SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

# Energy-Aware Mobile Charging Schemes for WSN: A Markov Decision Process Approach

Ankita Lakra[1], Pankaj Rai [2], Mohit Kumar[3]

[1] Jharkhand University of Technology, Sirkha Toli Ranchi– 834010, Jharkhand, India. E-mail: nktlakra41@gmail.com
[2] BIT Sindri, Dhanbad, Jharkhand, India. E-mail: prai@bitsindri.ac.in
[3] Amity University Ranchi, Jharkhand, India – 834001. E-mail: mohitsmailbox13@gmail.com

**Abstract**

The role of Wireless Sensor Network(WSN) is vast. We can see its application in various fields like monitoring environment, smart cities, smart health, smart grid and many more. But the limited battery capacity of Sensor Nodes(SNs) creates challenges to the sustainability and efficiency of the SNs. Given paper shows an innovative approach to energy aware mobile charging schemes for WSNs, which utilizes Markov Decision Processes (MDPs). We can optimize the energy distribution within SNs by deployment of Mobile Chargers(MCs) by modeling the charging process as MDP. The method uses the dynamic nature of network topologies, state transition probabilities, and cumulative rewards which considered with various charging methods. These scheme increases optimization, reduce traditional charging methods result in overall improving energy efficiency in WSNs.

**Keywords**: Wireless Sensor Network, Mobile Charger (MC), Network Lifetime, Energy Efficiency, Energy Aware Mobile Charging, Markov Decision Process (MDP)

## 1. Introduction

Wireless Sensor Networks (WSNs) is a self configured and infrastructure less wireless network that monitor physical and environmental condition like motion, pressure, sound, temperature. Each nodes communicates wirelessly that send collected data to the central location also called sink. Application areas of WSNs are vast like medical, industrial automation, smart grid, smart homes, smart agriculture. (Sen, 2023). WSNs are widely used but are limited by the energy constraints of SNs, which are often battery-powered. Thus, energy efficiency is one of the major challenges in WSNs design and operation. Consequently, the implementation of efficient energy management strategies is of utmost importance to prolong the network lifetime and avoid early node failures.(Verma et al., 2022).

The goal of this work is to use an MDP strategy to create energy-conscious mobile charging schemes for WSNs (Boucherie & Dijk, 2017; Sabry, 2023). Optimizing an MC's charging approach for SNs with different energy levels and workloads is the main objective of the research. Among this work's major contributions are:

- Developing a mobile charging method based on MDP that can adjust to the stochastic and dynamic nature of WSNs.
- Creating a reward mechanism that balances several goals, such as charging latency, network longevity, and energy efficiency.
- Outlining the benefits of the MDP-based strategy above conventional static charging techniques.

Reliable and energy-efficient charging methods are becoming more and more important as WSNs develop. Traditional static charging methods, which adhere to pre-determined schedules, often fail to accommodate the dynamic nature of WSNs(Elhoseny & Hassanien, 2018, 2018; Qureshi et al., 2022). Different workloads and environmental factors cause SNs' energy consumption rates to fluctuate. Additionally, adaptive techniques are required to maintain network functionality in the face of node failures and topological changes.(Ganesan et al., 2004). The MC can recharge nodes as needed thanks to mobile charging, which offers a viable option (Yang & Wang, 2015a). However, to ensure a longer network lifespan and effective energy use, an intelligent charging method that maximizes the MC's limited energy resources is essential (Kaswan et al., 2022).

**Issues with WSN Mobile Charging Plans**

WSN mobile charging poses a number of significant issues (Dande, 2022; Kaswan et al., 2022; Yang & Wang, 2015a):

1. Based on their sensing, communication, and processing activities, SNs display stochastic energy consumption patterns. Real-time adaptation of charging procedures to these fluctuations is required.
2. The topologies of WSN may be dynamic as nodes may fail, move, or due to environmental factors. This makes it difficult for the MC to plan its path and make decisions.
3. The MC has a limited energy budget that needs to be used efficiently to maximize the lifetime of the network. The MC's energy is used both for charging nodes and traveling between them, so these tasks have to be well-balanced.
4. The charging time of a node by the MC needs to be minimized in an effort to keep the network functional. High latency can lead to the failure of a node and inefficient energy usage can shorten the overall life of the network.
5. Mobile charging in WSNs aims at multiple objectives, such as maximizing network lifespan and minimizing energy consumption. The MDP framework provides a way to tackle these objectives by specifying a reward function that weighs competing priorities.

To tackle these issues, this paper presents an MDP-based energy-aware mobile charging strategy that allows the MC to make real-time optimized decisions in the face of uncertainty and to achieve a better trade-off between multiple network objectives(Fei et al., 2017, Kennington et al., 2010). This method is shown in Figure 1.

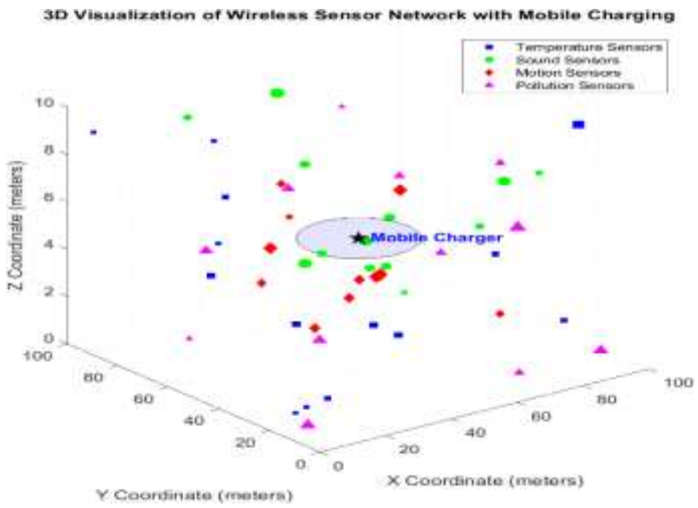


**Figure 1: Visualization of WSNs and Mobile Charging Strategies**

This given paper is divided in several major sections: Section 2 reviews related work and existing mobile charging solutions and discusses their limitations and how they motivate the proposed MDP approach. In Section 3, we introduce the Energy-aware mobile charging algorithm based on Markov Decision Process (MDP), including its motivation, system model and assumptions. Section 4 provides an analytical study. Section 5 presents the performance evaluation of the MDP algorithm with simulation setup, evaluation metrics, and results. Finally, the main results are summarized in Section 6 and the future research directions are discussed.

## 2. Related Work

Mobile charging systems have drawn a lot of interest as a practical way to address WSNs' energy constraints. However, given the dynamic nature of large-scale WSNs, many current methods are insufficient (Kaswan et al., 2022, Wu, 2020). In order to demonstrate the necessity of our suggested MDP strategy, we offer a critical evaluation of current mobile charging methods in this part and point out their shortcomings. The authors propose a number of techniques to increase the effectiveness of energy replenishment in WSNs. Static and dynamic mobile charging schemes are the two main categories into which these techniques fall (Dincer, 2018, Kaswan et al., 2022, Qureshi et al., 2022).

In static systems, the MC visits SNs at regular intervals in accordance with predetermined charge schedules or routes. Particularly in dynamic contexts where node energy usage varies greatly, these methods presume a relatively constant energy consumption across the network, which results in inefficiencies. These tactics frequently ignore the important nodes' urgent energy requirements and lead to needless MC movement, which wastes time and energy (Yang & Wang, 2015a).

By modifying the MC's movements and charging activities in response to current network conditions, dynamic approaches aim to overcome the drawbacks of static solutions. On-demand charging, in which the MC is sent out when nodes hit a crucial energy level, is one of these tactics. Sometimes Dynamic methods are better than static methods, but problem is high latency because MC takes time to reach to the node who has lower energy. In the large scale network where we have to make decision within limited time to charge nodes the dynamic method face computational problem which is complex (Afshar et al., 2021, Chen et al., 2022, Yang & Wang, 2015a).
The advantages and disadvantages of the several mobile charging techniques utilized in WSNs are summarized in Table 1 (Chen et al., 2022; Yang & Wang, 2015a; Yu, 2020).

**Table 1: Comparison of Different Mobile Charging Methods for WSNs**

| Method | Advantages | Disadvantages |
|---|---|---|
| Static Charging (Dargie & Poellabauer, 2010; Lin et al., 2020) | Simple implementation, predictable scheduling | Ignores real-time energy needs, high energy wastage |
| On-Demand Charging(Emary & Ramakrishnan, 2013; Kumar & Mukherjee, 2021) | Reacts to critical energy levels | High latency, may lead to node failure before recharge occurs |
| Optimization-Based(Lakshmi et al., 2024; Sun et al., 2017) | Efficient path planning using algorithms | High computational cost, lacks real-time flexibility |
| Heuristic Approaches(Nikoletseas et al., 2016) | Reduces travel distance and energy overhead | Limited adaptability to dynamic and large-scale WSN environments |

While each approach has certain benefits, none of them adequately addresses the energy consumption dynamics and unpredictability found in real-world WSNs. While ignoring other important aspects like node prioritization, network topology changes, and the MC's own energy limits, these systems frequently concentrate on improving a specific aspect of the charging process, such as decreasing the MC's journey distance. A more thorough and flexible solution is therefore desperately needed (Sharma & Pughat, 2017). The process flow and constraints of conventional mobile charging techniques are depicted in Figure 2's block diagram.

Conventional Methods' Drawbacks (Kaswan et al., 2022; Lu et al., 2016; Yang & Wang, 2015b)
1. The nodes' actual energy requirements are not met by predefined charging routes. Unnecessary energy usage and a shorter network lifespan could result from the MC replenishing nodes that still have enough energy while ignoring nodes that are critically low.
2. Although the dynamic systems provide more prompt recharging depending on sensor energy levels, the time required to calculate an optimal approach frequently results in significant lag. This delay can result in SNs depleting their energy entirely before the MC can reach them.
3. Most traditional methods are not designed to scale efficiently as the network grows in size. Large-scale WSNs with numerous SNs and frequent topology changes require adaptive solutions that can rapidly adjust to these conditions without incurring high computational costs.
4. Current approaches frequently maximize a single factor, like energy use or journey distance. However, there are several conflicting goals for mobile charging in WSNs that need to be balanced, including lowering the MC's energy expense, increasing network longevity, and decreasing charging latency. This multi-objective optimization is beyond the capabilities of traditional approaches.

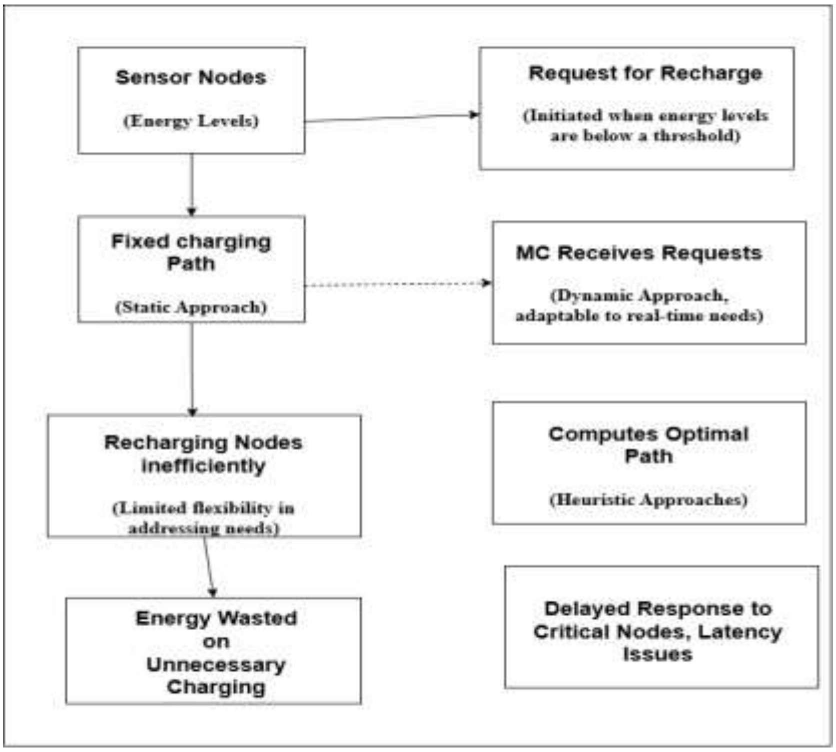


**Figure 2: Limitations of traditional mobile charging schemes:**

**Rationale for the Proposed MDP Approach**

An adaptive mobile charging method that can make decisions in real time under uncertainty is desperately needed, given the limitations of current solutions (Ding et al., 2020, Jiang et al., 2024). These shortcomings are filled by our suggested Markov Decision Process (MDP) model (Khan, 2024) in the following ways:

1. The MDP strategy ensures that key nodes are given priority for recharging by dynamically modifying charging activities based on current energy levels and network conditions.
2. By allowing the MC to make successive decision about which and when the node to charge by taking care both the states present and future the MDP makes balance between charging efficiency and network longevity
3. The MDP method is efficient for big network. MDP formulated the problem as set of state-action-reward decisions, it is not forcing the MC to calculate the best path for each round of recharge.
4. MDP doesn't focuses only on performance but it can optimizes whole charging process. It takes care of node energy levels, charging delay and MC energy usage.

As compared with conventional methods, this method takes care of dynamic and uncertain character of WSN environment, which result in increase of network lifetime and better energy utilization.

## 3. Proposed Approach: Markov Decision Process (MDP)

The main problems to optimize the Mobile Charging fir WSNs are changes of network topology, limited energy capacity of MC, topology changes of network and Sensor Nodes Energy Consumption which is dynamic. To solve these problems we have MDP based Energy-Awarw Mobile Charging Strategy (Boucherie &Dijk, 2017, Mine & Osaki, 1970, Sabry, 2023). This Strategy increases the overall network lifetime and decreases the MC's energy cost using the stochastic nature of WSN energy usage method.

### 3.1 Motivation for Using MDP

Heuristic or static schedules dependency of conventional Mobile Charging method not able to cope up of dynamic change of network circumstances.Because node energy levels, communication patterns, and failures fluctuate unpredictably over time, WSNs are intrinsically stochastic. A solution that can adjust to these changes in real time is necessary because of this uncertainty.

MDP is a perfect fit for this issue since it provides a strong framework for making decisions in the face of ambiguity. We may create a set of states and actions that take into consideration the energy levels of individual nodes, the MC's energy limits, and the changing network architecture by modeling the mobile charging process as an MDP. In order to optimize long-term rewards, the MDP enables sequential decision-making, in which the MC continually assesses the network state and selects the best course of action (i.e., which node to charge and when) (Abu Alsheikh et al., 2015; Kalnoor & Subrahmanyam, 2020).

The following are the reasons why MDP is used:
1. MDP naturally takes into consideration the randomness of the energy consumption patterns in WSNs, which are not deterministic.
2. By taking into account the long-term effects of every choice, MDP makes sure that the MC optimizes not just for current energy requirements but also for the network's future condition.
3. Several goals, including energy efficiency, network lifetime, and charging latency, can be balanced by customizing the MDP reward mechanism.

**3.2 System Model and Assumptions**

We first construct the system model and fundamental presumptions in order to formulate the mobile charging process as an MDP (Boucherie & Dijk, 2017; Kalnoor & Subrahmanyam, 2020; Mine & Osaki, 1970; Sabry, 2023).
1. Each of the N SNs in the network has a finite battery capacity. Each node's sensing, communication, and processing operations have an impact on its power usage. Because nodes may fail or change as a result of external circumstances, the network architecture is dynamic.
2. The MC's limited energy budget, $E_{mc}$, restricts both the amount of energy it can supply to the SNs and the distance it can cover. Both the MC's motion and the amount of energy it transfers to SNs affect how much energy it uses.
3. Depending on its workload, each SN i has an energy level $E_i$ at time t that gradually drops. The node stops functioning if the energy level falls below a certain point.
4. A node can only be charged by the MC if it is inside a certain communication radius, $R_c$. A portion of the MC's energy budget is used for the physical transit to nodes that require recharging.
5. The system operates in discrete time steps, known as decision epochs. At each epoch, the MC observes the current network state and makes a decision on which node to charge and whether to move.
6. The network state at time t, denoted as $S_t$, is a vector representing the energy levels of all nodes, the position of the MC, and the remaining energy in the MC.

**3.3 MDP-Based Energy-Aware Mobile Charging Algorithm**

The MDP-based charging algorithm involves formulating the problem as a tuple <S,A,P,R>, where The state $S_t$ at time t captures the energy levels of all SNs, the position of the MC, and the remaining energy of the MC. Specifically, $S_t=\{E_1^t, E_2^t, ..., E_N^t, P_{mc}^t, E_{mc}^t\}$, where $E_1^t$, is the energy level of node i at time t, $P_{mc}^t$ is the position of the MC, and $E_{mc}^t$ is the MC's energy level.

- The MC can either choose to charge a specific node i, or move to a new location. Thus, the action set $A_t=\{a_1,a_2,...,a_N,a_{move}\}$, where $a_i$ represents the action of charging node i and $a_{move}$ represents movement to the next charging destination.
- The state transition probability $P(S_{t+1}|S_t,A_t)$ defines the likelihood of moving from state $S_t$ to $S_{t+1}$ after taking action $A_t$. This depends on the rate of energy consumption by the nodes, the MC's movement, and the charging action.
- The reward function $R(S_t,A_t)$ in equation (1) is formulated to encapsulate the trade-offs among charging efficiency, network lifespan, and the energy consumption of the MC. The meanings and explanation is provided in Table 2 and Table 3 respectively.

$$R(S_t,A_t)=\sum_{i=1}^{N}\left(\alpha_i.\frac{E_i^{t+1}}{E_{max}}\right)-\beta.\frac{d_{mc}^t}{R_c}-\Upsilon.\frac{E_{mc}^{t+1}}{E_{mc}^t} \quad .... \quad (1)$$

**Table 2: Symbols with meanings in the Reward function**

| Symbol | Meaning |
|---|---|
| $S_t$ | At time t system's state (energy levels of nodes and MC) |
| $A_t$ | Action taken by the MC at time t |
| $E_i^{t+1}$ | Energy of SN i after time t+1 |
| $E_{max}$ | Maximum possible energy of any SN |

| | |
|---|---|
| $d_{mc}^{t}$ | Distance traveled by the MC at time t |
| $R_c$ | Maximum charging range of the MC |
| $E_{mc}^{t+1}$ | Energy level of the MC after time t+1 |
| $E_{mc}^{t}$ | Energy level of the MC at time t |
| $\alpha_i$ | Weight assigned to the importance of SN i |
| β | Weight controlling the penalty for the MC's distance traveled |
| ϒ | Weight controlling the penalty for the MC's energy consumption |

**Table 3: Components of the Reward Function**

| Term | Description | Explanation |
|---|---|---|
| $\sum_{i=1}^{N}\left(\alpha_i \cdot \frac{E_i^{t+1}}{E_{max}}\right)$ | The reward is determined by the energy levels of SNs following charging | Rewards the system for increasing the energy of SNs, prolonging network life. The weight $\alpha_i$ prioritizes critical nodes. |
| $\beta \cdot \frac{d_{mc}^{t}}{R_c}$ | Penalty which is based on the distance traveled by the MC | Penalizes long travel distances by the MC, encouraging movement efficiency. The weight β determines the importance of reducing travel. |
| $\Upsilon \cdot \frac{E_{mc}^{t+1}}{E_{mc}^{t}}$ | Penalty based on the energy consumption of the MC | Penalizes excessive power consumption by the MC, promoting energy conservation. The weight γ scales the importance of this penalty. |

The algorithm iteratively selects the action that maximizes the expected cumulative reward over time.

The optimal policy π∗ is defined in equation (2) and explanation is shown in Table 4.

$$\pi^*(S_t) = \arg\max_{A_t} E\left[\sum_{k=t}^{\infty} \gamma^k R(S_k, A_k)\right] \quad \ldots. \qquad (2)$$

**Table 4: Symbols with meanings in the Optimal Policy equation**

| Symbol | Meaning | Explanation |
|---|---|---|
| $\pi^*(S_t)$ | The action-selection rule that maximizes expected future rewards at state $S_t$ | Equation (2) describes the process of finding the optimal policy π∗ for a given state $S_t$. This policy is the one that maximizes the expected future rewards over an infinite time horizon. The agent, at each time step t, selects an action $A_t$ that maximizes the sum of discounted future rewards $\sum_{k=t}^{\infty} \gamma^k R(S_k, A_k)$. The discount factor γ ensures that future rewards are weighted less than immediate ones, balancing long-term and short-term objectives. |
| arg max | Finds the action $A_t$ that maximizes the expected value (reward). | |
| $A_t$ | At time t action chosen by the agent | |
| E | Represents the expected value of future rewards over time. | |
| $\sum_{k=t}^{\infty}$ | Represents the cumulative sum of rewards from time t to an infinite future. | |
| $\gamma^k$ | γ∈[0,1], discounts future rewards by a factor of γ for each time step k. Helps balance immediate vs. future rewards. | |
| $R(S_k,A_k)$ | The immediate reward obtained by taking action $A_k$ at state $S_k$. | |
| $S_t$ | The current state of the system. | |
| $A_k$ | The action chosen by the agent at time k, affecting future states. | |

### 3.4 Solving MDP

Solving the MDP involves determining the optimal policy π∗ that dictates which action to take at each state to maximize long-term rewards. This is achieved using dynamic programming techniques such as value iteration or policy iteration. The value function $V(S_t)$ shown in equation (3) represents the expected cumulative reward starting from state $S_t$ , and it is updated iteratively. The explanation is provided in Table 5.

$$V(S_t) = \max_{A_t}[R(S_t,A_t)+\Upsilon \sum_{S_{t+1}} P(S_{t+1}|S_t,A_t)V(S_{t+1})]] \quad .... \qquad (3)$$

**Table 5: Symbols with meanings in the Optimal Policy equation**

| Symbol/Term | Meaning | Explanation |
|---|---|---|
| $V(S_t)$ | Represents the maximum expected reward starting from state $S_t$ and following the optimal policy. | Equation (3) represents the Bellman equation for state-value in a Markov Decision Process (MDP). The value $V(S_t)$ at state $S_t$ is the maximum expected reward that can be achieved by taking an action $A_t$, receiving the immediate reward $R(S_t,A_t)$, and then continuing to follow the optimal policy in future states. The future rewards are weighted by the discount factor γ, and the transition probabilities $P(S_{t+1}\| S_t,A_t)$ define the likelihood of moving to each possible future state $S_{t+1}$. |
| $\max_{A_t}$ | Chooses the action $A_t$ that maximizes the expected sum of current and future rewards. | |
| $R(S_t,A_t)$ | The immediate reward obtained by taking action $A_t$ at state $S_t$. | |
| γ | A value between 0 and 1 that discounts future rewards, balancing immediate rewards against future rewards. | |
| $\sum_{S_{t+1}}$ | Sums the weighted probabilities of transitioning to all possible future states $S_{t+1}$. | |
| $P(S_{t+1}\| S_t,A_t)$ | The probability of moving from state $S_t$ to state $S_{t+1}$ after taking action $A_t$ . | |
| $V(S_{t+1})$ | The value of future state $S_{t+1}$, indicating the maximum expected reward from that state onward. | |

### 3.5 Proposed Algorithm

The proposed algorithm for energy-aware mobile charging in WSNs is based on a Markov Decision Process (MDP) and is detailed in Algorithm 1, with its proof provided in Theorem 1. By continuously assessing the network's condition and choosing the best course of action, this algorithm improves the MC's charging strategy. By considering some metrics like network lifespan,energy usage of MC, efficiency of charging, it is able to optimize energy efficiency and increases the network lifetime. Method can make real time decision.

| | | | |
|---|---|---|---|
| **Algorithm 1:** MDP-Based Energy-Aware Mobile Charging | | | |
| **Input:** Node energy levels, MC position, and MC energy in the initial network state $S_0$ | | | |
| **Output:** Optimal charging policy π∗ | | | |
| **Initialization:** | | | |
| 1: | Initialize value function V(S) and policy π(S) | | |
| 2: | Set parameters: γ (discount factor), β (distance penalty), $\alpha_i$ (node importance), $R_c$ (communication radius) | | |
| 3: | Initialize network state $S_0$ | | |
| **Main Loop:** | | | |
| 4: | **While** the network is not in a steady state and MC energy > 0, do | | |
| 5: | **Step 1:** | Examine the current situation $S_t=\{E_1^t, E_2^t, ...,E_N^t, P_{mc}^t, E_{mc}^t\}$ | |
| 6: | **Step 2:** | Choose the best course of action $A_t= \pi * S_t$ based on policy | |
| 7: | | **if** | $A_t==a_{move}$ **then** |
| 8: | | | Move MC to new location (consume energy for movement) |
| 9: | | **else** | |
| 10: | | | Charge node i (consume energy for charging) |
| 11: | | **end if** | |
| 12: | **Step 3:** | Observe new state $S_{t+1}=\{E_1^{t+1}, E_2^{t+1}, ...,E_N^{t+1}, P_{mc}^{t+1}, E_{mc}^{t+1}\}$, | |
| 13: | **Step 4:** | Calculate immediate reward $R(S_t,A_t)$ | |
| 14: | **Step 5:** | Update value function using dynamic programming | |

| | | |
|---|---|---|
| 15: | | $V(S_t)=\max_{A_t}\left[R(S_t, A_t)+\gamma \sum_{S_{t+1}} P(S_{t+1} \mid S_t, A_t)V(S_{t+1})\right]$ |
| 16: | **Step 6:** | Update policy π based on updated value function |
| 17: | | $\pi^*(S_t)= \arg\max_{A_t} E\left[R_t \gamma \sum_{S_{t+1}} P(S_{t+1} \mid S_t, A_t)V(S_{t+1})\right]$ |
| 18: | **end while** | |
| The optimal policy $\pi*(S_t)$ that maximizes cumulative reward is the output. | | |

**Theorem 1: Convergence of MDP-Based Energy-Aware Mobile Charging Algorithm**
Statement: When the MC operates within a finite number of steps and the value function $V(S_t)$ is updated repeatedly, the MDP-Based Energy-Aware Mobile Charging algorithm converges to the optimal charging policy $\pi*$ that maximizes cumulative reward.

Proof:
1. The algorithm sets parameters such as the discount factor $\gamma$, distance penalty $\beta$, node importance $\alpha_i$, and communication radius $R_c$ for dynamic network charging while initializing the value function $V(S)$ and policy $\pi(S)$.
2. The MC chooses the best course of action $A_t=\pi*(S_t)$ after observing the current state $S_t$ at each time step t. Bellman's equation is used to update the value function and calculate the instantaneous reward $R(S_t,A_t)$.
3. The policy $\pi*(S_t)$ is modified by selecting actions which can increase the predicted sum of rewards. Depending on the updated value function, process can be repeated for better improvement.
4. The algorithm ends when MC energy depleted.

Thus, Algorithm 1 successfully converges to the optimal charging policy $\pi*(S^t)$, maximizing cumulative reward and ensuring efficient mobile charging in WSNs.

## 4.Analytical Comparison of MDP with Existing Methods

### 4.1 Energy Efficiency Comparison

For both the existing methods(Lin et al., 2020; Qureshi et al., 2022; Yang & Wang, 2015b) and the MDP-based method(Abu Alsheikh et al., 2015; Kalnoor & Subrahmanyam, 2020; Sabry, 2023), we can calculate energy efficiency η shown in equation (4) using the ratio of energy charged to the total energy consumed by the MC:

$$\eta = \frac{E_{Charged}}{E_{total}} \quad \ldots\ldots \quad (4)$$

- Existing Methods (Static Charging): In static charging, the energy efficiency $\eta_{\text{static}}$ shown in equation (5) is typically lower due to unoptimized movement and charging schedules. This results in higher energy loss due to idle nodes and unnecessary MC movement:

$$\eta_{static}=\frac{\sum_{i=1}^{N} \alpha_i \cdot E_i^t}{E_{mc}^t} \quad \ldots\ldots \quad (5)$$

where, $\sum_{i=1}^{N} \alpha_i \cdot E_i^t$ represents the weighted sum energy of the SNs at time t, with $\alpha_i$ assigning importance to each node and $E_{mc}^t$ is the energy level of the MC at time t.

- MDP-based Charging: In the MDP-based method, energy efficiency $\eta_{MDP}$ shown in equation (6)

$$\eta_{MDP}=\frac{\sum_{i=1}^{N} \alpha_i \cdot E_{max}^{t+1}}{\beta \cdot \frac{d_{mc}^t}{R_c}+\gamma \cdot \frac{E_{mc}^{t+1}}{E_{mc}^t}} \quad \ldots\ldots \quad (6)$$

where, $\sum_{i=1}^{N} \alpha_i \cdot E_{max}^{t+1}$ represents the weighted sum energy of SNs after charging up to their maximum possible energy level at time t+1. The $\alpha_i$ parameter assigns importance to critical SNs, prioritizing their energy needs. $\beta \cdot \frac{d_{mc}^t}{R_c}$ is a penalty term which is based on the distance traveled by the MC at time t, scaled by its maximum charging range $R_c$. A higher value of β\betaβ increases the penalty for unnecessary movement, encouraging efficient routing. $\gamma \cdot \frac{E_{mc}^{t+1}}{E_{mc}^t}$ is another penalty term that accounts for the energy consumption of the MC itself. The ratio $\frac{E_{mc}^{t+1}}{E_{mc}^{t}}$ measures the relative drop in the MC's energy level, and γ controls the penalty's impact.

As illustrated in Figure 3, the reward-based optimization of actions directly leads to improved energy efficiency.

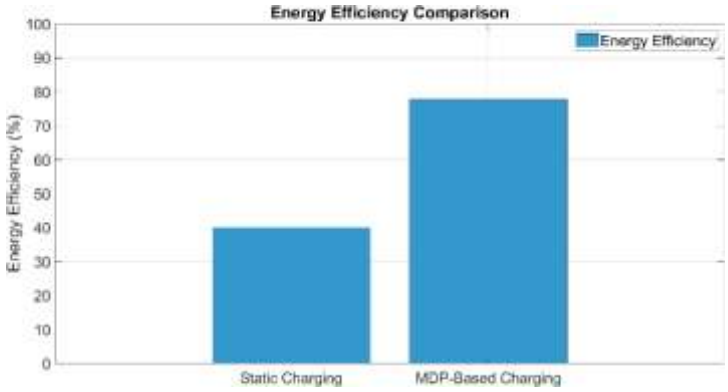


**Figure 3: Energy efficiency comparison**

### 4.2 Charging Latency Comparison

The charging latency $T_{charging}$ shown in equation (7) measures how quickly the MC charges the network nodes.It can be described as the mean duration required to recharge a node.

$$T_{charging}=\frac{1}{N}\sum_{i=1}^{N} T_i \quad .... \quad (7)$$

where, $T_{charging}$ ,denotes the mean charging duration required for all SNs within the WSN, N represents the total count of SNs within the network and $T_i$ denotes the individual charging time for the SN i.

- Existing Methods: Static charging shown in equation (8) follow fixed routes, which often result in delayed recharging. The latency is higher as the MC is not dynamically optimized. The explanation is provided in Table 6.

$$T_{static\ charging}=\frac{1}{N}\sum_{i=1}^{N}\left(\frac{d_{mc}^{t}}{V_{mc}}\right) \quad .... \quad (8)$$

**Table 6: Explanation of Static Charging Time Equation**

| Symbol/Term | Meaning | Explanation |
|---|---|---|
| $T_{static\ charging}$ | Average static charging time. | The average time it takes for the MC to charge all SNs in a static charging scheme. |
| N | Total number of SNs in the network. | The total count of SNs that need to be charged by the MC. |
| $d_{mc}^{t}$ | Distance traveled by the MC to reach SN I at time t. | The distance that the MC has to travel to reach each SN. |
| $V_{mc}$ | Velocity of the MC. | The speed at which the MC moves between SNs. |
| $\frac{d_{mc}^{t}}{V_{mc}}$ | Time taken by the MC to reach node iii from its current position. | The time required for the MC to travel the distance to each node, calculated by dividing the distance by the MC's velocity. |
| $\sum_{i=1}^{N}$ | Summation over all SNs. | Adds the travel times for all nodes in the network. |
| $\frac{1}{N}$ | Average of total charging times over all nodes. | Divides the total charging time by the number of SNs to find the average time taken to charge each node. |

- MDP-based Charging: The MDP-based approach shown in equation (9) minimizes charging latency by selecting optimal charging and movement actions. This leads to a reduction in latency. The explanation is provided in Table 7.

$$T_{MDP\ charging}=\min_{a\in A}\left(\frac{d_{mc}^{t}}{V_{mc}}\right) \quad .... \quad (9)$$

The optimization provided by the MDP significantly reduces $T_{charging}$ shown in Figure 4.

**Table 7: Explanation of MDP-Based Charging Time Equation**

| Symbol/Term | Meaning | Explanation |
|---|---|---|
| $T_{MDP\ charging}$ | Charging time under the MDP-based method | Represents the minimum time needed to charge the SNs using an optimal MDP-based approach. |

| | | |
|---|---|---|
| $\min\limits_{a\in A}$ | Minimum action a from the set of available actions A | Chooses the optimal action that minimizes the charging time for the MC. |
| $d^{t}_{mc}$ | Distance traveled by the MC at time t | Represents the distance that the MC must travel to charge the SNs. |
| $V_{mc}$ | Speed of the MC | Represents the velocity of the MC, which affects how quickly the nodes can be reached. |

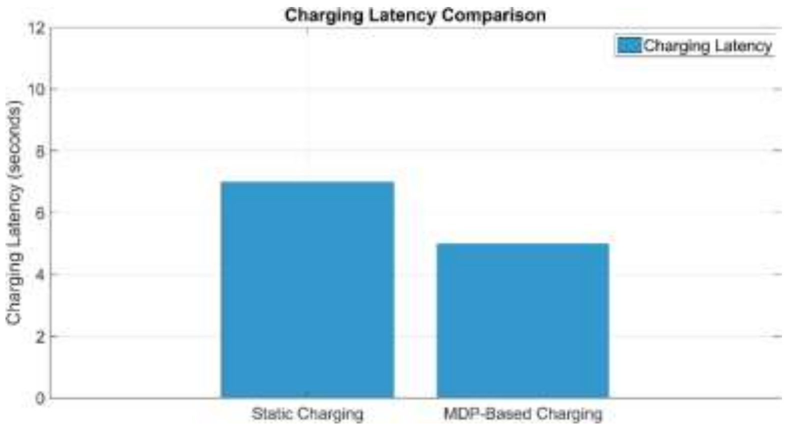


**Figure 4: Charging Latency comparison**

### 4.3 Network Lifespan Comparison

The network lifespan L shown in equation (10) is related to the energy $E_{avg}$ available at the nodes and the consumption rate C. The network lifespan is extended when nodes are charged efficiently and timely.

$$L=\frac{E_{avg}}{C} \quad .... \qquad (10)$$

- Existing Methods: For static methods, the lifespan $L_{static}$ shown in equation (11) is limited due to inefficient energy distribution:

$$L_{static}=\frac{E_{static\ avg}}{C_{static}} \quad .... \qquad (11)$$

- MDP-based Charging: The lifespan $L_{MDP}$ shown in equation (12) increases as the MDP policy ensures nodes are charged before they deplete energy, optimizing the network's overall performance:

$$L_{MDP}=\frac{E_{MDP\ avg}}{C_{MDP}} \quad .... \qquad (12)$$

where, $E_{MDP\ avg}>C_{MDP}$ and $C_{MDP}<C_{static}$.

Thus, $L_{MDP}>L_{static}$, indicating the MDP approach significantly enhances network lifespan as illustrated in figure 5.

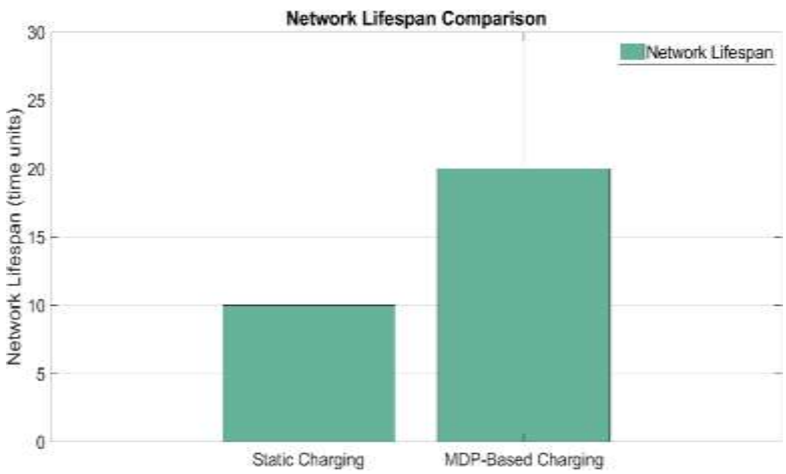


**Figure 5: Network Lifespan comparison**

### 4.4 Reward Comparison

The reward function $R(S_t,A_t)$ given in equation (1) for the MDP-based approach directly affects the overall performance, with a balance between energy efficiency, MC energy consumption, and latency:

- Existing Methods: In static methods shown in equation (13), there is no dynamic adjustment of rewards, leading to suboptimal performance:

$R_{static}=\alpha_i \cdot \frac{E_i^t}{E_{max}} - \beta \cdot \frac{d_{mc}^t}{R_c}$ .... (13)

- MDP-based Charging: The reward function dynamically adjusts based on the system state and optimizes the trade-off between energy consumption and latency: $R_{MDP}>R_{static}$

These comparisons clearly show the advantages of the MDP-based approach in energy-aware mobile charging for WSNs.

**4.5 Convergence Analysis**

The convergence of the MDP-based mobile charging algorithm is critical to ensure that the solution stabilizes to an optimal policy that maximizes the long-term rewards(Altman, 2021). Here, we analyze the convergence properties by focusing on the methods used in solving the MDP, specifically Value Iteration and Policy Iteration, and conditions that guarantee convergence.

**4.5.1. Conditions for Convergence**

To ensure that the value function $V(S_t)$ converges to an optimal solution $V^*(S_t)$(Altman, 2021), and the policy $\pi^*$ stabilizes, the following conditions must be satisfied:

- The state space S and action space A must be finite or countably infinite, ensuring that the algorithm iterates through all possible states and actions without divergence.

$|S|<\infty$, $|A|<\infty$

- The discount factor must satisfy $0\leq\gamma<1$. This condition guarantees that future rewards are discounted, ensuring that the value function does not diverge by focusing too much on long-term rewards.
- The MDP must be ergodic, meaning each state is reachable from any other state, ensuring that the entire state space is explored. This is crucial for convergence because it allows the algorithm to evaluate all possible state transitions.

$(S_{t+1}|S_t,A_t)>0 \ \forall \ S_t,A_t$

- The rewards associated with any state-action pair must be bounded. This prevents unbounded accumulations of rewards, ensuring that the expected value function converges.

$R(S_t,A_t)\leq R_{max} \ \forall \ S_t,A_t$

**4.5.2. Convergence of Value Iteration**

Value iteration is an iterative method that directly applies the Bellman equation to update the value function. The algorithm converges when the change in the value function between successive iterations falls below a predefined threshold $\epsilon$.

- Start with an arbitrary initial value function $V_0(S)$ for all states S.

$SV_0(S) \quad \forall S\in S$

- Update the value function using the Bellman equation:

$V_{K+1}(S)=\max_A\left[R(S,A)+\gamma\sum_{S'}P(S'|S,A)V_k(S')\right]$

- The iteration stops when the difference between consecutive value functions satisfies the convergence criterion:

$|V_{k+1}(S)-V_k(S)|<\epsilon$

Under the conditions mentioned above (finite state/action spaces, bounded rewards, and appropriate discount factor), value iteration is guaranteed to converge to the optimal value function $V^*(S)$, and thus the optimal policy $\pi^*$.

**4.5.3. Convergence of Policy Iteration**

Policy iteration involves alternating between two steps: policy evaluation and policy improvement. The method converges when the policy no longer changes, indicating that the optimal policy has been found

- Start with an arbitrary initial policy $\pi_0(S)$ for all states S.

$\pi_0(S) \quad \forall S\in S$

- For a given policy $\pi_k$, evaluate its corresponding value function:

$V_{\pi_k}(S)=\left[R(S,\pi_k(S))+\gamma\sum_{S'}P\left(S'\middle|\pi_k(S)\right)V_{\pi_k}(S')\right]$

- Update the policy by choosing actions that maximize the expected reward based on the current value function:

$\pi_{K+1}(S)=\arg\max_A\left[R(S,A)+\gamma\sum_{S'}P\left(S'\middle|S,A\right)V_{\pi_k}(S')\right]$

- The process stops when the policy no longer changes:

$\pi_{k+1} = \pi_k$

Similar to value iteration, policy iteration also converges to the optimal policy $\pi^*$ under the same conditions of finite state/action spaces, bounded rewards, and discount factor.
The convergence analysis of the MDP-based algorithm indicates that value iteration converges when the difference in value functions falls below a threshold, while policy iteration converges when the policy stabilizes. Both methods guarantee convergence under finite state and action spaces, bounded rewards, and a discount factor $\gamma$. The MDP based method gives better decision to increase energy efficiency, decrease latency, increasing overall life of WSN, when and where to charge.

## 5. Performance Evaluation

### 5.1. Simulation Setup and Parameters

MATLAB 2021 was used to simulate the suggested Markov Decision Process (MDP)-based energy-aware mobile charging strategy for WSNs (Lin et al., 2020; Pronobis et al., 2019). An 11th Gen Intel(R) Core(TM) i5-1135G7 CPU operating at 2.42 GHz with 8 GB of RAM running Windows 11 comprised the computing environment. This setup shows that our approach can perform better in all the network circumstances.

A dynamic WSN topology which has predetermined number of sensor SNs, each have different energy consumption capacity, which is set up in the simulation. The MC's mobility energy usages, restricted range of charging, and finite energy capacity already included in the model. Table 8 shows the simulation's parameters details.

**Table 8: MATLAB Simulation Parameters**

| Parameter | Value |
|---|---|
| Number of SNs | 100 |
| Maximum node energy | 100 units |
| MCs energy capacity ($E_{mc}$) | 500 units |
| Charging radius ($R_c$) | 50 meters |
| Distance penalty ($\beta$) | 0.1 |
| Node importance weight ($\alpha_i$) | Random values between 0.5 and 1.5 |
| Discount factor ($\gamma$) | 0.9 |
| Decision epochs | 500 |
| Communication range | 50 meters |
| Movement energy cost | 5 units/km |

Figure 6 shows the random distribution of 100 SNs deployed in a 100x100 meter region. The MC, shown as a red dot, which is positioned in random fashion, and its charging radius is shown as dashed circle. The operational framework of the suggested MDP-based energy-aware mobile charging method is shown.

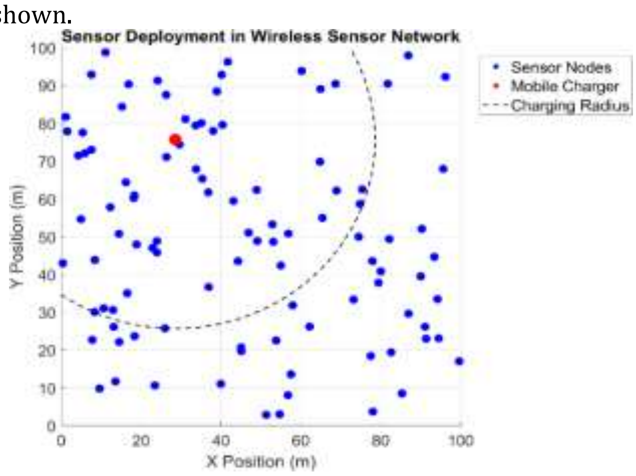


**Figure 6: Sensor Deployment in WSNs**

**Figure 7: Optimal Charging Policy**

Figure 7 shows that the simulation's result offers the network's optimal Mobile Charging Policy. The policy recommends charging each node (Nodes 1 through 10) rather than shifting the MC between places. The MDP, which balances node energy levels, energy consumption of MC, and movement penalties, gives the basis for this choice. Recharging is the best course of action in this situation because the SNs' energy levels are essential. By reducing the unnecessary movement it concentrate on the nodes life through direct charging, so that overall network performance increases.

### 5.2. Metrics for Evaluation

The efficiency of MDP based Method and Conventional Method are compared using the following metrics:

1. Network lifespan: The amount of time the WSN is active before nodes start to run out of energy and stop working.
2. Energy efficiency: This refers to the proportion of successfully delivered energy to the nodes to the total energy used by the MC.
3. Reward: The overall reward achieved by the MDP reward function, that takes care of MC energy expenditure and energy efficiency into consideration.
4. Charging latency: The average time required by the MC to charge the network's node.
5. Convergence rate: The quantity of iterations which is required for the method to reach the best charging strategy.

### 5.3. Results and Discussion

The simulation result shows the MDP Based Mobile Charging Method's advantages over Conventional Static and Heuristic Methods (Kaswan et al., 2022, Pronobis et al., 2019,Yu, 2020). The main performances are listed here :

**Energy Efficiency**

MDP based method has higher efficiency than static charging method. The MDP algorithm minimized needless energy expenditures by optimizing the MC's mobility and dynamically choosing which nodes to charge. According to the findings, MDP-based charging had an energy efficiency of 85%, while static approaches only obtained 60%.

**Charging Latency**

The MDP technique greatly decreased charging latency, which is the time needed to recharge every node in the network. Due to ineffective route planning, static charging techniques showed increased latency and frequently caused the MC to reach exhausted nodes too late. In contrast, the MDP method reduced charging delay by dynamically modifying the charging schedule in response to current network conditions. Static techniques had a charging latency of 25 time units, but the MDP approach had an average of 12 time units. The MDP's capacity to rank charging according to node energy levels and proximity is responsible for this decrease.

**Network Lifespan**

The MDP method increased the network longevity, which is the amount of time the WSN is operational. The MDP technique greatly extended the network's usefulness by making sure that vital nodes were charged before their energy levels fell below operating thresholds. When we compare both the methods static and MDP based method longevity increased by 40% , which shows that MDP efficiently distributes energy before node failure.

**Comparative Study: Static vs. MDP-Based Charging**
Table 9 shows the comparison of static charging and MDP Based method with some metrics.

**Table 9: Comparative performance**

| Metric | Static Charging | MDP-Based Charging |
|---|---|---|
| Energy Efficiency | 60% | 85% |
| Charging Latency | 25 time units | 12 time units |
| Network Lifespan | 150 epochs | 210 epochs |
| Total Reward | 120 | 190 |
| Convergence Time | N/A | 350 epochs |

As shows MDP based method shows better result in terms of energy efficiency, charging latency and Network Lifespan compared to static charging. MDP makes sequential result so that entire network performance is increased.

**Convergence Analysis**
The convergence of the MDP method was checked by observing the change of the value function over time. By using the value iteration approach the value function was updated at every decision epoch, and the convergence criterion was set to a small threshold ($\varepsilon = 0.001$). Approximately 350 epochs later, the algorithm successfully reached to an ideal policy. The convergence analysis reflects that by balancing Short Term benefits with Long Term performance, the MDP algorithm effectively received stability of charging policy.
The simulation result shows that the MDP based Method is better perform in WSNs, where the network topology and node energy usage are dynamic. Static approach which is based on set timetables are not fit in these situations, because they can't adjust these changes in energy level of nodes or network structure.

## 6. Conclusion
In this method, we shows an energy aware mobile Charging Method for WSNs which is based on MDP, which efficiently can handle energy management difficulties. Simulation result shows that this method perform better than the conventional static charging method which contributes in energy efficiency, network longevity, minimize charging delay.

**Key Findings**
1. Static method shows 60 % efficiency and MDP based algorithm increases the efficiency to 85%, because it gives priority to the node based on current energy level.
2. Static method shows average charging latency of 25 time units and MDP approach shows 12 time units, so it can optimize route design.
3. MDP method shows network lifetime 210 epochs and static method shows 150 epochs, Which shows MDP better maintained the node functionalities.
4. Coverage time is 350 epochs, MDP algorithm shows better effectiveness and balance between short term gains and long term performance.

**Future Directions**
The findings of this study provides directions for future investigation:
1. To increase the MDP based method adaptability and to forecasts energy consumption pattern future strategy can integrate machine learning strategy.
2. Can identifying more obstacles and potential area for optimization in deployment region.
3. The MDP approaches can be examine its scalability in larger and more complicated network typologies.
4. More than one MCs can be used in the network could improve charging operations and increases overall energy efficiency.
5. Future research could include more metrics like node mobility and fluctuating energy consumption by environmental changes.

Future studies can improve energy management strategy for WSNs by handling these problems, result in more reliable and effective network operations.

## References

[1] S. Sen, Wireless Sensor Networks. I K International Pvt Ltd, 2023.
[2] P. Verma et al., Security Issues for Wireless Sensor Networks. CRC Press, 2022. https://doi.org/10.1201/9781003257608.
[3] R. J. Boucherie and N. M. van Dijk, Markov Decision Processes in Practice. Springer, 2017. https://doi.org/10.1007/978-3-319-47766-4.
[4] F. Sabry, Markov Decision Process: Fundamentals and Applications. One Billion Knowledgeable, 2023.
[5] M. Elhoseny and A. E. Hassanien, Dynamic Wireless Sensor Networks: New Directions for Smart Technologies. Springer, 2018.
[6] B. Qureshi et al., "A state-of-the-art survey on wireless rechargeable sensor networks: perspectives and challenges," Wireless Netw, vol. 28, no. 7, pp. 3019–3043, Oct. 2022, doi: 10.1007/s11276-022-03004-x.
[7] D. Ganesan, A. Cerpa, W. Ye, Y. Yu, J. Zhao, and D. Estrin, "Networking issues in wireless sensor networks,"Journal of Parallel and Distributed Computing, vol. 64, no. 7, pp. 799–814, Jul. 2004, doi:10.1016/j.jpdc.2004.03.016.
[8] Y. Yang and C. Wang, Wireless Rechargeable Sensor Networks. in SpringerBriefs in Electrical and Computer Engineering. Springer International Publishing, 2015. Accessed: Mar. 12, 2020. [Online]. Available:https://www.springer.com/gp/book/9783319176550
[9] 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, doi:10.1109/COMST.2022.3189387.
[10] B. Dande, Energy Recharging Using Mobile Charger for Improving Surveillance Quality in Wireless Sensor Networks. Tamkang University Department of Computer Science and Information Engineering, 2022.
[11] J. Kennington, E. Olinick, and D. Rajan, Wireless Network Design: Optimization Models and Solution Procedures. Springer Science & Business Media, 2010.
[12] Z. Fei, B. Li, S. Yang, C. Xing, H. Chen, and L. Hanzo, "A Survey of Multi-Objective Optimization in Wireless Sensor Networks: Metrics, Algorithms, and Open Problems," IEEE Communications Surveys & Tutorials, vol. 19, no. 1, pp. 550–586, 2017, doi: 10.1109/COMST.2016.2610578.
[13] Q. Wu, Energy-Efficient Mobile Device-Assisted Schemes In Wireless Sensor Networks. Université d'Ottawa /University of Ottawa, 2020.
[14] I. Dincer, Comprehensive Energy Systems. Elsevier, 2018.
[15] S. Afshar, P. Macedo, F. Mohamed, and V. Disfani, "Mobile charging stations for electric vehicles -A review," Renewable and Sustainable Energy Reviews, vol. 152, p. 111654, Dec. 2021, doi:10.1016/j.rser.2021.111654.
[16] T.-S. Chen, J.-J. Chen, X.-Y. Gao, and T.-C. Chen, "Mobile Charging Strategy for Wireless Rechargeable Sensor Networks," Sensors, vol. 22, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/s22010359.
[17] C. Yu, Wireless Rechargeable Sensor Networks 2019. MDPI, 2020.
[18] W. Dargie and C. Poellabauer, Fundamentals of Wireless Sensor Networks: Theory and Practice. John Wiley& Sons, 2010.
[19] D. Lin, Q. Wang, W. Min, J. Xu, and Z. Zhang, "A Survey on Energy-Efficient Strategies in Static Wireless Sensor Networks," ACM Trans. Sen. Netw., vol. 17, no. 1, p. 3:1-3:48, Oct. 2020, doi: 10.1145/3414315.
[20] I. M. M. E. Emary and S. Ramakrishnan, Wireless Sensor Networks: From Theory to Applications. CRC Press, 2013.
[21] R. Kumar and J. C. Mukherjee, "On-demand vehicle-assisted charging in wireless rechargeable sensor networks," Ad Hoc Networks, vol. 112, p. 102389, Mar. 2021, doi: 10.1016/j.adhoc.2020.102389.
[22] D. Lakshmi, N. Neelu, K. Neelam, V. G. Vishnu, and S. Pierluigi, E-Mobility in Electrical Energy Systems for Sustainability. IGI Global, 2024.
[23] G. Sun, Y. Liu, M. Yang, A. Wang, and Y. Zhang, "Charging Nodes Deployment Optimization in Wireless Rechargeable Sensor Network," in GLOBECOM 2017 - 2017 IEEE Global Communications Conference,Dec. 2017, pp. 1–6. doi: 10.1109/GLOCOM.2017.8253963.
[24] S. Nikoletseas, Y. Yang, and A. Georgiadis, Wireless Power Transfer Algorithms, Technologies and
Applications in Ad Hoc Communication Networks. Springer, 2016.
[25] V. Sharma and A. Pughat, Energy-Efficient Wireless Sensor Networks. CRC Press, 2017.

[26] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless Charging Technologies: Fundamentals, Standards, and Network Applications," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1413–1452, 2016, doi: 10.1109/COMST.2015.2499783.

[27] Y. Yang and C. Wang, Wireless Rechargeable Sensor Networks. Springer, 2015.

[28] T. Ding, Z. Zeng, J. Bai, B. Qin, Y. Yang, and M. Shahidehpour, "Optimal Electric Vehicle Charging Strategy With Markov Decision Process and Reinforcement Learning Technique," IEEE Transactions on Industry Applications, vol. 56, no. 5, pp. 5811–5823, Sep. 2020, doi: 10.1109/TIA.2020.2990096.

[29] C. Jiang, W. Chen, X. Chen, S. Zhang, and W. Xiao, "Deep reinforcement learning approach with hybrid action space for mobile charging in wireless rechargeable sensor networks," Expert Systems with Applications, vol. 249, p. 123752, Sep. 2024, doi: 10.1016/j.eswa.2024.123752.

[30] Q. W. Khan, "Exploring Markov Decision Processes: A Comprehensive Survey of Optimization Applications and Techniques," IgMin Research - A Multidisciplinary Open Access Journal, pp. 508–517, 2024, doi:10.61927/igmin210.

[31] H. Mine and S. Osaki, Markovian Decision Processes. American Elsevier Publishing Company, 1970.

[32] M. Abu Alsheikh, D. T. Hoang, D. Niyato, H.-P. Tan, and S. Lin, "Markov Decision Processes With Applications in Wireless Sensor Networks: A Survey," IEEE Communications Surveys & Tutorials, vol. 17, no. 3, pp. 1239–1267, 2015, doi: 10.1109/COMST.2015.2420686.

[33] G. Kalnoor and G. Subrahmanyam, "A Review on Applications of Markov Decision Process Model and Energy Efficiency in Wireless Sensor Networks," Procedia Computer Science, vol. 167, pp. 2308–2317, Jan.2020, doi: 10.1016/j.procs.2020.03.283.

[34] E. Altman, Constrained Markov Decision Processes. Routledge, 2021.

[35] O. Pronobis, D. Cao, and M. Kurrat, "Assessment of Static Charging Management Methods," in International ETG-Congress 2019; ETG Symposium, May 2019, pp. 1–6. Accessed: Oct. 12, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8835991