tag{4}$$

$$\sum_{(i,j)\in\mathcal{A}} x_{ij} \le 1, \quad \forall j \in \{\mathcal{N} \setminus Start\} \tag{5}$$

$$E_j = \sum_{(i,j)\in\mathcal{A}} \left(E_i + xC_{ij}e_{ij}^c - xD_{ij}e_{ij}^d - l_{ij}e_l\right)x_{ij}, \quad \forall j \in \{\mathcal{N} \setminus Start\} \tag{6}$$

$$E_{\min} \le E_j \le E_{\max}, \quad \forall j \in \mathcal{N} \tag{7}$$

$$xC_{ij} + xD_{ij} = 0, \quad \forall (i,j) \in \mathcal{A} \tag{8}$$

The objective function (1) minimizes both the total travel time defined in Eq. (2) and the total cost defined in (3), where α is the economic value of $Time$ and β is the weight to adjust the two sub-objectives. In (2), l_{ij} is the length of arc (i, j) and v_{ij} is the speed of EV traveling through arc (i, j). In Eq. (3), e_{ij}^c and ρ_{ij}^c are charging capacity and corresponding electricity price, e_{ij}^d and ρ_{ij}^d are discharging capacity and corresponding price. Constraints (4)-(5) ensure the route continuity and that the EV can only visit each node once. E_j is the carried energy when the EV arrives at node *j*, and the conservation of energy is defined in Eq. (6), where e_l is the power consumption per unit length. The upper and lower bounds for E_j are described in Eq. (7). Eq. (8) imposes constraints on the energy management strategy of the EV, meaning it cannot be charged and discharged at the same time when passing through a certain dynamic WER.

In order to solve the above MIP model, it is necessary to know the overall structure and all system parameters. Furthermore, it is a typical NP-hard problem with bilinear terms in both objective functions and constraints. Therefore, if the system scale is large, it can be very time-consuming to generate solutions.

4 Proposed Reinforcement Learning-Based Approach

4.1 EV Agent

The EV agent starts from node 0 and travels to node $|\mathcal{N}|-1$ along the WEVEN represented by a directed graph. It can reduce costs and even obtain benefits by selecting proper routes and energy strategies. The system state s_k at step *k* is a tuple of two variables $s_k = (p_k, SoC_k)$, where p_k is the vehicle location represented by the node number, and SoC_k is the SoC of EV battery. The action space $a_k = (a_{mk}, a_{ek})$ consists of two parts: one is mobility action a_{mk}, namely moving to one of the adjacent nodes along the directed arcs, and the other is energy management action $a_{ek} = \{xC_k, xD_k\}$, namely charging ($xC_k = 1$, $xD_k = 0$), discharging ($xC_k = 0$, $xD_k = 1$), or idle ($xC_k = 0$, $xD_k = 0$).

4.2 Environment Agent

Given the system state $s_k = (p_k, SoC_k)$ and the action $a_k = (a_{mk}, a_{ek})$ taken at step *k*, the next system state $s_{k+1} = (p_{k+1}, SoC_{k+1})$ can then be determined according to the following state transition rules:

$$p_{k+1} = a_{mk} \tag{9}$$

$$SoC_{k+1} = SoC_k + \left(-l_{ij} \cdot e_l + xC_k \cdot e_{ij}^c - xD_k \cdot e_{ij}^d\right) / E_{bat}, \quad \text{for } i = p_k \text{ and } j = a_{mk} \tag{10}$$

where the definitions of parameters e_l, e_{ij}^c and e_{ij}^d are the same as in the MIP model, and E_{bat} is the energy storage capacity of the EV.

The reward function can be defined as:

$$R(s_k, a_k) = -(1-\beta)\cdot\alpha\cdot\frac{l_{ij}}{v_{ij}} + \beta\left(xD_k \cdot e_{ij}^d \cdot \rho_{ij}^d - xC_k \cdot e_{ij}^c \cdot \rho_{ij}^c\right), \quad \text{for } i = p_k \text{ and } j = a_{mk} \tag{11}$$

The constraints for mobility actions in Eqs. (4)-(5) and energy management actions in Eq. (8) are involved in the definition of action space. The constraints on the energy carried energy by the EV are equivalent to those on the SoC. There are two methods proposed to limit the SoC for the EV agent to a certain range. One is to subtract a large value from the reward when the SoC exceeds the limit in a certain step; the other is to define the available action space based on the SoC value, which means only charging is permitted when the SoC value falls below the lower limit, and vice versa.

4.3 Algorithm

The action-value-based RL algorithm is adopted to find the best course of action for the above problem. The first-visit Monte Carlo (MC) method is used to evaluate the action-value function $Q(s,a)$. MC methods do not assume complete knowledge of the environment and require only experience sample sequences of states, actions, and rewards from actual or simulated interaction with an environment. Therefore, the environment agent defined in Section 4.2 needs only generate sample transitions. The pseudocode of the first MC method is shown in Algorithm 1.

Algorithm 1 First-visit MC method

Initialize:
$\pi \leftarrow$ an arbitrary policy
$Q(s,a) \leftarrow$ empty dictionary, for all $s \in S,\ a \in A(s)$
$Returns(s,a) \leftarrow$ empty dictionary, for all $s \in S,\ a \in A(s)$

while *True* **do**:
    generate an episode following policy π : $s_1,\ a_1,\ r_1,\ \cdots,\ s_K,\ a_K,\ r_K$
    $G \leftarrow 0$
    for $k = K,\ K-1,\ \cdots,\ 1$ **do**
        $G \leftarrow G + r_k$
        if $s_k \in \{s_1, \cdots, s_{k-1}\}$ **then**
            next k
        else
            append G to $Returns(s_k, a_k)$
            $Q(s_k, a_k) \leftarrow average(Returns(s_k, a_k))$
            $A^* \leftarrow \arg\max_a Q(s_k, a)$
            for all $a \in A(s_k)$ **do**
                $\varepsilon-$greedy policy: $\pi(a \mid s_k) \leftarrow \begin{cases} 1-\varepsilon+\varepsilon/|A(s_k)| & \text{if } a = A^* \\ \varepsilon/|A(s_k)| & \text{if } a \neq A^* \end{cases}$
            end for
        end if
    end for
end while

In order to achieve enough exploration for data collection, epsilon greedy is introduced to select the action from a given state. For any policy π, any ε-greedy policy π' with respect to value function q_π is guaranteed to be better than or equal to π, which can be proved by the policy improvement theorem in [37]. The equality can hold only when both π' and π are optimal among the policies.

5 Case Studies

5.1 Case 1

As shown in Fig. 2, a simple WEVEN with three traffic nodes and seven directed arcs is employed in Case 1. All the arcs are electrified as dynamic WERs enabling wireless energy trading. The parameters are shown in both Fig. 2 and Table 1. In this case, a per-unit (p.u.) approach is employed to quantify a number of different variables, and the (dis)charging capacities are assumed the same on each arc for convenience.

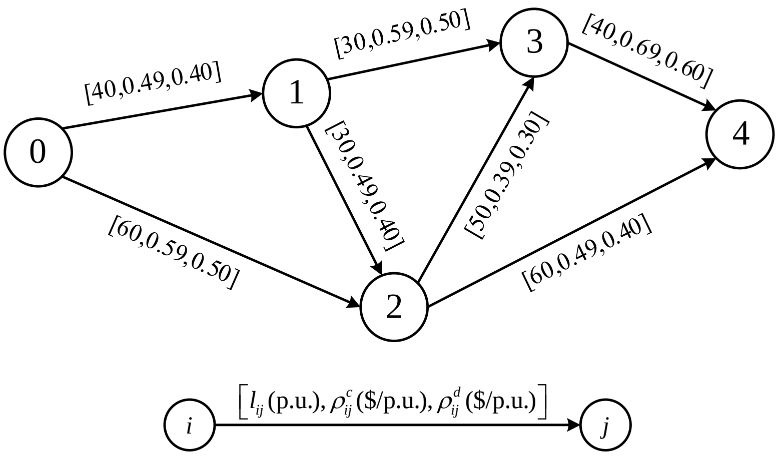


Figure 2: Dynamic WEVEN for Case 1

Table 1: Parameter Values for Case 1

Description	Parameter	Value	Unit
Charge consumed per kilometer	e_l	0.15	p.u.
Charging capacity on each arc	e_{ij}^c	10	p.u.
Discharging capacity on each arc	e_{ij}^d	10	p.u.
Initial battery charge	E_0	25	p.u.
Maximum battery charge	$E_{\max}$	50	p.u.
Minimum battery charge	$E_{\min}$	10	p.u.
Value of time	α	5	\$/p.u.

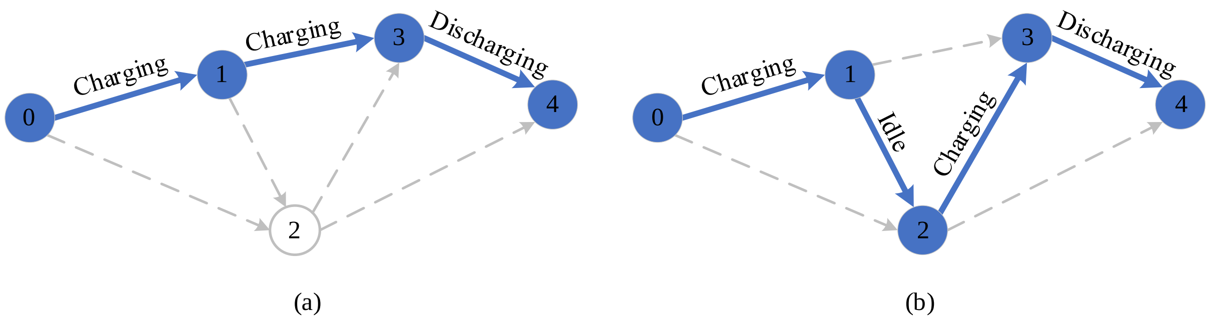


Figure 3: Optimal routes and energy strategies for Case 2. (a) $\beta = 0.2$ and 0.5. (b) $\beta = 0.8$.

In order to investigate the impact of the weighting in the objective function on the final result, three cases where $\beta = 0.2$, 0.5 and 0.8 are studied. The larger the β, the more the vehicle tends to earn more revenue, and the smaller the β, the more it tends to reduce travel time. The optimal routes and energy strategies are shown in Fig. 3. The total reward of the proposed RL method is compared with the objective value of the MIP model in Table 2, from which it can be seen that they are just opposite numbers, thus verifying the optimality of the proposed RL method.

Table 2: Optimal value comparison between RL and MIP methods

β	Total reward of RL	Objective value of MIP
0.2	-8.29	8.29
0.5	-6.98	6.98
0.8	-4.91	4.91

5.2 Case 2

Case 2 tests the proposed RL model using a real traffic system. The real road map of Santa Clara, California provided by Google is used in Fig. 4. Suppose there is one region with abundant renewable energy and low electricity prices and another region with heavy electric loads and high electricity prices in this WEVEN. An EV participating in wireless energy trading will travel from Node 0 to Node 19. The parameters of the EV are summarized in Table 3. The detailed travel time and (dis)charging prices on each road are given at Santa Clara.xlsx, in which three scenarios with different distributions of the two regions are provided. Note that the (dis)charging capacities e_{ij}^{c} and e_{ij}^{d} are determined by the product of (dis)charging power P and traveling time t_{ij} on each arc.

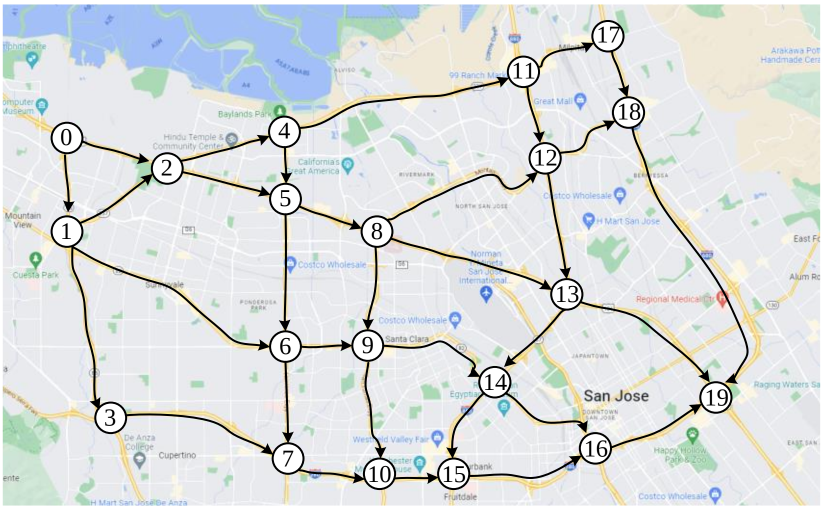


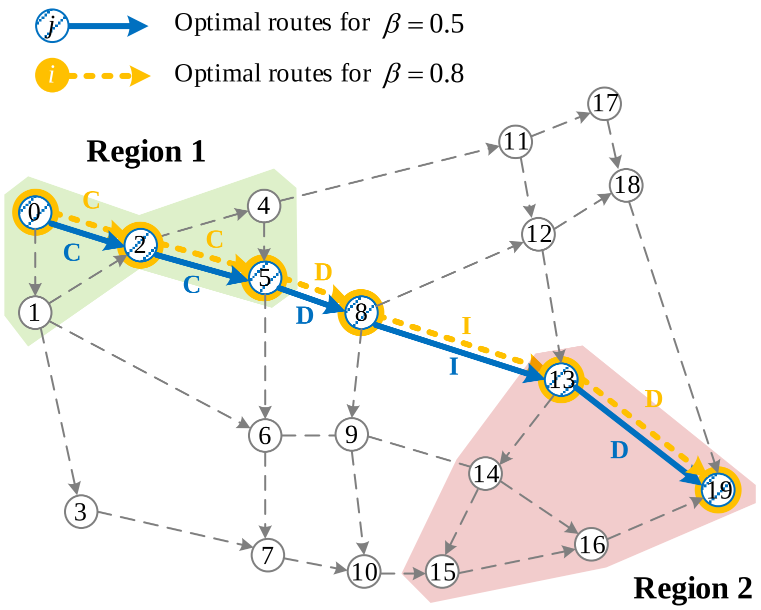
Figure 4: The real road map of Santa Clara, California

Table 3: Parameter Values for Case 1

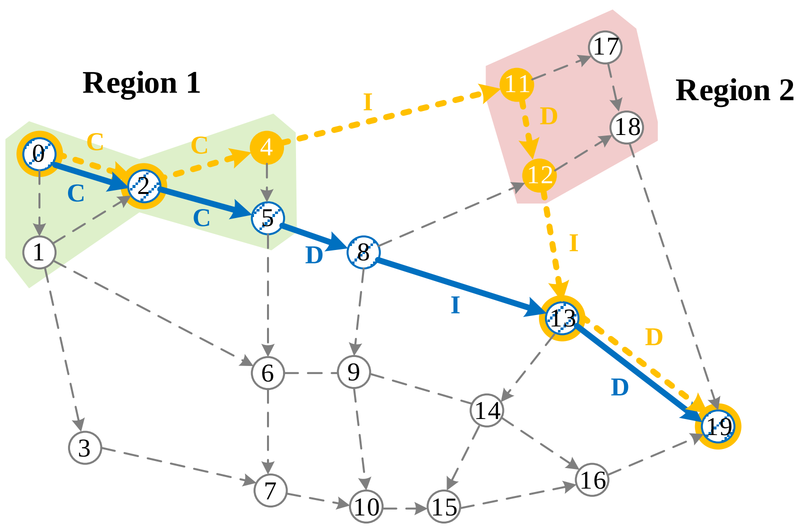
Description	Parameter	Value	Unit
Charge consumed per mile	e_l	0.25	kWh/mile
Wireless (dis)charging power	P	30	kW
Initial battery charge	SoC_0	30%	kWh
Maximum energy storage SoC	$SoC_{\max}$	100%	kWh

Minimum energy storage SoC	$SoC_{\min}$	20%	kWh
Value of time	α	0.2	\$/min

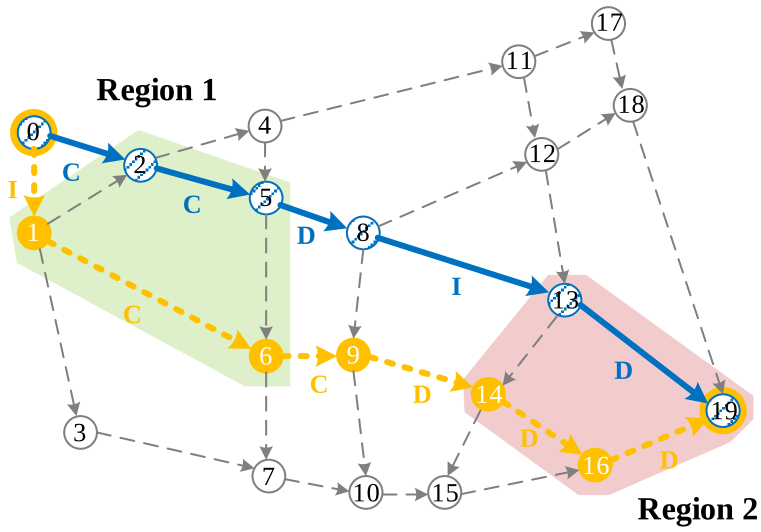
Fig. 5 shows the optimal routes and energy strategies for different electricity price region distributions, in which “D”, “I”, and “C” labeled on the selected routes represent discharging, idle, and charging, respectively. It can be seen that the EV agent tends to charge in Region 1 and discharge in Region 2. The EV agent has a greater possibility of (dis)charging in the desired regions when the start and end nodes are located nearby or even included in the two regions. Table 1 is a comparison of the optimization results of the three scenarios. When $\beta = 0.5$, the EV agent can get the maximum benefits in Scenario 1 and 3; when $\beta = 0.8$, the maximum benefits are also obtained in Scenario 3. The maximum electricity transferred from Region 1 to Region 2 can be achieved in Scenario 3 as well.



(a)



(b)



(c)

Figure 5: Optimal routes and energy strategies for Case 2 in three scenarios (Region 1 has abundant renewable energy and low electricity prices, and Region 2 has heavy electric loads and high electricity prices). (a) Scenario 1. (b) Scenario 2. (3) Scenario 3.

Table 1: Comparisons of the optimization results for Case 2

Scenario	β	Total reward of RL ($)	Charge capacity in Region 1 (kWh)	Discharge capacity in Region 2 (kWh)
1	0.5	-1.19	2	1
	0.8	0.02	2	1
2	0.5	-1.39	2	0
	0.8	-0.18	2	1
3	0.5	-1.20	1	2
	0.8	0.30	5.5	4.5

Although the action-value-based RL algorithm presented in this paper does not need the complete environment agent model, it does require billions of experience data for MC process to get the near-optimal strategy. In addition, since the real traffic system has quite a large action and state space to explore, the training can be very time-consuming. Other potential RL algorithms, such as deep Q-learning, can be investigated for better performance.

6 Conclusion

WEVEN is a promising application of the Energy Internet, and the mobility of EVs can help transport energy across a wide geographical area. This paper presents an action-value-based RL algorithm to find the optimal vehicle routing and energy scheduling strategy for EV owners in the profit-oriented WEVEN. Such an approach just requires sample sequences with prior experience rather than comprehensive knowledge of the environment. The SoC constraints for the EV are also well-guaranteed during the trip. Two cases, including a real traffic system, are examined to show that the proposed method can efficiently solve large-scale problems after offline training. Scenarios with different weights of the two objectives and price region distributions are simulated and analyzed. Future research directions can be uncertainty in the WEVEN, the coordination of multiple participating EVs, and the interaction between the WEVEN and the power grid.

Acknowledgments

This work was supported by a grant from the Hong Kong Research Grants Council, Hong Kong Special Administrative Region, China, under Project 17206222.

References

[1] W. Liu, T. Placke, and K. T. Chau, *Overview of batteries and battery management for electric vehicles,* Energy Reports, ISSN 2352-4847, 8, 4058-4084

[2] C. C. Chan and K. T. Chau, *Modern electric vehicle technology,* ISBN 0-19-850416-0, Oxford, Oxford University Press, 2001.

[3] P. Machura and Q. Li, *A critical review on wireless charging for electric vehicles,* Renewable and Sustainable Energy Reviews, ISSN 1364-0321, 104, 209-234

[4] C. Liu, K. T. Chau, and J. Zhong, *Development of a smart DC micro-grid for plug-in electric vehicle charging and discharging,* World Electric Vehicle Journal, ISSN 2032-6653, 4(4), 939-942

[5] C. Qiu, K.T. Chau, T. W. Ching, and C. Liu*, Overview of wireless charging technologies for electric vehicles,* Journal of Asian Electric Vehicles, ISSN 1883-6038, 12(1), 1679-1685

[6] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, *5G for vehicular communications,* IEEE Communications Magazine, ISSN 1558-1896, 56(1), 111-117

[7] S. Zeadally, M. A. Javed, and E. B. Hamida, *Vehicular communications for ITS: Standardization and challenges,* IEEE Communications Standards Magazine, ISSN 2471-2833, 4(1), 11-17

[8] S. M. S. Hussain, F. Nadeem, M. A. Aftab, I. Ali, and T. S. Ustun, *The emerging energy internet: Architecture, benefits, challenges, and future prospects,* Electronics, ISSN 2079-9292, 8(9),

[9] C. Liu, K. T. Chau, D. Wu, and S. Gao, *Opportunities and challenges of vehicle-to-home, vehicle-to-vehicle, and vehicle-to-grid technologies,* Proceedings of the IEEE, ISSN 1558-2256, 101(11), 2409-2427

[10] A. Y. S. Lam, K. C. Leung, and V. O. K. Li, *Vehicular energy network,* IEEE Transactions on Transportation Electrification, ISSN 2332-7782, 3(2), 392-404

[11] B. Lin, B. Ghaddar, and J. Nathwani, *Electric vehicle routing with charging/discharging under time-variant electricity prices,* Transportation Research Part C: Emerging Technologies, ISSN 0968-090X, 130, 103285

[12] T. Chen, B. Zhang, H. Pourbabak, A. Kavousi-Fard, and W. Su, *Optimal routing and charging of an electric vehicle fleet for high-efficiency dynamic transit systems,* IEEE Transactions on Smart Grid, ISSN 1949-3053, 9(4), 3563-3572

[13] X. Tang, S. Bi, and Y.-J. A. Zhang, *Distributed routing and charging scheduling optimization for internet of electric vehicles,* IEEE Internet of Things Journal, ISSN 2327-4662, 6(1), 136-148

[14] D. Kosmanos *et al.*, *Route optimization of electric vehicles based on dynamic wireless charging,* IEEE Access, ISSN 2169-3536, 6, 42551-42565

[15] Y. Jin, J. Xu, S. Wu, L. Xu, and D. Yang, *Enabling the wireless charging via bus network: Route scheduling for electric vehicles,* IEEE Transactions on Intelligent Transportation Systems, ISSN 1558-0016, 22(3), 1827-1839

[16] S. Lv, Z. Wei, G. Sun, S. Chen, and H. Zang, *EV charging-driving navigation in electrified highway network,* in 2020 IEEE Power & Energy Society General Meeting (PESGM), 2-6 Aug. 2020, 1-5

[17] C. Li, T. Ding, X. Liu, and C. Huang, *An electric vehicle routing optimization model with hybrid plug-in and wireless charging systems,* IEEE Access, ISSN 2169-3536, 6, 27569-27578

[18] M. Nazari, A. Oroojlooy, L. Snyder, and M. Takác, *Reinforcement learning for solving the vehicle routing problem,* Advances in Neural Information Processing Systems, 31(2018)

[19] B. Lin, B. Ghaddar, and J. Nathwani, *Deep reinforcement learning for the electric vehicle routing problem with time windows,* IEEE Transactions on Intelligent Transportation Systems, ISSN 1558-0016, 23(8), 11528-11538

[20] J. J. Q. Yu, W. Yu, and J. Gu, *Online vehicle routing with neural combinatorial optimization and deep reinforcement learning,* IEEE Transactions on Intelligent Transportation Systems, ISSN 1558-0016, 20(10), 3806-3817

[21] M. Alqahtani and M. Hu, *Dynamic energy scheduling and routing of multiple electric vehicles using deep reinforcement learning,* Energy, ISSN 0360-5442, 244, 122626

[22] T. Qian, C. Shao, X. Wang, and M. Shahidehpour, *Deep reinforcement learning for EV charging navigation by coordinating smart grid and intelligent transportation system,* IEEE Transactions on Smart Grid, ISSN 1949-3061, 11(2), 1714-1723

[23] J. Jin and Y. Xu, *Shortest-path-based deep reinforcement learning for EV charging routing under stochastic traffic condition and electricity prices,* IEEE Internet of Things Journal, ISSN 2327-4662, 9(22), 22571-22581

[24] P. Yi, T. Zhu, B. Jiang, R. Jin, and B. Wang, *Deploying energy routers in an energy internet based on electric vehicles,* IEEE Transactions on Vehicular Technology, ISSN 1939-9359, 65(6), 4714-4725

[25] W. Liu, K. T. Chau, C. C. T. Chow, and C. H. T. Lee, *Wireless energy trading in traffic internet*, IEEE Transactions on Power Electronics, ISSN 1941-0107, 37(4), 4831-4841

[26] P. Liu *et al.*, *Joint route selection and charging discharging scheduling of EVs in V2G energy network*, IEEE Transactions on Vehicular Technology, ISSN 0018-9545, 69(10), 10630-10641

[27] C. C. T. Chow, A. Y. S. Lam, W. Liu, and K. T. Chau, *Multisource–multidestination optimal energy routing in static and time-varying vehicular energy network,* IEEE Internet of Things Journal, ISSN 2327-4662, 9(24), 25487-25505

[28] K. Mahmud, G. E. Town, S. Morsalin, and M. J. Hossain, *Integration of electric vehicles and management in the internet of energy,* Renewable and Sustainable Energy Reviews, ISSN 1364-0321, 82, 4179-4203

[29] A. Haldar and N. Sethi, *Environmental effects of Information and Communication Technology - Exploring the roles of renewable energy, innovation, trade and financial development,* Renewable and Sustainable Energy Reviews, ISSN 1364-0321, 153, 111754

[30] X. Cheng, R. Zhang, S. Chen, J. Li, L. Yang, and H. Zhang, *5G enabled vehicular communications and networking*, China Communications, ISSN 1673-5447, 15(7), iii-vi

[31] P. Bhattacharya, S. Tanwar, U. Bodkhe, A. Kumar, and N. Kumar, *EVBlocks: A blockchain-based secure energy trading scheme for electric vehicles underlying 5G-V2X ecosystems,* Wireless Personal Communications, ISSN 1572-834X, 127(3), 1943-1983

[32] J. J. Q. Yu and A. Y. S. Lam, *Autonomous vehicle logistic system: Joint routing and charging strategy,* IEEE Transactions on Intelligent Transportation Systems, ISSN 1558-0016, 19(7), 2175-2187

[33] Z. Liao, M. Taiebat, and M. Xu, *Shared autonomous electric vehicle fleets with vehicle-to-grid capability: Economic viability and environmental co-benefits,* Applied Energy, ISSN 0306-2619, 302, 117500

[34] Y. Wang, Z. Su, and N. Zhang, *BSIS: Blockchain-based secure incentive scheme for energy delivery in vehicular energy network,* IEEE Transactions on Industrial Informatics, ISSN 1941-0050, 15(6), 3620-3631

[35] Y. Wang, H. T. Luan, Z. Su, N. Zhang, and A. Benslimane, *A secure and efficient wireless charging scheme for electric vehicles in vehicular energy networks,* IEEE Transactions on Vehicular Technology, ISSN 1939-9359, 71(2), 1491-1508

[36] A. Miglani, N. Kumar, V. Chamola, and S. Zeadally, *Blockchain for internet of energy management: Review, solutions, and challenges,* Computer Communications, ISSN 0140-3664, 151, 395-418

[37] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction,* ISBN 978-0-262-03924-6, Cambridge, MA, MIT press, 2018

Presenter Biography

Chaoran Si received the B.Eng. and B.M. degrees from Tianjin University, Tianjin, China, in 2018, and the M.Eng. degree from Zhejiang University, Hangzhou, China, in 2021. She is currently pursuing a Ph.D. degree at the Department of Electrical and Electronic Engineering at the University of Hong Kong. Her research interests include wireless vehicular energy network, electric vehicle technologies, and smart grid.