

A DRL-Based Partial Charging Scheme of Multiple Mobile Chargers for Maximizing Survival Rate and Energy Usage Efficiency

Tingwei Li, Shuangjuan Li*

College of Mathematics and Informatics, South China Agricultural University, China
 litingwei@stu.scau.edu.cn, lishj2016@scau.edu.cn

Abstract—Wireless Rechargeable Sensor Network (WRSN) is a new paradigm that prolongs the lifetime of Wireless Sensor Network (WSN). To improve the survival rate of sensor nodes and energy usage efficiency, this paper studied how to schedule charging paths of multiple mobile chargers and allocate charging time simultaneously. We propose an on-demand partial charging scheme for multiple mobile chargers. First, to balance the charging workload of multiple mobile chargers, we proposed a new clustering algorithm to divide the sensor nodes into several clusters. Then, we proposed an algorithm of determining the target charging sensor node. Finally, based on deep reinforcement learning technique, the charging duration calculation strategy is designed to automatically allocate charging duration for the sensors. Extensive simulations show that, compared with baseline algorithms, our scheme can increase the survival rates of sensor nodes by 3.62%~11.95% and increase the energy usage efficiency by 1.57%~6.15% in different network scale.

Index Terms—partial charging scheme, multiple mobile chargers, deep reinforcement learning, wireless rechargeable sensor network

I. INTRODUCTION

The wireless sensor network (WSN) consists of sensor nodes for sensing, data collecting, and processing [1]. Decades of continuous research and development have led to wireless sensor networks being widely used in agriculture, medical, military, and other fields [2]. However, the energy in wireless sensors is limited, which constrains the further development of wireless sensor networks [3]. The recent breakthrough in wireless power transfer (WPT) technology provides a promising way to empower sensors wirelessly [4]. With the help of WPT, in a wireless rechargeable sensor network, mobile chargers (MCs) can move and wirelessly charge the sensors by equipped with high-capacity batteries and transmitting antennas. Besides, the sensors are equipped with receivers to capture energy transmitted by the MC. The WRSN provides controllable and reliable energy replenishment, facilitating sustainable sensor network operations [5].

In a large-scale sensor network, a single mobile charger may not be able to satisfy all charging requests in time, as there exists a large number of sensors to be charged to avoid their energy depletion [6]. To ensure that sensor nodes receive timely energy replenishment, multiple mobile chargers are required to deploy in the sensor network. Besides, the

charging process varies from fast to slow charging according to the charging characteristics of batteries [7]. The full charging method will lower charging efficiency in the late charging period. In this paper, we designed an on-demand partial charging scheme for multiple chargers in WRSN. The main contributions of this work are as follows:

- The charging scheduling problem is modeled as a multiple-objective joint optimization problem, which aims at simultaneously maximizing survival rate and energy usage efficiency.
- The proposed algorithm based on deep reinforcement learning will train a deep neural network that schedules mobile chargers automatically allocate an appropriate charging duration to the sensors.
- We conducted the experiment showing the superiority of our algorithm compared to existing on-demand methods in several critical performance metrics.

The remainder of the paper is constructed as follows: Section II reviews related work. Section III introduces the system model and definition of the problem. Section IV demonstrates the proposed charging algorithm in detail. Section V presents the results of the simulations. Section VI concludes the paper and points out future work.

II. RELATED WORK

A. Full Charging Schemes

The Nearest Job Next Preemption (NJNP) charging scheme, as proposed in [8], selects the closest node from the mobile charger to be charged next. However, MC would give up charging when receiving a new charging request from a closer sensor. Therefore, frequent preemption makes NJNP deficient. In [9], Lin et al. proposed an on-demand TSCA charging scheme for WRSNs. The scheme aims to minimize the number of dead nodes while maximizing the charging efficiency so as to prolong the network lifetime. Compared with NJNP, TSCA achieved promising performance in terms of charging throughput and charging efficiency. Further, Lin et al. [10] investigated the issue of multiple mobile chargers' online collaborative charging schedules. They divided the sensor network into several subdomains and designated an MC in each subdomain. Then, they proposed an mTS charging scheme that combines temporal requirements and spatial features. Dong et al. [11]

*Corresponding Author: Shuangjuan Li

combined the temporal and spatial requirements and event features into a united charging priority metric that is used to sort instant charging requests for MC to serve. Abhinav et al. [12] designed a fuzzy inference system which determines the charging schedules of the nodes by wisely integrating multiple attributes.

B. Partial Charging Schemes

Aiming to maximize the sum of sensor survival times and minimize the traveling distance of the mobile charger, Xu et al. [13] proposed a partial charging scheme to increase sensor survival opportunities. However, in their work, sensors are always replenished by a fixed unit power each time. Based on the characteristics of the charging curve on lithium batteries, Guo et al. [14] proposed a partial charging scheme where sensor nodes are charged to their 80% energy level. To achieve the sum of charging rewards collected from all charged sensors by the mobile charger per tour maximized, Liang et al. [15] determined the energy level for partial charging using the residual energy of the sensor nodes. Wang et al. [16] prioritize sensors based on their contributions to the charging tasks and always select core nodes to serve so as to minimize the overall node dead time and prolong the network lifetime. Noting that through partial charge, more sensors can be served before their energy exhaustion and charging delay can be shortened, Lin et al. [17] developed an MPF charging scheme suitable for large-scale sensor networks.

III. MODEL AND PROBLEM STATEMENT

A. Network Model

A wireless rechargeable sensor network consists of a base station (BS), mobile chargers, and wireless sensor nodes. As shown in Figure 1, there are m wireless mobile chargers (denoted as $W = \{w_1, w_2, \dots, w_m\}$) and n stationary sensor nodes (denoted as $S = \{s_1, s_2, \dots, s_n\}$) deployed in the $X \times Y$ two dimensions area. The sensor $s_i \in S$ is powered by a rechargeable battery with energy b .

The sensors consume battery energy for sensing and transmitting data. We assume that ec_i and re_i denote the energy consumption rate of s_i and its residual energy. So, the residual lifetime of s_i can be calculated by $rl_i = re_i/ec_i$. When rl_i falls below a threshold Θ , sensor s_i will send a *charging request* $req(s_i) = (p_i, re_i, ec_i)$ to the BS, where $p_i = (x_i, y_i)$ denotes the position of s_i . After receiving a charging request, BS inserts the request into MC's *charging queue* $\Delta = \{req_1, req_2, \dots\}$. Subsequently, the MC selects a request $req(s_j)$ from its charging queue for responding. It moves to s_j and charges the sensor for t_i seconds. This process is referred to a *charging task*, denoted by $k_i = (s_j, t_i)$. The energy required for MC to perform the charging task k_i and return to BS is calculated as:

$$e_{k_i} = (d(s_{i-1}, s_i) + d(s_i, BS))ec_{trl} + ec_{chg}t_i, \quad (1)$$

where $d(s_{i-1}, s_i)$ represents the distance between the sensors in the $i-1$ th and i th charging task. ec_{chg} and ec_{trl} respectively represents the energy consumption rate for charging

and traveling spent by w_i . Before MC performs the charging task k_i , MC will calculate its remaining energy re_w . If re_w is lower than e_{k_i} , it will not executing k_i and return to the BS for replenishing energy. We assume that the MC can immediately replenish energy by replacing the battery, so the time for the MC to obtain energy from the BS is ignored.

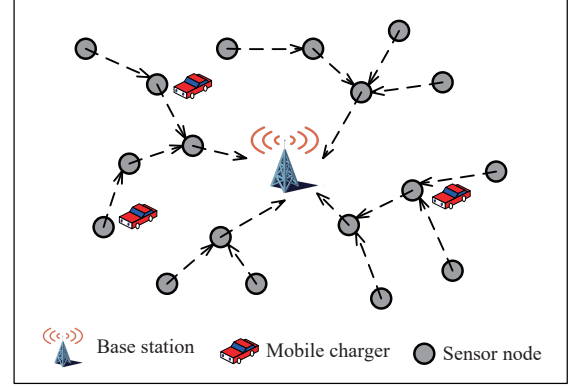


Figure 1. Network model.

B. Charging Model

During the process of charging the battery, the energy gained by the sensors is not at a stable rate but instead follows a pattern from fast to slow. The charging process where the sensor s_i gets energy from the MC can be calculated as [7]:

$$\int_0^t \frac{1}{\tau} (b - re_i) e^{-\frac{z}{\tau}} dz = (b - re_i) [1 - e^{-\frac{t}{\tau}}]_0^t, \quad (2)$$

where $b - re_i$ represents the amount of energy required to charge a battery fully, t is the charging duration, and τ is the time constant, which is a technology-dependent parameter related to the battery. The charging process reveals that the charging efficiency has a marginal effect. As the charging process continues, the charging speed will gradually slow down.

C. Problem Formulation

In a wireless sensor network, if sensors run out of energy and is unable to sense or transmit data, it will seriously impair system performance. Therefore, the survival rate of sensor nodes is a crucial performance metric of WRSN [10]. To measure the survival rate, the metric δ is defined as:

$$\delta = \frac{1}{t} \int_0^t \frac{|S_{alive}|}{|S|} dt, \quad (3)$$

where $|S_{alive}|$ is the number of survival sensor nodes at time t , $|S|$ is the total number of sensor nodes deployed in the network.

The energy cost of MCs consists of two parts: the energy consumed by MCs moving between charging positions and the energy consumed for charging the sensors. To ensure that energy-limited MCs transmit more energy to sensors, an effective charging scheme should consider how to make full

use of the battery energy of MCs. To measure the energy usage efficiency, the metric η is defined as:

$$\eta = \frac{E_{node}}{E_{chg} + E_{trl}} = \frac{\sum_{i=1}^{|S|} \sum_{k_j \in K_i} (b - re_i)(1 - e^{-t_j/\tau})}{\sum_{i=1}^{|W|} t_{chg}^i \cdot ec_{chg} + \sum_{i=1}^{|W|} vt_{trl}^i \cdot ec_{trl}}, \quad (4)$$

where E_{node} denotes the energy obtained by sensors and E_{chg} , E_{trl} denotes the energy consumed by MCs' charging and traveling. K_i is a set that contains all charging task $k_j(s_i, t_j)$ which is replenishing energy for s_i . t_{chg}^i and t_{trl}^i respectively represents the time for charging and moving spent by w_i .

In this work, we focus on studying the Survival rate and energy usage Efficiency MAXimation (SEMA) problem, which is given as follows:

Given a WRSN with multiple MCs, our objective is to schedule charging tasks for MCs (i.e. determine which sensor should be charged and corresponding charging duration) so as to maximize survival rate and energy usage efficiency.

This problem can be formulated as:

$$\max \delta, \quad (5)$$

$$\max \eta, \quad (6)$$

subject to:

$$(d(s_{i-1}, s_i) + d(s_i, BS)) * ec_{trl} + t_i * ec_{chg} \leq re_{w_i}, \quad (7)$$

$$0 \leq re_i \leq b \quad \forall i \in \{1, 2, \dots, |S|\}, \quad (8)$$

$$0 \leq re_{w_i} \leq b_w \quad \forall i \in \{1, 2, \dots, |W|\}, \quad (9)$$

where re_{w_i} denotes the residual energy of w_i and b_w denotes the battery capacity of w_i .

IV. PROPOSED SOLUTION

A. Determining the Service Boundary

As shown in Fig. 1, here are n sensor nodes and m mobile chargers in WRSN. Consequently, for the purpose of balancing the workload of each mobile charger, the objective of the sensor clustering algorithm is to divide the n sensors on the plane into m clusters, denoted as $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_m\}$, ensuring that the sensors within each cluster are spatially close and the number of sensors between clusters is balanced. We refer to this step as determining the service boundary of the mobile chargers. To achieve this, we propose a sensor clustering algorithm. Define $P_k(\bar{x}_k, \bar{y}_k)$ as the cluster center of Ω_k . The detailed pseudocode of this algorithm is given in Algorithm 1.

B. Selection of Target Charging Node

The maximum lifetime and remaining survival time of sensor s_i can respectively calculated as $ml_i = b/ec_i$ and $rl_i = re_i/ec_i$. s_i sends a charging request req_k to the base station when rl_i is less than a certain time threshold Θ . The following target charging node selection algorithm aims to

Algorithm 1: Sensor Clustering Algorithm

input : Sensor node set S , mobile charger set W , number of iterations M , parameters ϵ .
output: all clusters $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_m\}$.

- 1 Initialize each cluster center P_k randomly;
- 2 $\kappa \leftarrow \epsilon \times \left\lceil \frac{|S|}{|W|} \right\rceil$;
- 3 **for** $m \leftarrow 1$ **to** M **do**
- 4 $S' \leftarrow S$;
- 5 Randomly select a sensor node s from the sensor set S ;
- 6 **while** $S' \neq \emptyset$ **do**
- 7 Calculate the distance $d(s, P_k)$ between sensor node s and each cluster center P_k ;
- 8 **for** Ω_k **in** Ω **do**
- 9 **if** $|\Omega_k| \leq \kappa$ **then**
- 10 Add sensor node s into cluster Ω_k ;
- 11 **break**;
- 12 **end**
- 13 **end**
- 14 $S' \leftarrow S' \setminus \{s\}$;
- 15 Find the closest node s' to sensor s in S' ;
- 16 **end**
- 17 **for** Ω_k **in** Ω **do**
- 18 $\bar{x}_k \leftarrow \frac{1}{|\Omega_k|} \sum_{s \in \Omega_k} s(x)$;
- 19 $\bar{y}_k \leftarrow \frac{1}{|\Omega_k|} \sum_{s \in \Omega_k} s(y)$;
- 20 **end**
- 21 **end**
- 22 **return** All Ω_k ;

schedule mobile chargers to select a charging request from its charging queue for response.

Each request in the charging queue corresponds to a priority. If the queue is not empty, the mobile charger needs to select a charging request with the highest priority for response. The priority of a charging request consists of temporal priority and spatial priority. These two parts of priority are calculated as follows:

The temporal priority $\lambda_t(k)$ of the charging request $req_k(s_i)$ is:

$$\lambda_t(k) = \frac{rl_i}{ml_i} = \frac{re_i}{b}, \quad (10)$$

The spatial priority $\lambda_d(k)$ of the charging request $req_k(s_i)$ is:

$$\lambda_s(k) = \frac{d(s_i, w_j)}{D_j}, \quad (11)$$

where $d(s_i, w_j)$ denotes the distance between the mobile charger w_j and the sensor node s_i , and D_j denotes the distance between the farthest two sensor nodes in the service boundary.

Therefore the priority of the charging request $req_k(s_i)$ can be calculated as:

$$\lambda(k) = a_1 \lambda_t(k) + a_2 \lambda_s(k), \quad (12)$$

subject to:

$$a_1, a_2 \geq 0 \quad (13)$$

C. Learning Environment Construction

After selecting the target charging node, a mobile charger needs to schedule the charging duration at this chosen node. We propose a charging duration calculation algorithm based on deep reinforcement learning to address this. The mobile charger is an agent that interacts with the environment. In this section, we introduce how to construct a learning environment, including the state space, action space, and reward function for training the agent.

State Space. The environment state, denoted as $\langle S_{network}^k, S_{target}^k, S_{MC}^k \rangle$, comprises three components: the sensor network state, the target charging node state, and the mobile charger state. The sensor network state, represented as $S_{network}^k = \{S_1, S_2, \dots, S_h\}$, encompasses all sensor nodes serviced by the mobile charger w_k . Each S_i is a five-tuple that contains: the coordinates of the node s_i , its battery capacity, energy consumption rate, and current residual energy. For instance, $S_i = (40.0, 50.0, 1000.0, 0.25, 500.0)$ describes a sensor node s_i located at coordinates (40.0, 50.0) with a battery capacity of 1000J, an energy consumption rate of 0.25J/s, and a current residual energy of 500J. Similarly, S_{target}^k comprises the same five parts of information for the target charging node. S_{MC}^k is a seven-tuple which describes the state of mobile charger w_k : its current coordinates, residual energy, battery capacity, moving energy consumption rate, charging energy consumption rate, and moving speed.

Action Space. The action space of the mobile charger is represented as $A = \{t | 0 \leq t \leq 3\tau\}$. $a_t \in A$ denotes that the mobile charger w_k charges the target sensor node s_{target} for a_t seconds.

Reward Function. The reward function represents the benefits the agent can acquire after taking action a_t in the current state s_t . The reward signal guides the agent to learn a better strategy. To enhance survival rate and energy usage efficiency, we designed the reward function as follows:

$$r(s_t, a_t) = -\alpha \frac{N_{miss}}{N_{req}} + \beta \frac{E_{node}}{E_{chg} + E_{trl}}, \quad (14)$$

subject to:

$$\alpha, \beta \geq 0 \quad (15)$$

where N_{miss} denotes the number of dead sensor nodes after taking action a_t , and N_{req} denotes the number of sensor nodes that have been sent charging requests.

D. The Learning Algorithm

We train the agent with proximal policy optimization (PPO) algorithm [18]. Compared with other traditional RL algorithms, the PPO provides advantages in stability and sampling efficiency. PPO is a policy-based algorithm under the Actor-Critic framework. In our charging duration strategy learning algorithm, the charging duration calculation policy $\pi_\theta(s_t)$ is determined by the actor network. The state value function $V_\phi(s_t)$ which estimates the expected cumulative reward for a given state s_t can be calculated by critic network. Before training, the parameters of the actor and critic networks

are randomly initialized. Calculate the $\lambda(i)$ by Eq.(12), and select a sensor with lowest $\lambda(i)$ as target charging sensor s_{target} . Determine the charging duration a_t under the policy $\pi_\theta(s_t)$, where $s_t = \langle S_{network}, S_{target}, S_{MC} \rangle$. A charging task $k = (s_{target}, a_t)$ is constructed, and MC interacts with the environment to execute the task, receiving a reward $r(s_t, a_t)$. The agent observes the environment to obtain s_{t+1} , and the transition (s_t, a_t, r_t, s_{t+1}) is stored in the replay buffer.

In the PPO update phase, each transition in the replay buffer is retrieved. We use the generalized advantage estimation [19] to estimate the advantage. The advantage can be calculated as

$$A_{GAE}(s_t, a_t) = \sum_{l=t}^T (\gamma\lambda)^l (r_t + \gamma V_\phi(s_{l+1}) - V_\phi(s_l)) \quad (16)$$

where γ is discount factor, and λ is GAE hyperparameter. Thus, the target value V_t^{target} can be calculated as

$$V_t^{target} = A_{GAE}(s_t, a_t) + V_\phi(s_t) \quad (17)$$

A mini-batch of transition is randomly sampled from the replay buffer, and the policy π_θ is updated with the objective function

$$J(\theta) = \frac{1}{|RB|} \sum_{(s_t, a_t)} \min\left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} A_{GAE}(s_t, a_t), \right. \\ \left. clip\left(\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}, \epsilon\right) A_{GAE}(s_t, a_t)\right) \quad (18)$$

To learn the state-value function, a commonly used approach is to minimize a loss function based on mean squared error

$$L(\phi) = \frac{1}{|RB|} \sum_{(s_t, a_t)} \frac{1}{2} (V_t^{target} - V_\phi(s_t))^2 \quad (19)$$

With an objective function for policy improvement and a loss function for correcting the state-value function, the gradient of these functions enable gradient ascent on θ and gradient descent on ϕ

$$\theta_+ = \theta_- + \omega_\theta \nabla J(\theta)|_{\theta=\theta_-} \quad (20)$$

$$\phi_+ = \phi_- - \omega_\phi \nabla L(\phi)|_{\phi=\phi_-} \quad (21)$$

where ω_θ and ω_ϕ are the learning rates for policy and state-value functions respectively. The pseudocode of the algorithm training process is illustrated in Algorithm 2.

V. EXPERIMENTS

In this section, we conducted extensive experiments to evaluate the performances of our proposed algorithms and analyzed the simulation results by compared with the baseline algorithms.

Algorithm 2: Charging Duration Strategy Learning Algorithm

input : The target charging sensor s_{target} .
output: A charging duration of target charging sensor s_{target} .

```

1 Initialize actor network parameter  $\theta$  randomly;
2 Initialize critic network parameter  $\phi$  randomly;
3 for  $episode = 1, 2, \dots, E$  do
4   for  $t = 1, 2, \dots, T$  do
5     Calculate the  $\lambda(i)$  by Eq.(12), and select a
       sensor with lowest  $\lambda(i)$  as target charging
       sensor  $s_{target}$ ;
6     Use policy  $\pi_\theta(s_t)$  to determine the charging
       duration  $a_t$  of  $s_{target}$  when the state is
        $s_t = \langle S_{network}, S_{target}, S_{MC} \rangle$ ;
7     Construct the charging task  $k = (s_{target}, a_t)$ ;
8     MC executes the charging task and obtains
       reward  $r_t$  and  $s_{t+1}$ ;
9     Store the transition  $(s_t, a_t, r_t, s_{t+1})$  in replay
       buffer  $RB$ ;
10    if replay buffer  $RB$  is full then
11      Sampling minibatch size transition
         $(s_t, a_t, r_t, s_{t+1})$  from  $RB$ ;
12      Estimate the advantage  $A_{GAE}(s_t, a_t)$  and
        calculate the target value  $V_t^{target}$  by
        Eq.(16) and Eq.(17), respectively;
13      Update actor network parameters by
        Eq.(20);
14      Update critic network parameters by
        Eq.(21);
15      Empty the replay buffer  $RB$ ;
16    end
17  end
18 end

```

A. Simulation Setup

In the experiment, 100 to 300 sensor nodes are randomly placed in a $100m \times 100m$ region. The base station is located at the center of the network. Initially, all sensors are equipped with a full, rechargeable battery with a capacity of $b = 1000J$. The energy consumption rate of the sensor varies from $0.2J/s \sim 0.5J/s$. Meanwhile, the mobile chargers with high-capacity batteries are initially located at the base station. The simulation parameters are listed in Table I.

To evaluate the performance of the proposed TSPC-DRL algorithm, we compare it with the following charging algorithms: mTS [10], IOCS [11], and FLCSD [12]. These algorithms are full charging algorithms. We modified the algorithms which limit the charging duration to no more exceed τ

B. Performance Evaluation

Impact of Network Scale. We consider five networks with different scales, including 100, 150, 200, 250, and 300 sensor

TABLE I
SENSOR NETWORK SETTING

Parameter	Value
Network Size	$100m \times 100m$
Number of Nodes	100 ~ 300
Number of MCs	4 ~ 8
b_w , Energy capacity of MC	50000J
v , Speed of the MC	1m/s
ec_{trl} , Energy consumption rate for traveling	10J/m
ec_{chg} , Energy consumption rate for charging	20J/s
b , The energy capacity of sensor	1000J
ec_i , Energy consumption rate of sensor	$0.2J/s \sim 0.5J/s$
τ , time constant	100s

nodes, and respectively deploy 4-8 mobile chargers in different network sizes. We first study the performance of different algorithms by increasing the network scale. Figure 2(a) depicts the survival rate of sensor nodes when the number of sensors varies from 100 to 300. The survival rate tends to decrease along with the increasing number of sensors in all algorithms due to the workload on each mobile charger is rising. But the TSPC-DRL achieves a higher survival rate than the other algorithms. As shown in Table II, the energy usage efficiency in TSPC-DRL are around 0.33. FLCSD, mTS, and IOCS are respectively around 0.31, 0.29, and 0.27. TSPC-DRL still outperforms the other algorithms.

Impact of Moving Speed. In Figure 2(b), we discuss the impact of the moving speed of mobile chargers on the performance of different algorithms in a network with 100 sensor nodes. We found that the TSPC-DRL algorithm is more sensitive to the moving speed of mobile chargers. Increasing the movement speed of mobile chargers can increase about 5% the survival rate of sensor nodes in TSPC-DRL from 1m/s to 5m/s. However, increasing the movement speed of mobile chargers does not significantly improve the node survival rate for other algorithms. Therefore, the deep reinforcement learning-based charging algorithm proposed in this paper can better adapt to environmental changes. As shown in Table III, although increasing the movement speed of mobile chargers did not improve the energy efficiency of mobile chargers, the TSPC-DRL algorithm still exhibits the best energy efficiency compared to other algorithms.

Impact of Battery Capacity. In Figure 2(c), we investigate the impact of sensor battery capacity on the performance of different algorithms. We find that increasing the sensor's battery capacity will enhance the survival rate of sensor nodes. This is because a larger sensor node battery capacity results in a longer lifespan, reducing charging requests and alleviating the charging burden on mobile chargers. As shown in Table IV, increasing sensor battery capacity leads to greater energy usage efficiency, with the TSPC-DRL algorithm exhibiting superior energy usage efficiency compared to other algorithms.

VI. CONCLUSION

To improve the survival rate of sensor nodes and energy usage efficiency, we have proposed a temporal spatial partial charging scheme for WRSN with multiple mobile chargers in

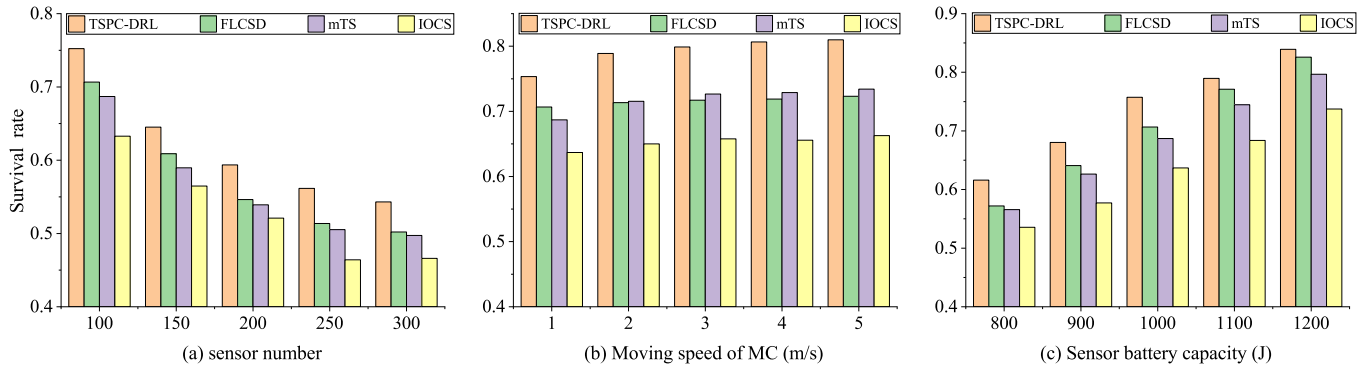


Figure 2. Impact of various parameters on the survival rate of sensor nodes.

TABLE II

IMPACT OF THE NETWORK SCALE ON THE ENERGY USAGE EFFICIENCY.

	100	150	200	250	300
TSPC-DRL	32.29%	33.13%	34.15%	34.06%	33.97%
FLCSD	30.72%	30.74%	31.01%	31.06%	31.09%
mTS	29.61%	29.74%	30.12%	30.37%	30.48%
IOCS	26.56%	27.59%	28.00%	28.02%	27.90%

TABLE III

IMPACT OF MOVING SPEED ON THE ENERGY USAGE EFFICIENCY.

	1.0m/s	2.0m/s	3.0m/s	4.0m/s	5.0m/s
TSPC-DRL	32.29%	33.46%	33.83%	33.90%	33.95%
FLCSD	30.72%	30.83%	30.91%	30.96%	31.04%
mTS	29.61%	30.26%	30.41%	30.51%	30.66%
IOCS	26.56%	26.93%	27.20%	27.04%	27.24%

TABLE IV

IMPACT OF SENSOR BATTERY CAPACITY ON THE ENERGY USAGE EFFICIENCY.

	800J	900J	1000J	1100J	1200J
TSPC-DRL	26.74%	29.52%	32.29%	34.84%	37.36%
FLCSD	24.69%	27.76%	30.72%	33.55%	36.04%
mTS	23.95%	26.73%	29.61%	32.32%	34.84%
IOCS	22.16%	24.20%	26.56%	28.98%	31.54%

this paper. Our charging strategy, termed TSPC-DRL, automatically determines the charging duration of the target charging node. Our proposed scheme has promising performance in prolonging the survival rate of the sensor node and improving the energy usage efficiency of MC compared with the modified mTS, IOCS, and FLCSD algorithm, which has been verified by the simulation experiments. Our future research will focus on designing a multi-node charging scheme for multiple MCs in WRSN.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 61702198. Shuangjuan Li is the corresponding author.

REFERENCES

- [1] M. S. BenSaleh, R. Saida, Y. H. Kacem, and M. Abid, "Wireless sensor network design methodologies: A survey," *Journal of Sensors*, vol. 2020, no. 1, p. 9592836, 2020.
- [2] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of wireless sensor networks: an up-to-date survey," *Applied system innovation*, vol. 3, no. 1, p. 14, 2020.
- [3] H. C. W. Y. and W. H., "Survey on charging programming in wireless rechargeable sensor networks," *Journal of software*, vol. 27, no. 1, pp. 72–95, 2015.
- [4] J. Liu, J. Peng, W. Xu, W. Liang, T. Liu, X. Peng, Z. Xu, Z. Li, and X. Jia, "Maximizing sensor lifetime via multi-node partial-charging on sensors," *IEEE Transactions on Mobile Computing*, 2022.
- [5] A. Kaswan, P. K. Jana, and S. K. Das, "A survey on mobile charging techniques in wireless rechargeable sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1750–1779, 2022.
- [6] W. Liang, W. Xu, X. Ren, X. Jia, and X. Lin, "Maintaining large-scale rechargeable sensor networks perpetually via multiple mobile charging vehicles," *ACM Transactions on Sensor Networks (TOSN)*, vol. 12, no. 2, pp. 1–26, 2016.
- [7] C. Zhao, H. Zhang, F. Chen, S. Chen, C. Wu, and T. Wang, "Spatiotemporal charging scheduling in wireless rechargeable sensor networks," *Computer Communications*, vol. 152, pp. 155–170, 2020.
- [8] L. He, L. Kong, Y. Gu, J. Pan, and T. Zhu, "Evaluating the on-demand mobile charging in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 9, pp. 1861–1875, 2014.
- [9] C. Lin, J. Zhou, C. Guo, H. Song, G. Wu, and M. S. Obaidat, "Tsca: A temporal-spatial real-time charging scheduling algorithm for on-demand architecture in wireless rechargeable sensor networks," *IEEE Transactions on Mobile Computing*, vol. 17, no. 1, pp. 211–224, 2017.
- [10] C. Lin, Z. Wang, J. Deng, L. Wang, J. Ren, and G. Wu, "mts: Temporal and spatial-collaborative charging for wireless rechargeable sensor networks with multiple vehicles," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 99–107.
- [11] Y. Dong, G. Bao, Y. Liu, M. Wei, Y. Huo, Z. Lou, Y. Wang, and C. Wang, "Instant on-demand charging strategy with multiple chargers in wireless rechargeable sensor networks," *Ad Hoc Networks*, vol. 136, p. 102964, 2022.
- [12] A. Tomar, L. Muduli, and P. K. Jana, "A fuzzy logic-based on-demand charging algorithm for wireless rechargeable sensor networks with multiple chargers," *IEEE Transactions on Mobile Computing*, vol. 20, no. 9, pp. 2715–2727, 2020.
- [13] W. Xu, W. Liang, X. Jia, and Z. Xu, "Maximizing sensor lifetime in a rechargeable sensor network via partial energy charging on sensors," in *2016 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 2016, pp. 1–9.
- [14] Y. Guo, X. Liu, and C. Chen, "Research on hybrid cooperative charging scheduling schemes in underwater sensor networks," *IEEE Access*, vol. 7, pp. 156452–156462, 2019.
- [15] W. Liang, Z. Xu, W. Xu, J. Shi, G. Mao, and S. K. Das, "Approximation algorithms for charging reward maximization in rechargeable sensor networks via a mobile charger," *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 3161–3174, 2017.
- [16] K. Wang, L. Wang, M. S. Obaidat, C. Lin, and M. Alam, "Extending network lifetime for wireless rechargeable sensor network systems through partial charge," *IEEE Systems Journal*, vol. 15, no. 1, pp. 1307–1317, 2020.

- [17] C. Lin, Y. Zhou, H. Dai, J. Deng, and G. Wu, "Mpf: Prolonging network lifetime of wireless rechargeable sensor networks by mixing partial charge and full charge," in *2018 15th annual IEEE international conference on sensing, communication, and networking (SECON)*. IEEE, 2018, pp. 1–9.
- [18] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [19] J. Schulman, P. Moritz, S. Levine, M. Jordan, and P. Abbeel, "High-dimensional continuous control using generalized advantage estimation," *arXiv preprint arXiv:1506.02438*, 2015.