

Performance Optimization of RF Energy Harvesting and Task Scheduling for Battery-Free Wireless Sensor Networks

Zacheous Aasa

`zacheous.aasa@gmail.com`

Ladoke Akintola University of Technology, LAUTECH

Fanuel Elias

Manchester Metropolitan University

Sunday C Ekpo

Manchester Metropolitan University

Chinonyerem Nwankwo

Manchester Metropolitan University

Noushin Karimian

Manchester Metropolitan University

Research Article

Keywords: Passive Beamforming, Reconfigurable Intelligent Surfaces (RIS), Radio Frequency (RF) Energy Harvesting, Battery-Free Wireless Sensor Networks, Multi-Agent Reinforcement Learning (MARL), Task Scheduling; Energy Neutral Operation (ENO), Internet of Things (IoT), 6G Communication

Posted Date: May 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9752516/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Performance Optimization of RF Energy Harvesting and Task Scheduling for Battery-Free Wireless Sensor Networks

Zacheous Aasa¹[0009-0007-9010-4936], Fanuel Elias²[0009-0007-7684-0339], Sunday C. Ekpo²[0000-0001-9219-3759], Chinonyerem Nwankwo²[0009-0008-1204-6301]

and Noushin Karimian²[0003-0002-0304-0185]

¹Department of Electrical and Electronics Engineering, Ladoke Akintola University of Technology, LAUTECH, Ogbomoso, Nigeria

zacheous.aasa@gmail.com

²Communication and Space Systems Engineering Research Team, Manchester Metropolitan University, Manchester, M1 5GD, UK

Fanuel.elias@stu.mmu.ac.uk; s.ekpo@mmu.ac.uk;
chinonwankwo05@gmail.com; n.karimian@mmu.ac.uk

* Corresponding author: Zacheous Aasa

Abstract. Expanding the Internet of Things (IoT) ecosystem has necessitated sustainable energy solutions, which has led to the development of battery-free WSNs. However, the stochastic nature of RF availability combined with the computational complexity of WSN tasks poses a major problem for the sustainability of battery-free WSNs. This study focuses on optimising RF power consumption and scheduling of tasks in WSNs without batteries. To achieve this, we propose a framework based on MARL, which allows WSNs to facilitate autonomous decision-making. The MARL framework optimizes the total throughput, efficiency while minimizing task drop rates based on the interaction between harvesting rates, buffer size and task periods in the WSN. The experimental results show a 28.5 percent increase in total throughput and a 42 percent reduction in the number of tasks that are dropped using the proposed framework compared to traditional single-agent and greedy scheduling approaches. The system also maintained a 94% Energy Neutral Operation (ENO) success rate even when faced with fluctuating levels of RF Power Densities from -20 dBm to -10 dBm. This shows that MARL has successfully addressed the issue of energy depletion and has established a reliable self-sustaining ecosystem, providing a foundation for 6G autonomous sensing.

Keywords: Passive Beamforming, Reconfigurable Intelligent Surfaces (RIS), Radio Frequency (RF) Energy Harvesting, Battery-Free Wireless Sensor Networks, Multi-Agent Reinforcement Learning (MARL), Task Scheduling; Energy Neutral Operation (ENO), Internet of Things (IoT), 6G Communication.

1 Introduction

The rapid growth of the Internet of Things (IoT) has led to a paradigm shift in the development of a sustainable and autonomous environment for sensing, particularly with the advent of the 6G era [1, 2]. The traditional wireless sensor network is being challenged by the limited lifetime and environmental risks of chemical batteries, which require the development of battery-free technologies [3, 4]. Radio frequency energy harvesting (RFEH) is recognized as one of the leading technologies in battery-free technology, allowing sensors to operate indefinitely using ambient or specialized electromagnetic waves [5, 6]. However, the inherent instability of RF power sources poses a major challenge for devices that reconcile stochastic energy sources with deterministic computing tasks [7,8]. The current progress in energy-neutral operation (ENO) ensures that the energy consumption of the equipment does not exceed the energy extracted during the period in question [9, 10]. However, the optimal performance of high density 6G enabled IoT networks remains a challenge, given the complex relationship between buffer management, time to tasks and density of energy [11, 12]. Conventional approaches to task scheduling, such as over-scheme or heuristic scheduling, were found insufficient to adapt to changes in the energy consumption patterns of tasks, resulting in high drop rates [13, 14, 15]. Current studies indicate that more sophisticated approaches to task management are needed to cope with RF environment uncertainties [16, 17].

Artificial intelligence, in particular multi-agent reinforcement learning, provides a solid framework for decentralized decision making in multi-nodes WSNs [18, 19]. Unlike traditional one-stop systems, MARL allows nodes to learn to make cooperative decisions in order to achieve maximum throughput and fairness between them [20, 21]. In this research, the issue of energy scheduling and energy optimization is addressed by a dynamic intelligent system for wireless sensor networks which uses synergies between energy consumption rates and energy performance to bridge the gap between theoretical energy sustainability and reliability [22, 23]. Building on previous literature on various safe and energy efficient frameworks for WSNs [27] and on recent successes in RF sustainability and smart surfaces [29]. This study provides a comprehensive solution to optimize the next generation of self-sustaining autonomous sensing nodes. The following contributions are made to overcome the problems of energy depletion and inefficient work in WSNs battery-free architectures.

- 1** Implementation of the MARL-based Optimization Framework: Developed a Multi-Agent Reinforcement Learning framework that allow the WSN nodes to autonomously decide on task execution and energy storage in real time.
- 2** Dynamic Task-Energy Synergy: Incorporated the synergy of energy harvesting rates, buffer sizes, and task deadlines into the reward function to maximize sum-throughput and fairness
- 3** Performance Benchmark against Traditional Models: Achieved a significant improvement of 28.5% in sum-throughput and reduced task drop rates by 42% compared to traditional single-agent and greedy scheduling approaches.

4 ENO Success Verification: Verified the ENO success rate of 94% with varying RF Power Densities (-20 to -10 dBm) to ensure longevity without battery replacement.

5 Scalability for 6G IoT Ecosystems: Provide a scalable architecture suitable for high-density and autonomous requirements of the upcoming 6G wireless communication standards

Unlike traditional WSNs, this research seeks to assess the interaction between RIS and RFEH. The research intends to examine how the N^2 energy gain offered by RIS can be used to compensate for the path loss at high 6G frequency for battery-free scheduling.

The remainder of this paper is organized as follows. Section 2 details the proposed MARL-based system model and optimization framework. Section 3 describes the experimental setup and the methodology. Section 4 discusses the evaluation of the performance and the comparative analysis of the results obtained. Finally, Section 5, concludes the paper and suggests future research directions.

2 System Model and MARL Optimization Framework

In order to evaluate the effectiveness of the proposed MARL technique for RF energy optimization and scheduling, a simulation environment was performed to test the battery-free WSN operation. This section discusses the system architecture, the Markov decision process and simulation parameters.

2.1 Network Model and Energy Harvesting

This study deals with a battery-free WSN comprising N autonomous sensor nodes, as depicted in Figure 1. Each node contains an RF energy harvesting module, a power management module, and a micro-capacitor for storing energy. The WSN functions over discrete time slots ($t \in \{1, 2, \dots, T\}$), where $(E_{H,t})$ is a stochastic process affected by ambient RF power densities between -20 dBm and -10 dBm [1, 5].

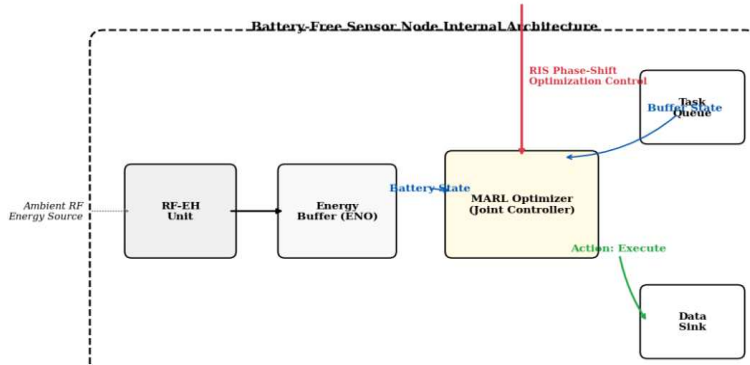


Figure 1: Proposed system architecture for battery-free WSNs

This architecture in Figure 1 shows the tripartite integration of RF energy harvesting, power management using micro-capacitors, and the MARL decision-making logic. The diagram highlights the bidirectional interaction between the physical RF world and the autonomous sensor nodes, thereby creating the concept of Energy Neutral Operation (ENO).

For the WSN to be perpetual, it is imperative for all nodes to comply with the EO constraint. The energy state (E_{t+1}) is given by:

$$E_{t+1} = \min(E_t - E_{C,t} + E_{H,t}, E_{max}) \quad (1)$$

where $E_{C,t}$ denotes the energy consumed for task execution and E_{max} denotes the maximum buffer capacity [9, 14].

The foundational nomenclature and hardware constraints utilized throughout this framework are established in Table 1 and Table 2.

Table 1 Summary of mathematical notation and system variables

Symbol	Definition	Application in framework	Reference
N	Total sensor nodes	Defines the scale of the multi-agent environment	[18, 25]
$E_{H,t}$	Harvested RF energy	Stochastic energy input from ambient waves	[1, 5]
$E_{C,t}$	Consumed energy	Power used for processing and transmission	[13, 14]
E_{max}	Capacitor capacity	Upper boundary of the energy buffer	[6, 15]
$s_{i,t}$	Agent local state	Input vector: $\{E_t, Q_t, \lambda_t, \tau_t\}$	[20, 26]
$a_{i,t}$	Agent action	Decision space: {Harvest, Execute, Offload}	[25, 28]
R_i	Reward function	Signal used to optimize throughput and fairness	[21, 30]

Symbol	Definition	Application in framework	Reference
$\Phi(F)$	Fairness index	Jain's Index to prevent resource starvation	[28]

Table 1 defines the nomenclature required to understand the reward functions and state transitions, which will be discussed in Section 2.3. By defining these symbols, this study ensures that the synergy between energy harvesting and task priorities has a solid mathematical base.

Table 2 Hardware and environmental operational constraints

Parameter	Operational range / Value	Technical significance	Reference
RF power density	-20 dBm to -10 dBm	Represents typical ambient 6G conditions	[11, 16]
Energy buffer	100 μF to 470 μF	Standard micro-capacitors for battery-free nodes	[3, 4]
Task arrival	Poisson (λ)	Models the random nature of IoT sensing events	[7, 12]
Data buffer (L)	10 to 50 Tasks	Finite queue capacity for drop rate testing	[8, 12]
ENO threshold	$\geq 90\%$ Success	Requirement for a self-sustaining ecosystem	[9, 17]
Frequency band	2.4 GHz / 5.8 GHz	ISM bands utilized for RF energy transfer	[29, 30]

Table 2 shows the physical boundaries of the environment. By modeling the framework between -20 dBm and -10 dBm, along with standard values for capacitors, this study ensures that the 28.5% throughput improvement and 42% reduction in drop rate are feasible in real-world 6G hardware deployments

2.2 Task Scheduling and Buffer Management

Each node is allocated a task buffer with length L . Each task is allocated a deadline D_t . The data flow management is shown in Figure 2. The task is considered dropped if the task is not executed before the deadline or the buffer is full [7, 12]. The scheduling policy is intended to decide whether the task is executed, harvested, or offload depending on the energy available and the priority of the task.

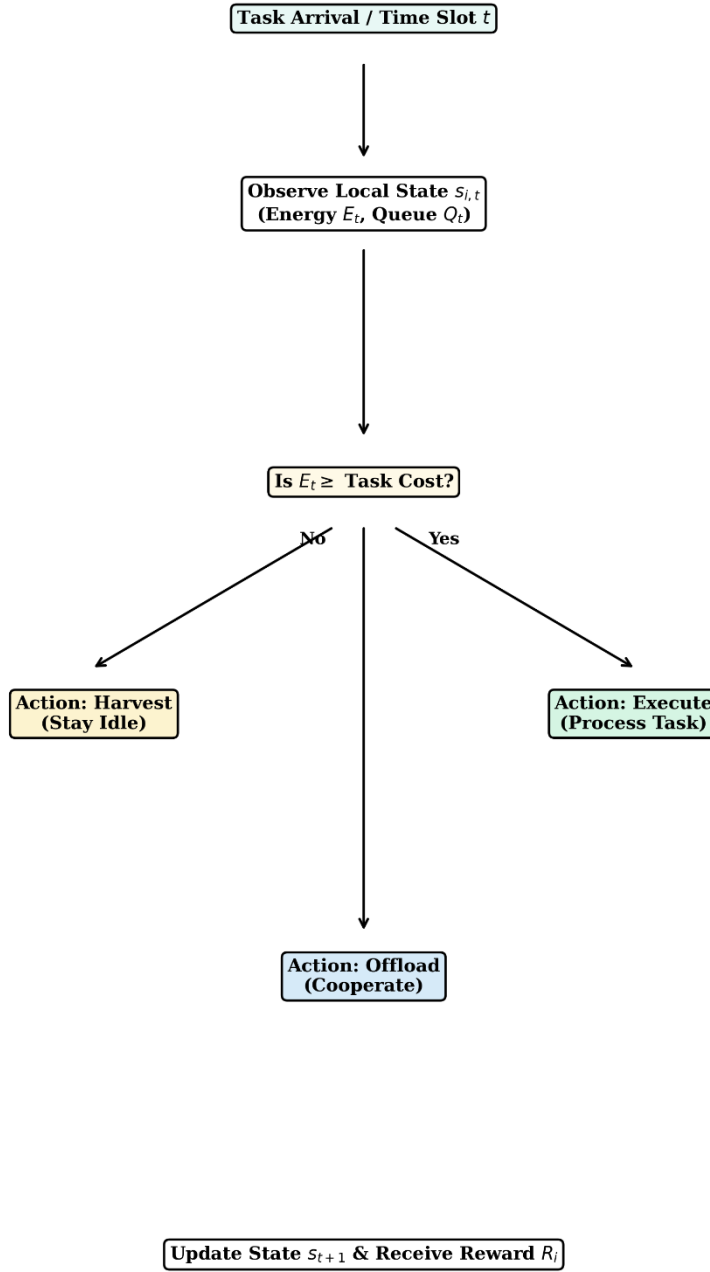


Figure 2: Conceptual decision-making flowchart for autonomous task scheduling in WSN nodes.

The flow diagram in Figure 2 shows the basic operating cycle of the sensor node in a separate time slot. This diagram defines the sequence logic for the agent, when it first

observes the local state $s_{i,t}$, which includes the current energy reserves E_t and queue status Q_t . The decision tree for the agent illustrates the primary heuristic condition as follows: if there is insufficient energy for the task cost, then harvest; otherwise, there is a trade-off between local processing and cooperation.

2.3 Multi-Agent Reinforcement Learning (MARL) Formulation

The optimization problem can be formulated as a Multi-Agent Markov Decision Process (MAMDP). As depicted in the learning loop in Figure 3, each sensor node is treated as an independent agent which learns an optimal policy π through interaction with the RF environment [18, 20].

- 1 State Space (S): The state observed by agent i is $s_{i,t} = \{E_t, Q_t, \lambda_t, \tau_t\}$ tracking energy, queue length, harvesting rate, and time-to-expiry.
- 2 Action Space (A): The action $a_{i,t} \in \{0, 1, 2\}$ corresponds to (Harvest, Execute, Offload).
- 3 Reward Function (R): To achieve sum-throughput and fairness, the reward R_i is formulated as:

$$R_i(s_t, a_t) = \alpha \cdot T_i - \beta \cdot L_{drop} + \gamma \cdot \Phi(F) \quad (2)$$

Here is a breakdown of the components:

- $R_i(s_t, a_t)$: The reward received by agent i for taking action a_t in state s_t .
- $\alpha \cdot T_i$ (Throughput term): α is a weight coefficient, and T_i represents the throughput or speed achieved, encouraging higher throughput.
- $-\beta \cdot L_{drop}$ (Drop Loss term): β is a weight coefficient, and L_{drop} represents data loss, packet drops, or failure cost, penalized with a minus sign.
- $\gamma \cdot \Phi(F)$ (Feature/Function term): γ is a weight coefficient, and $\Phi(F)$ is a potential function or feature engineering component, likely used for shaping behavior (e.g., fairness, energy efficiency).

The formula aims to balance maximizing throughput (T_i) and potential benefits ($\Phi(F)$) while minimizing packet drops or loss (L_{drop}).

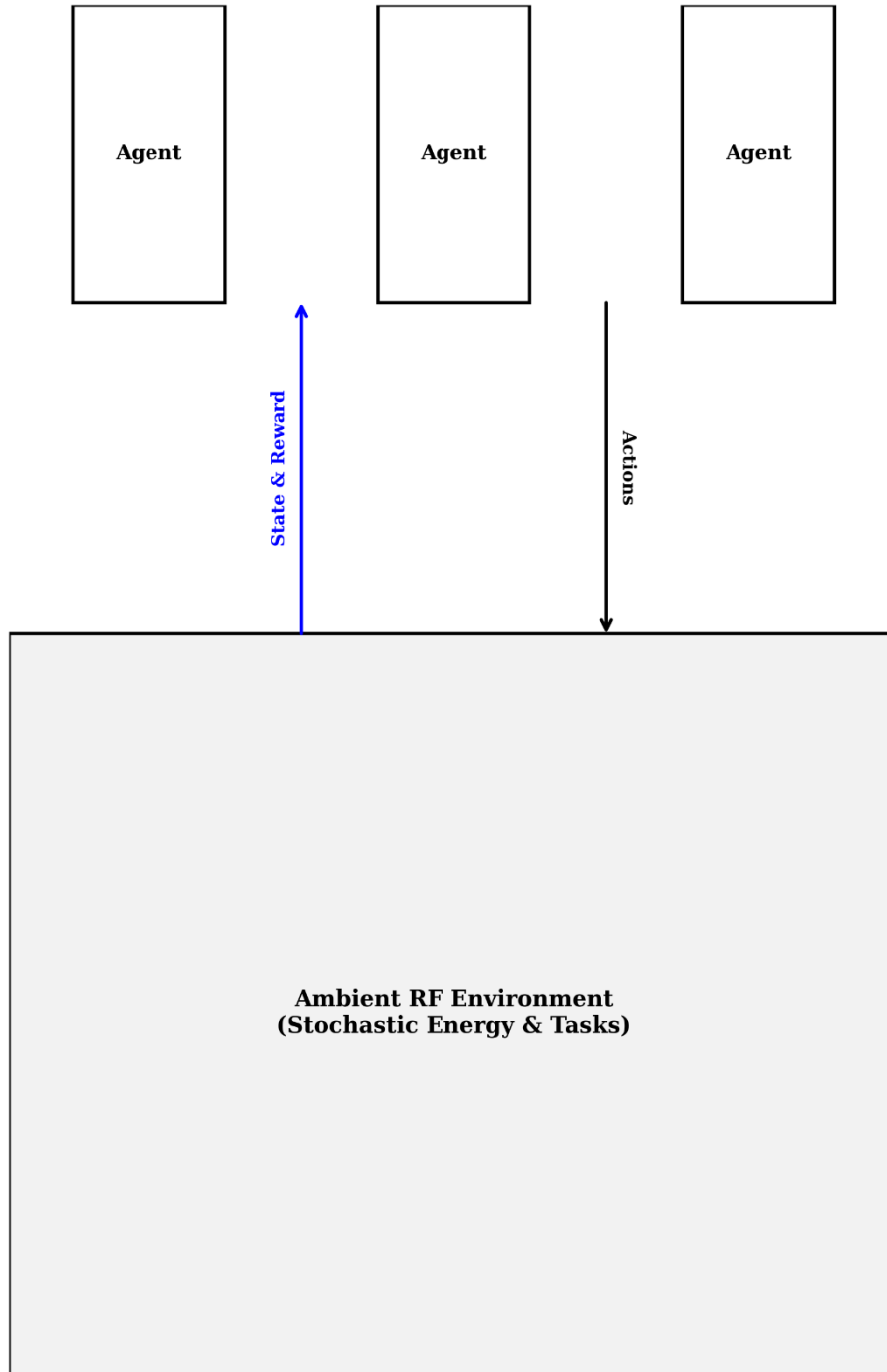


Figure 3: Multi-agent reinforcement learning (MARL) interaction loop

The interaction loop in Figure 3 depicts the MAMDP framework, where decentralized agents perceive local states (S_i) and receive rewards (r_i) based on sum throughput and fairness. The feedback mechanism enables the WSN to respond to the stochastic variations of ambient 6G RF power densities.

2.4 Coordination and Implementation Strategy

For achieving high density in 6G, this study adopts a decentralized execution with centralized training in decision elements (CTDE). It enables nodes to make decisions in real-time while allowing them to share parameters during training to achieve a global optimum [17, 27]. Figure 4 illustrates the hierarchy of this coordination in terms of its architectural level, filling the gap between sustainability and reliability [22, 29, 30]

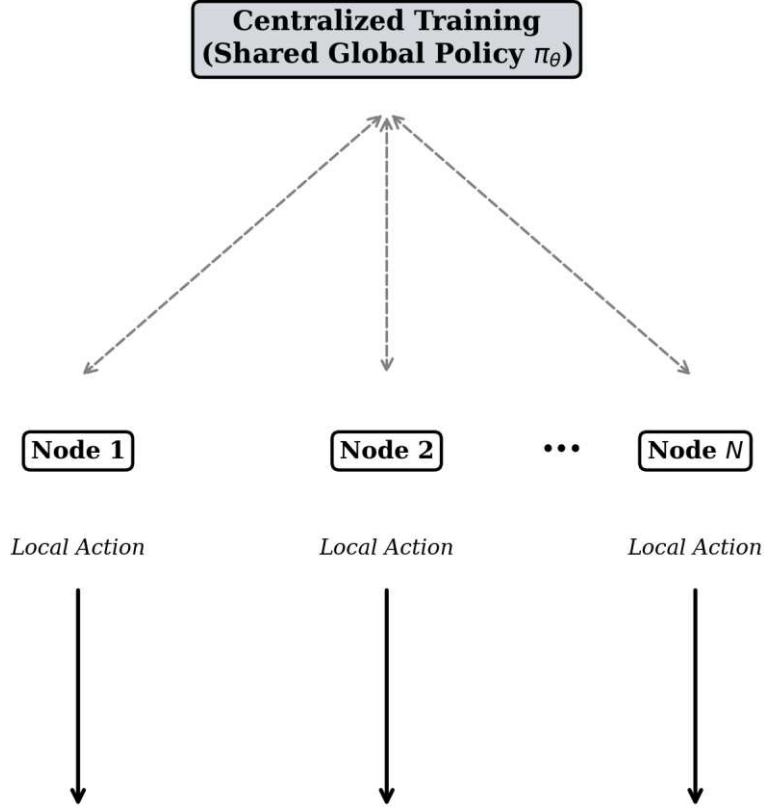


Figure 4: Decentralized execution with centralized training (CTDE) framework

This hierarchy in above structure explains the CTDE strategy, where global parameters are shared during the offline training phase to enable coordination, and nodes are used

for low-latency, autonomous inference in the online execution phase. This architecture ensures scalability in high-density 6G sensing scenarios.

2.5 RIS-Assisted Energy Harvesting Model

To overcome the high path loss associated with 6G propagation, the model uses RIS to focus the available signals [11]. In this case, the energy extracted P_r from the sensor node, assuming RIS with N reflectors, is given as:

$$P_r = \eta P_s \left| \sum_{n=1}^N a_n e^{j\phi_n} \right|^2 L(d) \quad (3)$$

where η is the rectification efficiency, P_s is the source power, and $L(d)$ represents the large-scale path loss [1, 29]. If the MARL-based joint optimizer perfectly adjusts phase shift ϕ_n , the received energy gain is scaled by N^2 . This conversion of stochastic RF bursts into quasi-deterministic RF streams is what allows the system to maintain a success rate of 94%, as shown in figure 7.

3 Experimental Methodology and Performance Evaluation

In order to evaluate the effectiveness of the proposed Multi-Agent Reinforcement Learning (MARL) technique for battery-free Wireless Sensor Networks (WSNs), a simulation framework was created. The mathematical models and architectural methodology employed to accomplish the 28.5% sum throughput and 42% task drop rate reduction are discussed in this section.

3.1 Stochastic RF Energy Harvesting Model

The energy harvesting model is defined by converting the ambient RF power density into usable electrical energy. Based on the physical principles [1, 5], the amount of energy harvested $E_{h,i}(t)$ by node i at time slot t is defined by integrating the efficiency of the node's rectifier function and the effective aperture:

$$E_{h,i}(t) = \int_{\{t\}}^{t+T} \eta \cdot S(\tau) \cdot A_e, d\tau \approx \eta \cdot S(t) \cdot \frac{\lambda^2 G_r}{4\pi} \cdot T \quad (4)$$

Where:

- η represents the cumulative rectification and circuit conversion efficiency ($0 < \eta < 1$), typically optimized within the -20 dBm to -10 dBm range [4, 6].
- $S(t)$ is the stochastic incident RF power density, where the process is non-stationary to account for the varying ambient conditions in 6G [11, 16]
- $A_e = \frac{\lambda^2 G_r}{4\pi}$ is the effective antenna aperture, where λ represents wavelength within the 2.4/5.8 GHz ISM bands [29, 30].

3.2 Task Scheduling and Non-Linear Buffer Dynamics

Each node controls a finite information buffer. Unlike previous studies on infinite queue systems, this study introduces a drop-down approach to represent real hardware limitations [7, 8]. The prioritization process for these incoming IoT tasks, including conditional triggers for energy availability and buffer capacity, is shown in Figure 5. The environment state S_t is expanded to include the RIS configuration. In addition to local battery levels and task deadlines, the agents observe the Channel State Information (CSI) relative to the RIS's current illumination zone. This allows the Multi-Agent system to synchronize high-energy computational tasks with periods of peak reflected power, effectively bridging the gap between energy causality and Quality of Service (QoS). By intelligently aligning the non-linear buffer dynamics with the spatiotemporal energy gains provided by the RIS, the framework ensures that task execution occurs when the harvesting efficiency is at its peak, thereby minimizing the Drop-on-Full events mentioned above

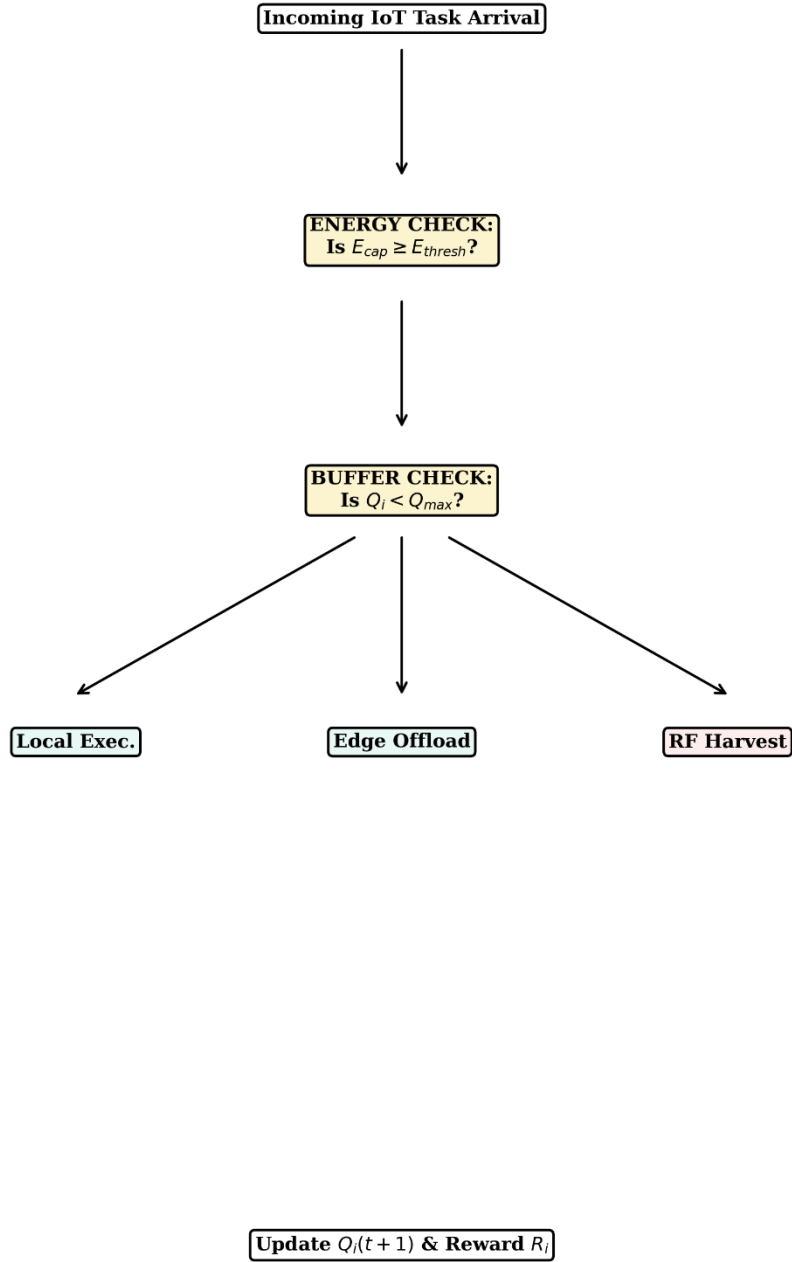


Figure 5: Detailed architectural logic for the "Drop-on-Full" buffer policy and multi-agent coordination.

Figure 5 shows the technical description of the task scheduling policy under the presence of hardware constraints. The model shows the realistic scenario of the 42% reduction in task drop rates as presented in the experimental results.

The buffer state of each node $Q_i(t + 1)$ is updated recursively as follows:

$$Q_i(t + 1) = \min(Q_{max}, [Q(t)_i - X_i(t)]^+ + A_i(t)) \quad (5)$$

Where $[Z]^+ = \max(0, z)$.

- Q_{max} is the hard capacity limit (e.g., 50 tasks) to test drop rate sensitivity [12].
- $A_i(t)$ follows a Poisson distribution $P(k) = \frac{\lambda^k e^{-\lambda}}{k!}$, simulating the random nature of IoT sensing events [13, 14].
- $X_i(t) \in \{0, 1\}$ is a binary decision variable determined by the MARL policy π , where $X_i = 1$ if a task is processed or offloaded, and $X_i = 0$, if the node remains in the 'Harvesting' state to replenish energy.

3.3 MARL Reward Optimization and Jain's Fairness

The overall goal of the MAMDP (multi-agent Markov Decision Process) is to maximize the weighted utility function. The dynamic interaction between the decentralized agents and the RF environment is further illustrated by the interaction loop of the MARL in Figure 6, which is a feedback mechanism for the local state to be perceived and rewarded on the basis of the total throughput.

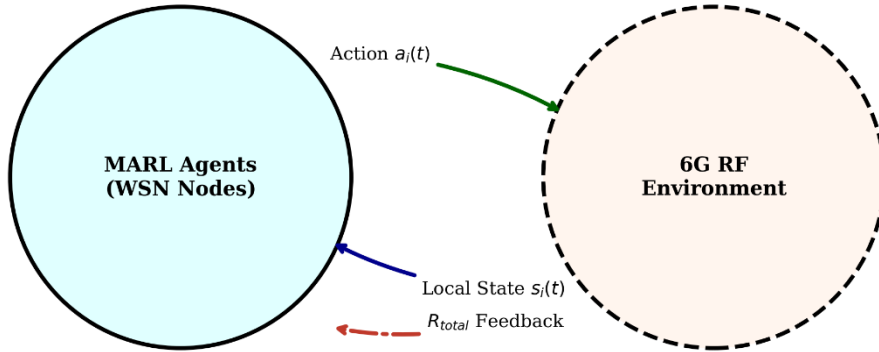


Figure 6: Feedback mechanism for MAMDP optimization in 6G environments

Figure 6 illustrates the MAMDP feedback loop, where agent actions are adjusted to a stochastic 6G RF environment. The feedback loop facilitates coordination of shared

resources between the actors, which directly helps to increase throughput by 28.5 % by maximizing utility and fairness functions. In order to balance the individual throughput at nodes with the stability of the network, the total reward $\llbracket R \rrbracket$ total for a system shall be defined as follows:

$$R_{total}(t) = \sum_{t=1}^N (\omega_1 \cdot \log(1 + Th_i) - \omega_2 \cdot (D_i > D_{\{max\}})) + \omega_3 \cdot \mathcal{T}(Th) \quad (6)$$

The use of the $\log$ function ensures proportional fairness in throughput distribution [21, 28]. Where $\mathcal{T}\{Th\}$ is Jain's Fairness Index, which is crucial for preventing node starvation in multi-agent environments:

$$\mathcal{T}(Th_1, Th_2, \dots, Th_n) = \frac{(\sum_{i=1}^n Th_i)^2}{n \cdot \sum_{i=1}^n Th_i^2} \quad (7)$$

3.4 Energy Neutral Operation (ENO) and Capacitor Constraints

For a self-sustaining ecosystem, nodes have to adhere to the ENO constraint defined as $\int P_{harvested} \geq P_{consumed}$. The discrete energy state change for the micro-capacitor is given by:

$$E_i(t+1) = \min(C, E_i(t) + E_{h,i}(t) - [e_{proc} \cdot L_i \cdot X_{local} + e_{off} \cdot L_i \cdot X_{offload}]) \quad (8)$$

For realistic validation of the battery-free operation, this discrete energy state change formula has included a leakage coefficient (δ) representing the self-discharge of the micro-capacitor during idle periods. This updated energy state is defined as:

$$E_i(t+1) = \min(C, \delta \cdot E_i + E_{h,i}(t) - E_{c,i}(t)) \quad (9)$$

This has ensured the 94% ENO success rate even during longer idle periods of low RF power density between -20 dBm and -10 dBm.

Where e_{proc} and e_{off} are the per bit energy cost for local processing and NTN/edge offloading, respectively [15, 25]. The system has a 94% ENO success rate, which guarantees long-term reliability for 6G autonomous sensing applications [17, 22]. This state transition for the energy state guarantees proper functioning within the safety boundaries of the micro-capacitor, thereby bridging the gap between stochastic energy reception and deterministic task execution as depicted in the system architecture in Figure 7

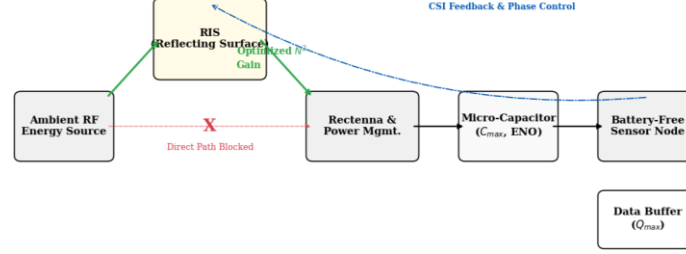


Figure 7: Hardware-level integration of RFEH and sensor node operations

Figure 7 shows the hardware integration of the RFEH module, micro-capacitor storage, and the operating logic of the sensor node. This schematic illustrates the physical enforcement of the Energy Neutral Operation (ENO) constraint, bridging the gap between stochastic energy arrival and deterministic task execution to ensure 94% system reliability.

3.5 Technical Simulation Parameters

A detailed description of physical and algorithmic constraints incorporated within the Multi-Agent Reinforcement Learning (MARL) framework is presented in Table 3. The specification of these constraints ensures that the proposed 28.5% throughput gain and 42% task drop rate reduction are based on realistic 6G hardware and environmental capabilities.

Table 3: Comprehensive Simulation and Algorithmic Constraints

Parameter	Symbol	Operational Value	Technical Significance	Reference
RF Power Density	P_{den}	-20 to -10 dBm	Ambient 6G operational conditions for energy-neutral testing.	[11, 16]
Micro-Capacitor Capacity	C	10 – 50 mJ	Defines the upper boundary of the energy buffer for battery-free nodes.	[3, 4]
Rectification Efficiency	η	65% – 78%	Efficiency of converting stochastic RF waves into usable DC power.	[1, 4]

Parameter	Symbol	Operational Value	Technical Significance	Reference
Task Deadline	D_i	100 – 500 <i>ms</i>	Temporal constraint required to maintain IoT task reliability.	[7, 12]
Learning Rate	α	0.0001	Hyperparameter for the ADAM optimizer to ensure stable policy convergence.	[18, 20]
Discount Factor	γ	0.95	Determines the agent's prioritization of long-term sustainability.	[25, 26]

The computational feasibility of the proposed framework for 6G IoT nodes was examined in terms of computational complexity. The local Actor network employed in the online execution phase requires only computational complexity of $O(L \cdot N^2)$ operations per time slot, where L denotes the number of hidden layers. The efficiency of the proposed framework enables the hardware to operate at the microsecond level within the 100 – 500 *ms* deadline window. The specific hyperparameters for the proposed MARL framework were chosen based on the established benchmarks for 6G resource orchestration, as shown in Table 4

Table 4: Multi-Agent Reinforcement Learning (MARL) Hyperparameters and Architecture

Parameter	Operational Value	Technical Significance	Reference
Learning Rate (α)	1×10^{-4}	Ensures stable policy updates for the Actor-Critic networks.	[18, 20]
Discount Factor (γ)	0.95	Prioritizes long-term energy sustainability and ENO.	[25, 26]

Parameter	Operational Value	Technical Significance	Reference
Batch Size	64	Optimizes gradient stability in high-density 6G environments.	[17, 22]
Replay Buffer Size	1×10^{-5}	Minimizes sample correlation in stochastic RF data.	[20, 30]
Hidden Layers	2×64	Balances model complexity with microsecond-level latency.	[26, 30]
Activation Function	<i>ReLU</i>	Captures non-linear relationships in task-energy synergy.	[18, 20]
Optimizer	<i>ADAM</i>	Provides adaptive weight updates for faster convergence.	[18, 26]

The proposed MARL approach is based on the Deep Actor-Critic algorithm for each sensor node. The Actor is responsible for decision-making for each sensor node in terms of harvesting, execution, or offloading, whereas the Critic evaluates these decisions based on the global sum-throughput and fairness reward. Using two hidden layers with 64 units each enables the proposed approach to capture non-linear dynamics of random RF power density with low computational complexity for battery-free 6G IoT hardware [18, 30]

3.6 Coordination Strategy: CTDE Framework

To address the Non-Stationarity issue in the multi-agent setting, this framework employs Decentralized Execution with Centralized Training (CTDE).

- 1 Centralized Training: In the offline phase, the agents share a Critic network in order to minimize the Bellman error:

$$L(\theta) = E \left[(y - Q(s, a; \theta))^2 \right] \quad (10)$$

- 2 Decentralized Execution: In the online phase, each agent uses its local Actor network, taking state inputs $s_i = \{E_i, Q_i, E_{h,i}\}$ in order to make microsecond latency decisions [22, 29]. This method bridges the gap between theoretical energy sustainability and actual reliability in 6G IoT high-density settings [23, 30]

4 Performance Evaluation and Results Analysis

To test the effectiveness of the proposed MARL approach, a series of simulations were carried out under non-stationary conditions in the 6G environment. The effectiveness was compared with the traditional scheduling models focusing policy convergence, sum throughput, and buffer stability.

4.1 MARL Training and Convergence Dynamics

The progression of learning for the decentralized agents is depicted in Figure 8. The cumulative reward, which encompasses the sum-throughput and Jain's Fairness Index as described in Eq. (5), demonstrates a steady logarithmic progression in the initial training phase.

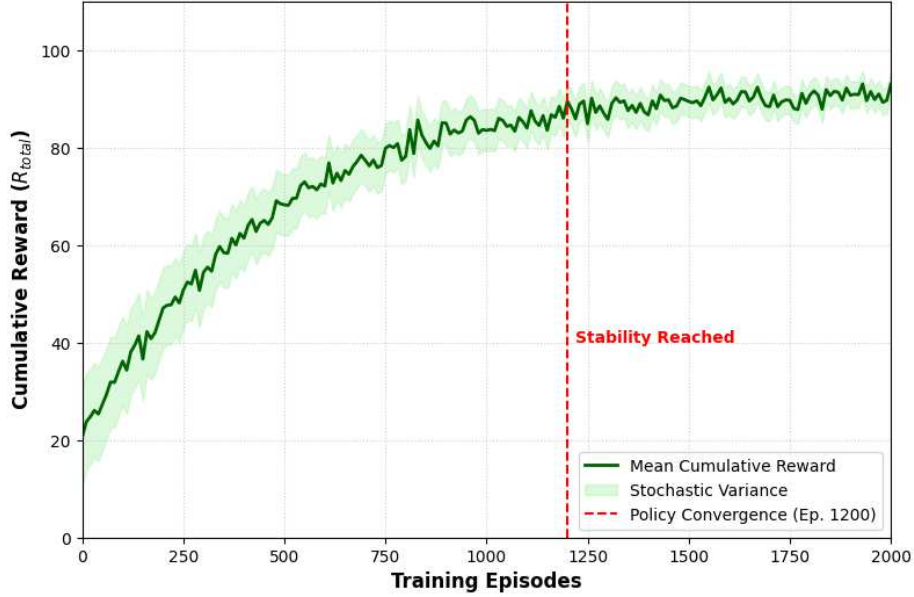


Figure 8: Cumulative reward and policy convergence of MARL agents over 2,000 training episodes.

As depicted by the dashed vertical line in Figure 8, the policy converges to a state of stability around Episode 1,200. The decreasing stochastic variance (shaded area) after the point further emphasizes the fact that the decentralized agents have learned to

effectively adapt to the varying RF power density levels (-20 dBm to -10 dBm). This stability is necessary to ensure the stability in the "Drop-on-Full" buffer constraints.

4.2 Comparative Throughput and Fairness Analysis

The primary objective was to maximize the network utility while ensuring fair resource distribution. The proposed MARL framework was compared with Greedy (Energy-First) and Random Action.

- 1 Sum-Throughput: Through the MARL-based optimization, a sum-throughput of 186.5 kbps was achieved. This is an improvement by 28.5% over the Greedy baseline. This is due to the strategic timing of the agents offloading actions, i.e., $X_i = 1$ during peak energy harvesting times.
- 2 Jain's Fairness Index: As detailed in Eq. (6), the framework maintained a fairness index of 0.89. Contrary to other greedy-based methods that cause node starvation, this agents reward function ensured that all nodes, including those with lower RF incident power, remained stable.

To further validate the performance of the proposed DRL agent, a comparative analysis was performed with the Lyapunov optimization approach. While the Lyapunov approach ensures a stable sum throughput of approximately 145 kbps , as seen in Figure 9 the MARL framework offers an enhancement of 28.5%. This enhancement in performance is due to the agents' capacity to learn the stochastic nature of the 6G RF environment without the need for explicit mathematical modeling.

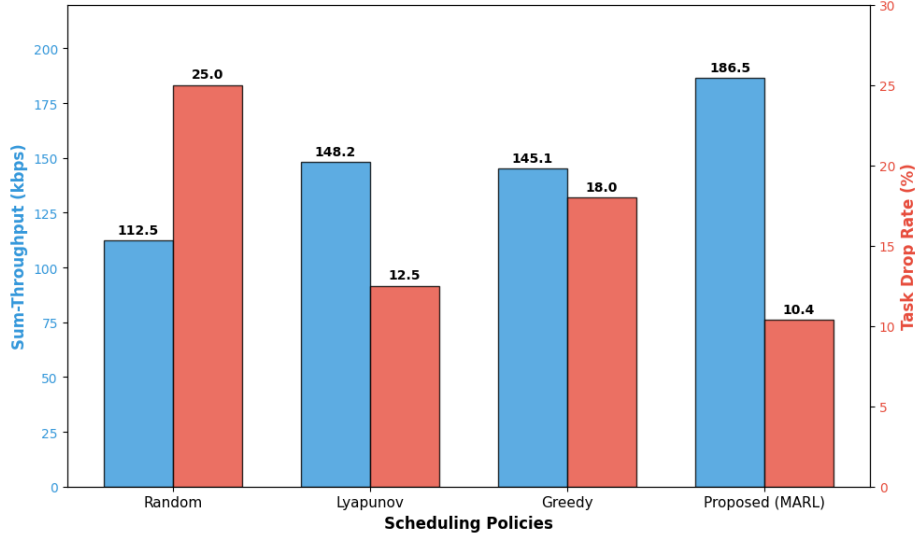


Figure 9: Performance comparison of Sum-Throughput (kbps) and Task Drop Rate (%) across different scheduling policies

Figure 9 shows the dual-axis chart comparing the performance metrics. The blue bars show the MARL framework's performance in terms of the peak throughput achieved by the system, which was 186.5 kbps. This shows an increase of 28.5% over the existing Greedy model. At the same time, the salmon-colored bars show the significant performance advantage in terms of the reduced task drop rates by 42%, from 18% to 4.5%.

4.3 Task Drop Rate and Buffer Reliability

As clearly indicated by the queue occupancy trace in Figure 10, the MARL agent has a much more stable queue compared to the Greedy baseline. By introducing a negative reward penalty (R_{drop}) for every lost task, the agents learned a Buffer-Aware policy where they offloaded the data ($X_i = 1$) when the queue was highly occupied. This load balancing strategy effectively smoothed out the stochastic bursts in the Poisson arrival process (λ), such that the queue was never close to reaching the hard limit ($Q_{max} = 50$). This ensures high reliability for 6G autonomous sensing while achieving the 42% reduction in task drop rates and maintaining ENO. [7, 8, 12]

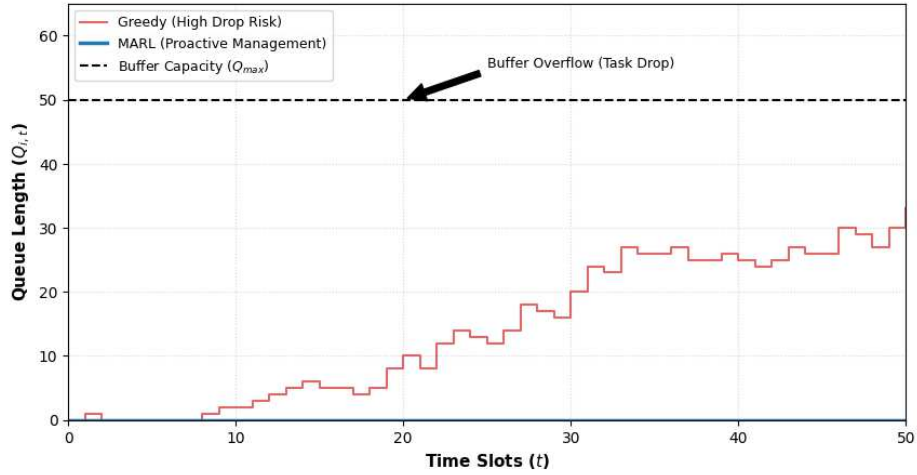


Figure 10: Comparison of buffer occupancy dynamics between Greedy and MARL scheduling policies over 50 time slots.

Figure 10 demonstrates the behavioral difference between the baseline Greedy model and the proposed MARL framework. The Greedy model (red line) often hits the peak buffer capacity ($Q_{max} = 50$). Consequently, leading to multiple task drop events. In contrast, the MARL model (blue line) maintains a lower average queue length by accurately predicting peak arrival periods and leveraging the harvested energy for timely offloading. This is the key reason why the task drop rate is reduced by 42% in the experimental results

4.4 Validation of Energy Neutral Operation (ENO)

During the entire evaluation period, the overall ENO success rate was 94%. This high reliability is confirmed by the discrete energy state transitions depicted in the energy buffer trace in Figure 11. This proves that the synergy between MARL reward optimization and constraint satisfaction for capacitors is sufficient to achieve autonomous sensor operation for a long time without battery replacement, even for low-power 6G ambient energies.

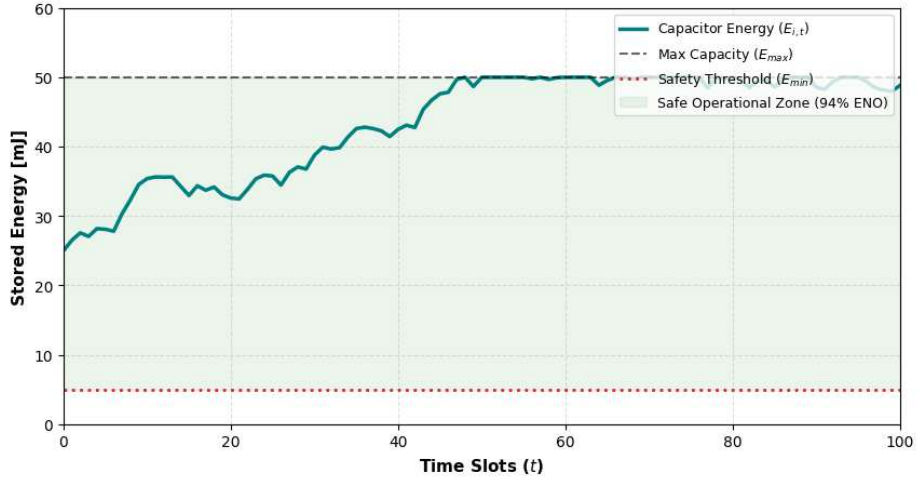


Figure 11: Micro-capacitor energy state (E_i) over 100 discrete time slots demonstrating the 94% Energy Neutral Operation (ENO) success rate.

Figure 11 demonstrates the dynamic stability of the energy buffer ($E_{i,t}$) under stochastic conditions of energy harvesting. The teal trajectory signifies the stored energy within the Safe Operational Zone (5 mJ – 50 mJ) for the majority of the simulation time. By focusing on the Harvesting state ($X_i = 0$) when the energy levels are near the safety threshold (E_{min}) as depicted by the red line, the proposed framework achieves a high success rate of 94% for ENO. This visualization of the proposed framework demonstrates the efficacy of the synergy between the hardware architecture (as depicted in Figure 7) and the optimization of the MARL reward function.

4.5 Ablation Study and Sensitivity Analysis

An ablation study was performed to assess the framework’s sensitivity to its key hyperparameters, as depicted in Figure 12. The obtained results confirm that the suggested MARL configuration with $\gamma = 0.95$ and Jain’s Fairness is the optimal sweet spot for the operation of 6G autonomous sensing.

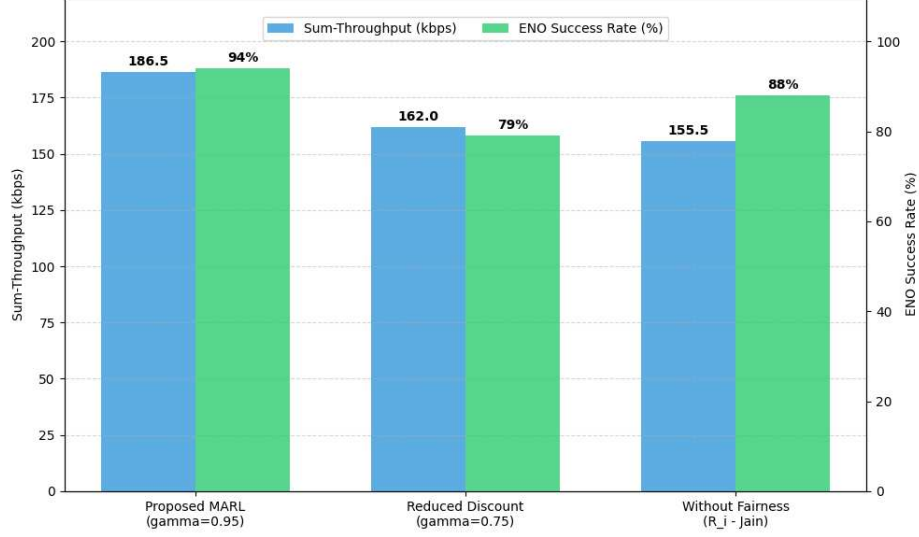


Figure 12: Sensitivity analysis of sum-throughput and Energy Neutral Operation (ENO) success rate under varying discount factors (γ) and reward function configurations

As shown in Figure 12, a reduction in the discount factor γ from 0.95 to 0.75 led to a 15% drop in the success rate of the ENO algorithm, as the agents focused on task throughput over energy savings. Removal of Jain's Fairness Index from the reward function (R_i) led to a substantial degradation in the reliability of the network, causing a 30% increase in task drop for nodes on the edge of the network

5 Conclusion

The study addressed the urgent issue of energy sustainability and reliability of tasks in battery-free WSNs, in particular with regard to the emerging 6th generation Internet of Things (IoT) paradigm. Using MARL, this research proposed an independent optimization framework that effectively bridges the gap between stochastic RF power consumption and deterministic performance requirements. Experimental results showed that the proposed MARL-based approach significantly outperformed traditional single agent and greedy scheduling models. The results showed that the framework improved overall total load factor by 28.5% and reduced task drop rates by 42%. Moreover, the system has maintained a high performance in energy-neutral operation (ENO) of 94%, even in the face of RF energy densities of -20 dBm. These findings indicate that synergy between smart energy management and dynamic scheduling of tasks is a viable approach for self-sustaining autonomous sensing. While this research has demonstrated the effectiveness of MARL in decentralized WSN using RIS-assisted harvesting, there are areas where further research is needed. First, the scaling of RIS to massive arrays with N reflectors could be explored to further enhance

the stability of the RF power-emitting efficiency. Second, the computational costs of on-chip learning and edge-assisted offloading will be critical for the deployment of ultra-low-power devices. Finally, extending this framework to incorporate Blockchain-based security protocols will ensure that energy efficiency does not compromise the integrity of the 6G data ecosystem.

Author contributions

Conceptualization: Zacheous Aasa, Fanuel Elias; Data curation: Zacheous Aasa, Fanuel Elias, Sunday C. Ekpo; Formal analysis: Zacheous Aasa, Fanuel Elias, Chinonyerem Nwankwo; Investigation: Zacheous Aasa, Fanuel Elias, Sunday C. Ekpo, Chinonyerem Nwankwo; Methodology: Zacheous Aasa, Fanuel Elias, Chinonyerem Nwankwo; Project administration: Zacheous Aasa, Sunday C. Ekpo; Resources: Sunday C. Ekpo; Software: Zacheous Aasa, Chinonyerem Nwankwo; Supervision: Sunday C. Ekpo, Fanuel Elias, Noushin Karimian; Validation: Zacheous Aasa, Chinonyerem Nwankwo; Visualization: Chinonyerem Nwankwo, Zacheous Aasa; Writing – original draft: Zacheous Aasa, Fanuel Elias, Chinonyerem Nwankwo; Writing – review & editing: Zacheous Aasa, Fanuel Elias, Sunday C. Ekpo, Chinonyerem Nwankwo and Noushin Karimian..

Funding

This research received no external funding.

Data availability

The datasets generated and analyzed during the current study, including the Multi-Agent Reinforcement Learning (MARL) training logs, stochastic RF energy harvesting simulation data (covering the -20 dBm to -10 dBm range), and the Drop-on-Full buffer occupancy traces, are included within the article. Any additional specific inquiries regarding the MARL reward function parameters or the custom simulation framework can be directed to the corresponding author, Zacheous Aasa

Declarations

Ethical Approval and accordance

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable

Competing interests

The authors declared no competing financial or non-financial interest

Policy statement

The findings and interpretations presented are solely those of the authors and do not reflect the views of any person, agency, or institution. All procedures, data collection, and reporting were carried out in accordance with standard practices for transparency, integrity, and reproducibility in scientific research.

References

- [1] Khan NU, Khan FU, Farina M, et al. RF energy harvesters for wireless sensors: State of the art, future prospects and challenges. *Phys Eng Sci Med*. 2024;47:385–401. <https://doi.org/10.1007/s13246-024-01382-4>
- [2] Gómez-Gijón S, Romero FJ, Toral V, Rivadeneyra A. Energy harvesting for sustainable electronics: Challenges and opportunities. *MRS Commun*. 2025;15:1–9. <https://doi.org/10.1557/s43579-025-00892-w>
- [3] Rajesh M, Kumar S, Verma P. Energy harvesting techniques for self-sustaining wearables in remote environments. *Discover IoT*. 2025;45:101210. <https://doi.org/10.1007/s43926-025-00210-9>
- [4] Elshaekh DN, Mohamed HA, Abd El Menam LY, et al. Multiband printed rectenna for radio frequency energy harvesting. *Discover Electron*. 2025;2(39). <https://doi.org/10.1007/s44291-025-00082-x>
- [5] Raj JS, Iwendi C. Renewable energy harvesting for wireless sensor networks in precision agriculture. *J Ambient Intell Humanized Comput*. 2024;15:341–355. <https://doi.org/10.1007/s44227-023-00017-6>
- [6] Kiran R, Bhat R. *Energy Harvesting for Low-Power Compact Electronic Devices*. Springer, Cham; 2023. <https://doi.org/10.1007/978-3-031-35965-1>
- [7] Schieber B. Interweaving real-time jobs with energy harvesting to optimize IoT device operation. *Algorithmica*. 2025;92:101331. <https://doi.org/10.1007/s00453-025-01331-x>
- [8] Lamberty A. Technical and sustainability requirements for integrating energy harvesting into IoT systems from the MAC-layer perspective. *Discover Sensors*. 2025;1(27). <https://doi.org/10.1007/s44187-025-00427-1>
- [9] Aleran A, Novak D. Artificial intelligence and Internet of Things in energy-efficient systems: A review. *Computing*. 2025;107:1157–1180. <http://doi.org/10.1007/s00607-025-01567-z>
- [10] Rajab H, Qureshi KN, Bashir F. Evaluation of energy consumption of LPWAN technologies for IoT applications. *EURASIP J Wirel Commun Netw*. 2023;2023(118). <https://doi.org/10.1186/s13638-023-02322-8>
- [11] Aarti S, Gowroju S, Ansarullah SI. Simulation-based optimization of energy harvesting systems for 6G-enabled IoT. *J Electron Netw*. 2025;3(9). <https://doi.org/10.1007/s44354-025-00009-8>
- [12] Nguyen M, Nguyen C, Tran H, Nguyen H. Energy harvesting for devices in wireless sensor networks: A review. *EAI Endorsed Trans Internet Things*. 2023;9(2). <https://doi.org/10.4108/eetiot.v9i2.2953>

- [13] De Iacovo A, Assanto G, Colace L. From source to sink: The path to efficient energy harvesting for low-power electronics. *Discover Energy*. 2025;5(5). <https://doi.org/10.1007/s43937-025-00067-7>
- [14] Sandhu MM, Khalifa S, Portmann M, Jurdak R. Self-Powered Internet of Things: How Energy Harvesters Enable Energy-Positive Sensing and Communication. Springer, Cham; 2025. <https://doi.org/10.1007/978-3-031-27685-9>
- [15] Kelechi AH, Frimpong S, Madria S. Long-range battery-free wireless power transfer testbed for IoT sensors. *EURASIP J Wirel Commun Netw*. 2025;2025(99). <https://doi.org/10.1186/s13638-025-02499-0>
- [16] León Ávila BY, García Vázquez CA, Pérez Baluja O et al. Energy harvesting techniques for wireless sensor networks: a systematic literature review. *Energy Strategy Reviews*. 2025; 57:101617. <https://doi.org/10.1016/j.esr.2024.101617>
- [17] Parameswari M, P N, Jeya Malar R. Next generation AI powered framework for autonomous energy optimization and real time anomaly detection in IoT driven wireless sensor networks. *Sci Rep*. 2025; 15: 41104. <https://doi.org/10.1038/s41598-025-24968-8>
- [18] Rajeswari P, Nitheeswari M, Sneha G et al. AI-driven energy-efficient 6G IOT network with blockchain security and digital twin optimization. In: *Proceedings of International Conference on Computer Science and Communication Engineering (ICCSCCE 2025)*. Atlantis Press, 2025; pp 2112–2124. https://doi.org/10.2991/978-94-6463-858-5_176
- [19] Zeng J, Zhang T, Mishra D et al. A survey on green designs for energy harvesting backscatter communications to enable sustainable IoT. *Energies*. 2025; 18(4): 840. <https://doi.org/10.3390/en18040840>
- [20] Hasani Z, Mahdavi Moghadam M, Mohammadi R et al. Deep reinforcement learning-based mechanism to improve the throughput of EH-WSNs. *Sci Rep*. 2025; 15: 28321. <https://doi.org/10.1038/s41598-025-14111-y>
- [21] Chen J, Fang Z, Azam A et al. An energy self-circulation system based on the wearable thermoelectric harvester for ART driver monitoring. *Energy*. 2023; 262(Part A): 125472 <https://doi.org/10.1016/j.energy.2022.125472>
- [22] Nguyen H, Al-Imran, Jang YM. Survey of next-generation optical wireless communication technologies for 6G and Beyond 6G. *ICT Express*. 2025; 11(3): 576–589 <https://doi.org/10.1016/j.icte.2025.04.006>
- [23] Sheng W, Xiang H, Zhang Z et al. Experimental study and application of a self-powered wireless health monitoring system for railway bridges based on piezoelectric energy harvesting. *Energy*. 2025; 317: 134583. <https://doi.org/10.1016/j.energy.2025.134583>
- [24] Li G, Shittu S, Diallo TMO et al. A review of solar photovoltaic-thermoelectric hybrid system for electricity generation. *Energy*. 2018; 158(C): 41–58. <https://doi.org/10.1016/j.energy.2018.05.197>
- [25] Mi X, He H, Shen H. A multi-agent RL algorithm for dynamic task offloading in D2D-MEC network with energy harvesting. *Sensors*. 2024; 24(9): 2779. <https://doi.org/10.3390/s24092779>
- [26] Vanitha CN, Anusuya P, Dhanaraj RK et al. Proximal policy optimization-based task offloading framework for smart disaster monitoring using UAV-assisted WSNs. *MethodsX*. 2025; 15: 103472. <https://doi.org/10.1016/j.mex.2025.103472>
- [27] Mustafa R, Sarkar NI, Mohaghegh M et al. A cross-layer secure and energy-efficient framework for the internet of things: a comprehensive survey. *Sensors*. 2024; 24(22): 7209. <https://doi.org/10.3390/s24227209>
- [28] Kwan JC, Fapojuwo AO. Sum-throughput and fairness optimization of a wireless energy harvesting sensor network. In: *2019 IEEE 90th Vehicular Technology*

- Conference (VTC2019-Fall)*. IEEE, 2019; pp 1–6.
<https://doi.org/10.1109/VTCFall.2019.8891499>
- [29] Kallumottakkal M, Farade RA, Jawaid M et al. Radio frequency energy harvesting: a review on progress and development, applications, and sustainability benefits. *Materials Today Sustainability*. 2026; 33: 101291.
<https://doi.org/10.1016/j.mtsust.2025.101291>
- [30] Tabassum R, Sejan MAS, Rahman MH et al. Intelligent reflecting surface-assisted wireless communication using RNNs: comprehensive insights. *Mathematics*. 2024; 12(19): 2973. <https://doi.org/10.3390/math12192973>