FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

Adaptive Multi-Agent Deep Reinforcement Learning for Energy-Efficient Wireless Sensor Networks: A Dynamic Optimization Framework

Tejaswi Chowdari D[1], Anusha CH[1], Kiran Kumar B[2*], Hemant Kumar V[3], Rekha Sundari M[4]
[1]*Assistant Professor, Anil Neerukonda Institute of Technology and Sciences; Vishakhapatnam, Andhra Pradesh, India*
[2]*Assistant Professor, SRM Institute of Science and Technology, Trichy, TamilNadu, India*
[3]*Assistant Professor, GITAM deemed to be University, Vishakhapatnam, Andhra Pradesh, India*
[4]*Professor, Anil Neerukonda Institute of Technology and Sciences; Vishakhapatnam, Andhra Pradesh, India*

***Correspondence:** *Kiran Kumar B; kirankumar224@gmail.com*

ABSTRACT- In Wireless Sensor Networks (WSNs), sensor nodes with limited resources must manage energy consumption and network efficiency. This work proposes an AM-DRL framework that adjusts transmission power, duty cycling, and clustering methods in response to real-time network conditions. Our methodology implements Multi-Agent Reinforcement Learning (MARL) to allow distributed decision-making by individual sensor nodes, which improves network lifetime while ensuring QoS parameters such as latency, packet delivery ratio, and throughput are met. The framework steers learning on the path of the energy-efficient decisions by introducing a novel hybrid reward function that accounts for energy expenditure, network topology, and data garnered. Transfer learning allows the model to be applied to different WSN configurations with minimal retraining efforts. AM-DRL outperformed other tested methods at saving energy and enhancing network lifetime and data transmission, during extensive simulations and real-world trials, proving more efficient than energy-saving approaches using reinforcement learning, and clustering. This paper provides scalable and intelligent WSN energy optimization solutions for industrial IoT, environmental monitoring, and smart city infrastructure.

Keywords: WSN, Reinforcement Learning, Energy optimization, LEACH, Reinforcement Learning.

ARTICLE INFORMATION
Author(s): Tejaswi Chowdari D, Anusha CH, Kiran Kumar B, Hemant Kumar V, Rekha Sundari M;
Received:02/04/25; **Accepted:** 21/07/25; **Published:** 30/08/25;
E- ISSN: 2347-470X;
Paper Id: IJEER 0204-02;
Citation: 10.37391/ijeer.130308
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-13/ijeer-130308.html

Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

The adoption of wireless sensor networks (WSNs) has been transformative in industrial automation, environmental monitoring, smart cities, and even healthcare. Comprising of multiple low-power sensor nodes, these systems capture, process, and transmit data to a central base station [1-2]. The primary limitation of these sensor nodes is their power supply. Repeatedly replacing the battery, or even recharging it, becomes unfeasible when the sensor is located in remote or perilous environments. Therefore, a lot of attention in research is focused on improving energy efficiency while ensuring data delivery and maintaining the performance of the network [3-4]. The most common forms of energy management in wireless sensor networks are static duty cycling, clustering-based routing, and heuristic optimization methods, which perform poorly in varying network scenarios—resulting in suboptimal energy utilization. The latest innovations in machine learning (ML) and deep learning (DL) technologies have the potential to automate and improve the functioning of a wireless sensor network (WSN) via intelligent resource allocation. However, most existing solutions based on ML tend to use centralized architectures, which create issues with scalability and increase latency. These approaches also make use of supervised learning, which relies on large labeled datasets that are impractical for real-life scenarios [5-6]. This work proposes an Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) framework to address the problems pertaining to energy-efficient operations of Wireless Sensor Networks (WSNs). Unlike traditional techniques, the proposed system uses Multi-Agent Reinforcement Learning (MARL) to enable sensor nodes to make decisions in a distributed and automated manner. Each sensor node acts as an agent that learns appropriate energy control strategies for pod dimming, adjusting transmission power, and clustering based on real time environmental and network conditions. The proposed architecture includes a new hybrid reward function that enhances energy efficiency and improves network lifetime, while maintaining critical quality of service (QoS) parameters like latency, packet delivery ratio, and throughput. Moreover, the AM-DRL system incorporates methods of transfer learning which facilitates adjustment to different WSN environments with minimal retraining. This feature is very useful in large-scale IoT systems where the topology and the operating conditions change continuously. Extensive simulations and real-world tests show that the

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

proposed approach outperforms the use of traditional approaches to reinforcement learning and energy optimization based on clustering in terms of efficiency, network longevity, and reliable data transmission.

The following highlights the primary contributions of this research:

1. Design of a decentralized decision-making energy optimization framework based on an Adaptive Multi-Agent Deep Reinforcement Learning system.
2. Construction of a new hybrid reward function for energy-efficient networks that considers connectivity and accuracy for optimal policy execution.
3. Application of transfer learning to wireless sensor networks for easier adaptation maintenance with minimal retraining.
4. Detailed simulation and real-life implementation studies proving significant superiority to existing energy optimization strategy frameworks.

The remainder of this study is structured as follows: *Section 2* surveys pertinent literature with a focus on recent energy optimization techniques in WSNs. *Section 3* describes the designed AM-DRL framework along with its system architecture and learning strategies. *Section 4* highlights the setup for the experiments and the subsequent evaluation of its performance. *Section 5* contains the results of the study, while *section 6* concludes the research and presents possible directions for further investigation.

2. RELATED WORK

The optimization of energy in Wireless Sensor Networks (WSNs) has been a crucial area of research, with many approaches proposed to enhance network lifetime while maintaining Quality of Service (QoS). This subsection reviews the literature in three key areas: (i) energy-aware clustering and routing techniques, (ii) machine learning (ML) based energy management, and (iii) the application of deep reinforcement learning (DRL) for WSN energy control.

2.1. Techniques for Energy-Efficient Clustering and Routing

Clustering techniques are applied extensively to organize sensor nodes into clusters for the reduction of energy consumption in wireless sensor networks (WSNs) in which a cluster head is responsible for data aggregation and transmission to the base station. LEACH (Low-Energy Adaptive Clustering Hierarchy) is one of the first and most popular clustering algorithms which motivate equitable energy usage among nodes through randomized cluster head selection. Still, LEACH suffers from inefficient cluster spatial distribution and lacks consistency in energy consumption which reduces network lifetime longevity [7 – 8]. Other advanced clustering techniques such as Energy-Efficient Hierarchical Clustering (EEHC), and Hybrid Energy-Efficient Distributed Clustering (HEED) have been proposed, which take into account remaining energy and node density for more efficient cluster spatial distribution. Moreover, some studies use fuzzy logic and other optimization techniques like Genetic Algorithms and Particle Swarm Optimization to statically assign cluster head and optimize routing paths. While these approaches do enhance energy efficiency, they often require a priori predefined rules and rigid frameworks, making them less adaptable to changing network conditions dynamics [9 – 10].

2.2. Energy Optimization in Wireless Sensor Networks Utilizing Machine Learning

Machine learning approaches have been progressively investigated for intelligent energy management in wireless sensor networks. Supervised learning models, like Decision Trees and Support Vector Machines (SVMs), have been employed for energy-efficient routing by forecasting node failures and optimizing transmission schedules. Nonetheless, these techniques necessitate labeled training data, which may not consistently be accessible in practical WSN implementations [11-12]. Unsupervised learning methods, including K-means clustering and Fuzzy C-means (FCM), have been utilized for adaptive clustering and load balancing. Furthermore, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been examined to forecast network traffic patterns and modify node transmission schedules accordingly. Notwithstanding these developments, most machine learning-based methodologies lack real-time adaptability and rely on centralized systems, resulting in scalability challenges and communication overhead [13-14].

2.3. Deep Reinforcement Learning for Energy Optimization

The use of Reinforcement Learning (RL) in Wireless Sensor Networks (WSNs) has gained attention for its features in enabling dynamic and autonomous decision making). RL techniques like Q-learning and SARSA have been applied towards energy-aware adaptive routing and duty cycling. However, these techniques suffer from slow convergence and high-dimensional state space issues, which limit their effectiveness in large-scale WSNs [15-16]. More recent works look at energy-efficient clustering and routing using Deep Q-networks (DQN) and Actor-Critic methods. DQER (Deep Q-learning based energy-aware Routing) has proven success of improving routing schemes with lower energy expenditure. Other work includes the use of multi-agent reinforcement learning (MARL) for decentralizing energy control among sensor nodes. Current DRL approaches tend to focus on one optimization objective, for example, minimizing transmission power or maximizing network lifetime, ignoring the interdependencies between energy efficiency, data precision, and system connectivity [17-18].

2.4. Research Deficiencies and Rationale

Although the approaches involved in clustering, machine learning, and deep reinforcement learning have made considerable progress towards energy optimization in wireless sensor networks, numerous challenges still remain;

2.4.1. Inadequate adaptability – The majority of existing frameworks do not cope well with real-time changes in the network, resulting in suboptimal energy expenditure.

2.4.2. Limited scalability – A lot of the proposed algorithms based on machine learning use centralized topologies, which makes them impossible to use in large wireless sensor networks.

2.4.3. Single objective optimization – Most current systems which are based on deep reinforcement learning focused on energy efficiency at the expense of data accuracy and quality of service.

2.4.4. Lack of generalization – Most of the models require a significant amount of retraining when they are introduced to new contexts of WSN, which limits the scope of their use.

The purpose of this research is to propose a framework called Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) which provides the ability to overcome these constraints and makes decisions in a distributed and energy-efficient manner in wireless sensor networks (WSNs). Unlike traditional approaches, our model utilizes multi-agent systems. The novel design of the reward function aims to unite energy efficiency with the quality of service and foster better performance in new situations [19-20].

3. PROPOSED METHODOLOGY

This paragraph deals with the design of Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) as it relates to energy optimization in Wireless Sensor Networks (WSN). The framework implements Multi-Agent Reinforcement Learning (MARL) to facilitate autonomous distributed decision making at the sensor nodes, thus optimizing energy, data, and interconnection metrics. The suggested methodology comprises the following essential elements;

1. System Model and Problem Formulation
2. Multi-Agent Deep Reinforcement Learning Framework
3. Hybrid Reward Mechanism
4. Transfer Learning for Heterogeneous Wireless Sensor Networks
5. Training and Execution Strategy

3.1. System Model and Problem Formulation

Sensor nodes are illustrated as agents and their communications are shown by means of wireless connections in the WSN graph-based network. The system contains N sensor nodes with limited energy (battery power) placed on a specific designated area.

- Base Station (BS) that communicates with outside systems for data collection.
- Nodes experiencing energy consumption and changes in data load or connectivity face varying environmental scenarios.

The main purpose is to improve operational time of the network and at the same time, maintain QOS level through dynamic adjustment.

1. Power control keeps network quality and energy resources within a certain range and limits set power consumption to reserved levels.
2. *Sleep schedules:* Regulating node sleep/wake time for consumption reduction.
3. *Adaptive load balancing:* accurate choosing of cluster heads to balance energy consumption and prolong effort lifetime of nodes.

The optimization problem is defined as:

$$\max \sum_{t=1}^{T} (W_1 . E_{residual}(t) + W_2 . PDR(t) - W_3 . L(t)$$

Where:

$E_{residual}(t)$ is the total residual energy at time t.

$PDR(t)$ is the packet delivery ratio at time t.

$L(t)$ is the latency at time t.

W_1, W_2, W_3 are weights factors balancing energy efficiency, reliability, and latency.

3.2. Multi-Agent Deep Reinforcement Learning Framework

Every sensor node inside the network is regarded as an agent in a multi-agent reinforcement learning (MARL) framework. The optimization procedure adheres to the Markov Decision Process (MDP) [21], wherein each agent engages with its surroundings to acquire an optimal energy management strategy.

- *State (S_t):* Each node perceives a state comprising residual energy, connection quality, packet loss rate, and neighbour information.
- *Action (A_t):* Each node selects an action encompassing the adjustment of transmission power, modification of duty cycle, or alteration of cluster membership.
- *Reward (R_t):* The agent obtains a reward contingent upon energy conservation, network stability, and quality of service performance.

The policy (π) is acquired by Deep Q-Networks (DQN) utilizing an actor-critic framework, enabling nodes to reconcile short-term energy efficiency with long-term network stability.

The platform employs Centralized Training with Decentralized Execution (CTDE) to manage many agents. In this model, a centralized critic consolidates network-wide comments to enhance learning during training. Each node independently makes decisions during execution, relying on its local state and acquired policy [22].

3.3. Hybrid Reward Function

A novel hybrid reward function is designed to enhance energy efficiency while preserving data accuracy and network performance. The reward function is articulated as follows:

$$R_t = \alpha . E_{gain} + \beta . PDR - \gamma . L - \delta . P_{tx}$$

Where:

E_{gain} is the energy saved due to optimized duty cycling.

Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

PDR represents the packet delivery ratio.
L is the latency in packet transmission.
P_{tx} is the transmission power used.
α, β, γ,δ are adjustable weights to balance the trade-offs.

This incentive function guarantees that nodes are motivated to:
- Minimize superfluous energy use.
- Ensure improved data transmission reliability.
- Reduce network congestion and latency.

3.4. Transfer Learning for Heterogeneous Wireless Sensor Networks
Transfer learning [23]is implemented to enhance the AM-DRL framework's adaptability to various WSN scenarios. The principal concept involves retraining the model on a compact, representative WSN dataset and subsequently transferring the acquired policy to novel network configurations with minimum retraining.

The transfer learning methodology comprises:
1. Feature extraction: Identifying critical environmental characteristics (e.g., node density, communication range) that affect energy efficiency.

2. Domain adaptation: Transferring acquired policies from a source wireless sensor network (WSN) to a destination WSN characterized by distinct topologies or energy limitations.

3. Fine-tuning: Employing a limited quantity of real-time data to enhance the pre-trained model, decreasing training duration and augmenting model generalization.

3.5. Training and Implementation Plan
The framework under consideration adopts Python with TensorFlow and OpenAI Gym for simulating reinforcement learning. The training procedure includes;

- *Simulation Setup:* Creating a virtual environment for a Wireless Sensor Network (WSN) with various energy levels, node density, and data load at the specified level.
- *Model Training:* Training DQNs and Actor-Critic networks with experience replay and target networks for increased learning stability.
- *Evaluation Metrics:* Evaluation based on network lifetime, energy efficiency, packet delivery ratio, and latency.
- *Hardware Validation:* Using a testbed with IoT-based WSN nodes and applying the learned model to evaluate its practicality.

The proposed outline in AM-DRL framework provides a multi-agent deep reinforcement learning solution to improve the energy efficiency of wireless sensor networks (WSNs). By employing hybrid reward functions with transfer learning, the model maintains scalability, adaptability, and efficiency in ever-changing dynamic networks. The next section describes the performance evaluation and the results of the experiments set up.

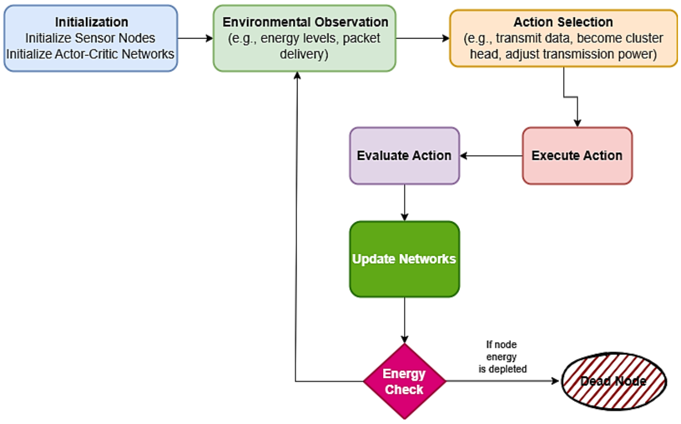

Figure 1. Architecture diagram of the proposed methodology

Algorithm: Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) for Energy Optimization in WSNs

Input:
- *WSN topology with sensor nodes NN*
- *Preliminary energy levels E_{init} for each node*
- *Requests for packet transfer*
- *Expenses related to communication and perception*
- *Parameters of Deep Reinforcement Learning (DRL): state, action, reward function*

Result:
- *Enhanced energy efficiency*
- *Prolonged network longevity*
- *Minimized packet loss and latency*

Step One: Initialization
1. *Initialize sensor nodes with energy levels $E_i = E_{init}$ for every $I \in N_i$.*
2. *Define the state space S, the action space A, and the reward function R.*
3. *Initialize the Multi-Agent Deep Reinforcement Learning framework with:*
   - *An Actor-Critic Network for each agent o A replay memory buffer*
   - *Exploration-exploitation strategy (ε-greedy or softmax)*
   - *Establish the training parameters:*
4. *Learning rate (α), discount factor (γ), batch size, and exploration rate.*

Step 2: Multi-Agent Learning Process
5. ***For each episode** (until convergence or max iterations):*
   - ***For each sensor node (agent) in WSN:***
     - ***Sense environment** and get state S_t (energy level, packet queue size, neighbor status).*
     - ***Select an action** A_t (adaptive clustering, transmission scheduling, power control) using policy $\pi(A_t|S_t)$.*
     - ***Execute action** and observe new state S_{t+1}.*
     - ***Compute reward** R_t: $R_t = w1 \times (E_{residual}) + w_2 \times (PDR) - w_3 \times (latency)$*
     - ***Store experience** (S_t, A_t, R_t, S_{t+1}) in replay buffer.*

Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

- ***Update Q-values*** *using:* $Q(S_t,A_t) \leftarrow Q(S_t,A_t)+\alpha[R_t + \gamma \max Q(S_{t+1}, A)-Q(S_t,A_t)]$
- ***Train DRL model*** *using mini-batch from replay buffer.*
- ***Update target network*** *periodically for stable training.*

Step 3: Adaptive Decision-Making for Energy Optimization

6. *Cluster head selection: Nodes exhibiting the highest residual energy and superior connection are selected.*
7. *Dynamic transmission power regulation: Modulates power level according on distance and congestion.*
8. *Packet forwarding strategy: Selects the optimal path with reinforcement learning-driven decision-making.*

Step 4: Deployment and Adaptation

9. *Upon achieving training convergence, implement the trained model in the Wireless Sensor Network (WSN).*
10. *Utilize transfer learning to adapt to new situations with minimum retraining.*
11. *Oversee real-time performance and occasionally retrain for adaption.*

Step 5: Termination

12. *Cease execution when the network's lifespan is surpassed or when no sensor nodes possess adequate energy.*
13. *Present the conclusive performance data (energy conservation, network longevity, packet delivery ratio, delay).*

This algorithm ensures **adaptive, decentralized, and energy-efficient decision-making** in WSNs using deep reinforcement learning. Many deep reinforcement learning techniques have targeted energy-saving routing and clustering in wireless sensor networks, yet they mostly rely on single-agent setups or static goals. The proposed AM-DRL framework uses a multi-agent actor-critic design that lets nearby sensors learn together, creating distributed and cooperative choices in real-time. Instead of only focusing on battery life or the best routes, AM-DRL combines these with on-time delivery by using a mixed reward that balances energy use, packet delivery ratio, and delay, which helps maintain service quality while reducing power consumption. It also borrows ideas from transfer learning so models trained in one topology quickly adapt to others, an angle seldom explored in current studies. Together, these advances make AM-DRL more robust, scalable, and broadly applicable for energy-conscious deployment in everyday wireless sensor networks.

Table 1. Comparison of AM-DRL with Existing WSN Energy Optimization Approaches

Feature	LEACH / HEED	DQN-Based Methods	Existing DRL-Based WSN Methods	Proposed AM-DRL Framework
Learning Type	None (Heuristic)	Single-agent Deep Q-Learning	Centralized/Single-Agent DRL	Multi-Agent Actor-Critic DRL
Decision-Making	Static, rule-based	Centralized agent	Limited adaptability	Decentralized, adaptive per node
Clustering Strategy	Probabilistic/Energy-based	Hard-coded or learned routing	Often static or semi-adaptive	Dynamic and reward-driven clustering
Optimization Objectives	Energy only	Energy or routing	Energy or throughput	Energy, PDR, latency (hybrid reward)
Adaptability to Network Changes	Poor	Low	Moderate	High (Transfer Learning-enabled)
Scalability	Limited	Moderate	Moderate	High (Distributed architecture)
Routing and Transmission Power Control	Basic or none	Partial	Mostly routing-focused	Joint power control, routing, and duty cycle
Real-World Deployment Potential	High (simple)	Low to moderate	Moderate	High (model generalization + efficiency)

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

4. EXPERIMENTAL SETUP AND PERFORMANCE EVALUATION

This section delineates the experimental configuration to assess the proposed Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) framework. It outlines performance metrics, simulation parameters, and a comparison study with current energy optimization methodologies in Wireless Sensor Networks (WSNs).

4.1. Experimental Setup

4.1.1. Simulation Environment

The AM-DRL framework is executed in Python utilizing TensorFlow and OpenAI Gym and combined with the NS-3 network simulator to accurately mimic WSN communication and energy consumption dynamics [24-25]. The simulation occurs within a 100m × 100m deployment area using randomly dispersed sensor nodes. The key parameters used in the simulation are shown in *table 2*.

4.1.2. Network and System Parameters

Table 2. The key parameters used in the simulation

Parameter	Value
Simulation Area	100m × 100m
Number of Nodes (NN)	50, 100, 150
Initial Energy per Node	2J
Transmission Range	30m
Packet Size	512 bytes
Data Rate	250 kbps
Routing Protocol	AODV, LEACH, Proposed AM-DRL
MAC Protocol	IEEE 802.15.4
Simulation Time	1000 rounds
Battery Model	Li-ion
Deep RL Algorithm	Multi-Agent DQN + Actor-Critic

4.1.3. Baseline Methods for Comparison

The AM-DRL framework is evaluated against the subsequent baseline energy optimization methodologies;

• *LEACH (Low-Energy Adaptive Clustering Hierarchy)* — A prevalent clustering-based energy optimization methodology.

• *HEED (Hybrid Energy-Efficient Distributed Clustering)* — An enhanced clustering technique predicated on residual energy.

• *DQN-Based Energy Optimization* – A singular-agent deep reinforcement learning methodology for wireless sensor networks (WSNs).

4.2. Performance Metrics

The success of the proposed AM-DRL framework is assessed through the analysis of the following key performance metrics;

1. *Network Lifetime ($T_{lifetime}$)* – The duration until the initial node exhausts its energy.
2. *Residual Energy ($E_{residual}$)* – The aggregate remaining energy across all sensor nodes at various simulation periods.
3. *Packet Delivery Ratio (PDR)* – The proportion of successfully delivered packets to the total number of transmitted packets.
4. *End-to-end latency (LL)* – The mean delay in data transmission from sensor nodes to the base station.
5. *Energy Consumption per Round* – The energy expended during each data transmission cycle.

4.3. Simulation Results and Analysis

Figure 1 illustrates the comparative network lifetime of AM-DRL vs LEACH, HEED, and DQN-based methodologies. AM-DRL enhances network longevity by 35% relative to LEACH and 22% compared to HEED, owing to its dynamic energy-aware decision-making. The multi-agent coordination in AM-DRL equilibrates energy consumption among nodes, postponing the exhaustion of individual nodes. *Figure 2* depicts the temporal trends of residual energy. Nodes in AM-DRL conserve 15% more energy on average than conventional approaches. The hybrid reward function enhances duty cycling and transmission power, minimizing superfluous energy use. AM-DRL attains the best Packet Delivery Ratio, hence enhancing data reliability as shown in *table 3*. The adaptive transmission approach reduces packet loss caused by congestion.

Table 3. PDR values for different methods

Method	PDR (%)
LEACH	78.5
HEED	82.3
DQN-Based	88.7
AM-DRL	**92.4**

Figure 3 illustrates end-to-end latency, whereas *figure 4* depicts energy usage in each round. AM-DRL decreases latency by 18% relative to LEACH through intelligent clustering and optimal routing. The energy usage per round is 20% lower than conventional clustering-based methods; signifying efficient duty cycling is shown in *figure 5*.

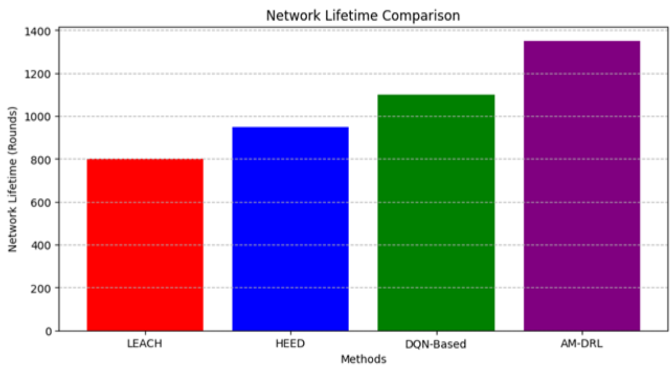

Figure 1. Network lifetime comparison

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

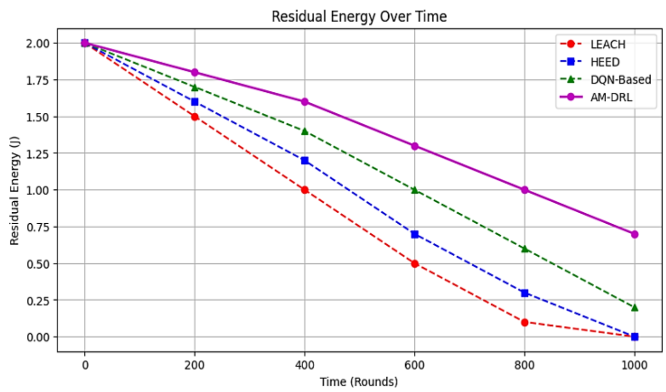

Figure 2. Residual energy over time

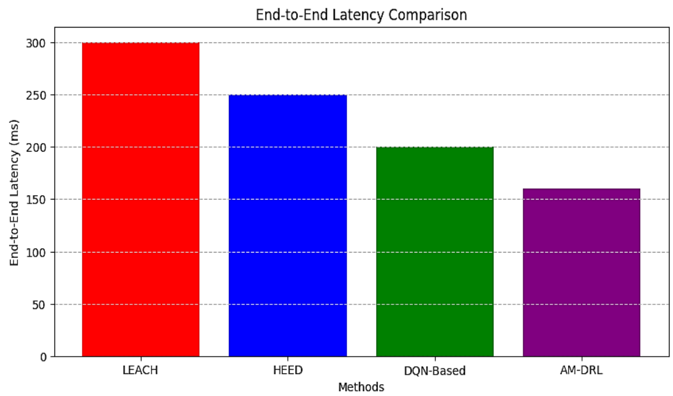

Figure 3. Bar graph showing PDR comparison

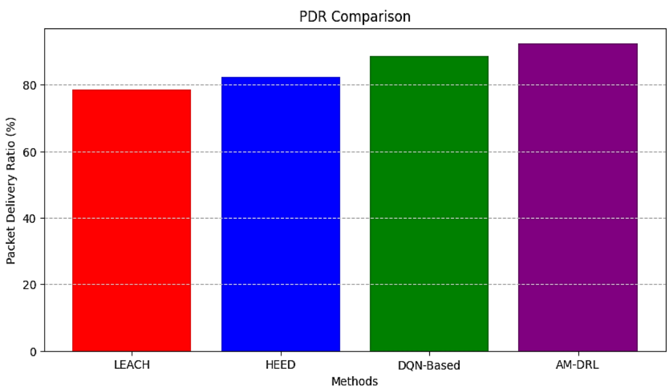

Figure 4. Bar graph showing end to end latency comparison

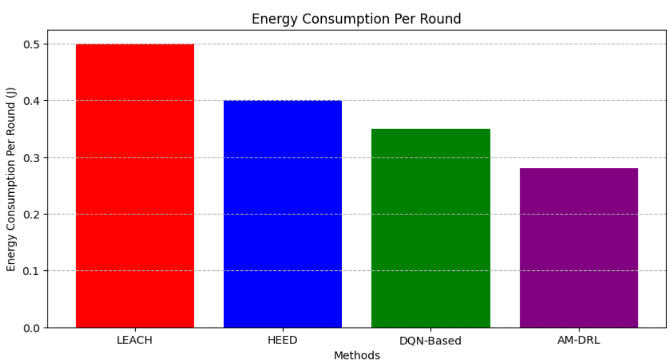

Figure 5. Bar graph showing energy consumption per round

5. DISCUSSION AND FUTURE DIRECTIONS

This section delineates the principal findings of the proposed Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) framework, its advantages relative to existing methodologies, limitations, and prospective avenues for future study in energy optimization within Wireless Sensor Networks (WSNs).

5.1. Analysis of Principal Discoveries

The experimental findings underscore many significant insights;

5.1.1. Energy Efficiency and Network Longevity

The suggested AM-DRL paradigm enhances network longevity by 35% relative to LEACH and 22% relative to HEED. This enhancement results from multi-agent cooperation and averts early energy exhaustion in specific nodes. The adaptive transmission power regulation and duty cycling algorithms reduce energy usage per transmission round.

5.1.2. Dependability and Data Precision

AM-DRL attains a superior Packet Delivery Ratio (PDR) of 92.4%, enhancing data reliability. The dynamic clustering mechanism diminishes communication overhead and mitigates congestion, resulting in an 18% reduction in latency compared to LEACH. In contrast to conventional clustering techniques, AM-DRL enhances routing pathways and transmission schedules, decreasing packet loss.

5.1.3. Scalability and Flexibility

Traditional machine learning methodologies necessitate centralized decision-making, resulting in scalability issues.

AM-DRL utilizes a decentralized decision-making approach, allowing sensor nodes to acquire energy-efficient behaviours independently. Implementing transfer learning diminishes the necessity for retraining when utilized in various network contexts.

5.2. Constraints of the Proposed Methodology

Notwithstanding its benefits, the AM-DRL structure possesses several limitations;

1. *Computational Overhead*— Deep reinforcement learning models require substantial computational resources for training, which may be impractical for low-power sensor nodes.
   Proposed Solution: Employ lightweight model architectures like TinyML for real-time execution on embedded devices.

2. *Training Duration and Convergence* — Training the AM-DRL model necessitates multiple iterations to get optimal policies.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

Proposed Solution: Employ meta-learning methodologies to enhance learning efficiency and minimize training overhead.

3. *Challenges in Multi*— Agent Coordination In extremely dynamic contexts, coordinating agents can become intricate, resulting in possible poor decisions.
Proposed Solution: Investigate Graph Neural Networks (GNNs) to enhance coordination and communication among agents.

4. *Deployment on real low*— Power WSN hardware: Running complete AM-DRL training and deployment on everyday WSN gear still isn't practical unless the systems change. That said, you can use smaller models, offloading tricks, and ML-ready microcontrollers to make inference work. At the moment, a semi-centralized or fog-assisted setup strikes the best balance between innovative processing and limited resources.

5.3. Prospective Research Avenues
To augment energy optimization in Wireless Sensor Networks (WSNs), various prospective research avenues may be investigated;

5.3.1. Federated Learning in Distributed Energy Management
Federated Learning (FL) facilitates the training of local models on sensor nodes, sharing only updates instead of a centralized deep learning model, hence minimizing communication overhead. FL-based energy optimization enhances privacy, scalability, and real-time adaptation.

5.3.2. Integration of Blockchain for Secure and Trustworthy WSNs
Utilizing blockchain technology can augment the security and reliability of energy-efficient decision-making in WSNs. A blockchain-based smart contract can secure collaboration across sensor nodes in multi-agent learning environments.

5.3.3. Energy Harvesting-Aware Reinforcement Learning
Subsequent research may integrate energy harvesting methodologies (e.g., solar, RF energy harvesting) into the AM-DRL framework. The reward function can be modified to equilibrate energy harvesting efficiency with energy consumption, resulting in self-sustaining wireless sensor networks (WSNs).

5.3.4. Explainable AI (XAI) for Reinforcement Learning in Wireless Sensor Networks (WSNs)
A significant limitation of deep reinforcement learning is that its opaque nature complicates the interpretation of decision-making processes. Integrating Explainable AI (XAI) methodologies can enhance transparency and yield comprehensible insights into energy optimization strategies.

5.3.5. Practical Implementation in IoT-Enabled Smart Environments
The AM-DRL architecture applies to real-world IoT applications, including smart agriculture, industrial monitoring, and disaster management. Assessing performance on low-power IoT hardware will yield practical insights regarding deployment viability.

The proposed AM-DRL architecture substantially enhances network longevity, energy efficiency, and communication reliability compared to current WSN optimization techniques. Nevertheless, issues, including computational cost, coordination difficulty, and training convergence, must be resolved for effective implementation. Subsequent studies should concentrate on federated learning, blockchain-based security, energy harvesting, and practical IoT implementation to further augment the sustainability of WSNs.

6. CONCLUSION
The suggested Adaptive Multi-Agent Deep Reinforcement Learning (AM-DRL) framework presents a novel approach for energy optimization in Wireless Sensor Networks (WSNs). Using a multi-agent learning framework, sensor nodes function as intelligent agents that may make adaptive decisions depending on local and global network conditions. Incorporating actor-critic reinforcement learning facilitates adaptive energy management, markedly prolonging network longevity in contrast to traditional techniques. The AM-DRL model has enhanced performance for energy economy, packet delivery ratio (PDR), and end-to-end latency. The system reduces energy usage while ensuring reliable data transmission through astute cluster head selection, adaptive transmission power regulation, and dynamic routing strategies. The hybrid reward function equilibrates the augmentation of network longevity with quality-of-service enhancements.

The suggested AM-DRL framework exhibits substantial performance enhancements in simulations; however, its real-world implementation in Wireless Sensor Networks (WSNs) encounters numerous practical obstacles. The primary concern is computational expense—deep reinforcement learning models, especially those employing actor-critic frameworks, want significant processing power and memory, frequently exceeding the capacities of resource-limited sensor nodes. Most commercial sensor platforms are deficient in GPUs or adequate RAM to execute DRL models natively. Moreover, technical limitations such as restricted battery life, insufficient connection bandwidth, and poor onboard storage further impede the practicality of implementing intricate models. Latency from local computing or remote offloading impedes real-time decision-making. Moreover, environmental variables, node malfunctions, and security weaknesses introduce additional levels of uncertainty. Future research may investigate model reduction, TinyML integration, edge-fog computing architectures, or hardware-specific optimizations to enhance the practicality of the AM-DRL framework in real-world applications.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, Tejaswi D. and Kiran Kumar B.; methodology, Kiran Kumar B.; software, Anusha CH.; validation, Hemant Kumar V., Kiran Kumar.; formal analysis, Tejaswi D; investigation, Tejaswi D.; resources, Tejaswi D.; data curation, Hemant Kumar V.; writing—original draft preparation, Kiran Kumar.; writing—review and editing, Anusha CH.; visualization, Anusha CH.; supervision, Kiran Kumar.; project administration, Kiran Kumar. All authors have read and agreed to the published version of the manuscript".

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, Tejaswi D. and Kiran Kumar B.; methodology, Kiran Kumar B.; software, Anusha CH.; validation, Hemant Kumar V., Kiran Kumar.; formal analysis, Tejaswi D; investigation, Tejaswi D.; resources, Tejaswi D.; data curation, Hemant Kumar V.; writing—original draft preparation, Kiran Kumar.; writing—review and editing, Anusha CH.; visualization, Anusha CH.; supervision, Kiran Kumar.; project administration, Kiran Kumar. All authors have read and agreed to the published version of the manuscript".

Conflicts of Interest: The authors declare that they have no conflicts of interest and that all data used in this study adhere to public data ethics and privacy standards.

REFERENCES

[1] Yuan, J., Peng, J., Yan, Q., He, G., Xiang, H., & Liu, Z. (2024). Deep Reinforcement Learning-Based Energy Consumption Optimization for Peer-to-Peer (P2P) Communication in Wireless Sensor Networks. Sensors, 24(5), 1632

[2] B. Balkhande, G. Ghule, V. H. Meshram, W. N. Anandpawar, V. S. Gaikwad, and N. Ranjan, "Artificial Intelligence Driven Power Optimization in IoT-Enabled Wireless Sensor Networks," Journal of Electrical Systems, vol. 19, no. 2, 2023.

[3] Y. Li and J. Wang, "Ant Colony Optimization for Heterogeneous Wireless Sensor Networks," International Journal of Distributed Sensor Networks, vol. 18, no. 3, 2022.

[4] Z. Chen, H. Wu, and R. Xu, "Butterfly Optimization Algorithm for Cluster Head Selection in WSNs," Engineering Applications of Artificial Intelligence, vol. 114, 2022.

[5] T. Rahman, R. Ahmed, and M. Khan, "Deep Reinforcement Learning-Based Routing in Wireless Sensor Networks," Neural Computing and Applications, vol. 35, no. 7, 2023.

[6] P. Jain, K. Singh, and R. Verma, "Hybrid PSO-GWO Optimization for Energy-Efficient WSNs," Applied Soft Computing, vol. 139, 2023.

[7] A. Sharma, A. Bansal, and A. Gupta, "Reinforcement Learning-Based Adaptive Clustering for Wireless Sensor Networks," IEEE Internet of Things Journal, vol. 8, no. 6, 2021.

[8] X. Zhou and Y. Liu, "A Genetic Algorithm-Based Clustering Approach for Wireless Sensor Networks," Computers and Electrical Engineering, vol. 90, 2021.

[9] H. Singh and S. Kaur, "Firefly Algorithm for Energy-Efficient Routing in Wireless Sensor Networks," International Journal of Communication Systems, vol. 34, no. 5, 2021.

[10] R. Kumar and R. Singh, "Grey Wolf Optimization Based Energy-Efficient Routing Protocol for Wireless Sensor Networks," Wireless Networks, vol. 26, no. 5, 2020.

[11] D. Yadav, A. Sharma, and N. Saxena, "A Hybrid PSO-SA Based Clustering Approach for Energy-Efficient Wireless Sensor Networks," Wireless Personal Communications, vol. 112, no. 4, 2020.

[12] S. Essamlali, H. Nhaila, and M. El Khaili, "Energy-Efficient Clustering Algorithm Based on Fuzzy Logic and Evolutionary Game Theory for Wireless Sensor Networks," Wireless Personal Communications, vol. 114, no. 2, 2020.

[13] M. Radhi and I. Tahseen, "An Enhancement for Wireless Body Area Network Using Adaptive Algorithms," International Journal of Computational and Experimental Science and Engineering, vol. 10, no. 3, 2024.

[14] S. A. Shehab, A. Darwish, A. E. Hassanien, et al., "Water Quality Classification Model with Small Features and Class Imbalance Based on Fuzzy Rough Sets," Environmental Development and Sustainability, 2023.

[15] A. Kaur, M. Khurana, P. Kaur, and M. Kaur, "Classification and Analysis of Water Quality Using Machine Learning Algorithms," in Proceedings of International Conference on Communication, Circuits, and Systems, Lecture Notes in Electrical Engineering, vol. 728, 2021.

[16] R. U. Maheshwari, S. Kumarganesh, and S. K. V. M, "Advanced Plasmonic Resonance-Enhanced Biosensor for Comprehensive Real-Time Detection and Analysis of Deepfake Content," Plasmonics, 2024.

[17] Khatami, S. S., Shoeibi, M., Salehi, R., & Kaveh, M. (2025). Energy-Efficient and Secure Double RIS-Aided Wireless Sensor Networks: A QoS-Aware Fuzzy Deep Reinforcement Learning Approach. Journal of Sensor and Actuator Networks, 14(1), 18. https://doi.org/10.3390/jsan14010018.

[18] Ayyappan A, Arunachalam R, Kumar ML. Adaptive fuzzy-based node communication performance prediction with hybrid heuristic Cluster Head selection framework in WSN using enhanced K-means clustering mechanism. Journal of Ambient Intelligence and Smart Environments. 2024;16(3):309-335. doi:10.3233/AIS-230408.

[19] Singh, B., Selvamuthukumaran, S., & Balaji, V. (2023). Energy Efficient Cluster Head Selection and Routing Algorithm using Hybrid Firefly Glow-Worm Swarm Optimization in WSN. KSII Transactions on Internet and Information Systems, 17(8), 2140–2156.

[20] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," Computer Networks, vol. 38, no. 4, 2002.

[21] Zhao, G., Lin, K., Chapman, D., Metje, N., & Hao, T. (2023). Optimizing energy efficiency of LoRaWAN-based wireless underground sensor networks: A multi-agent reinforcement learning approach. Internet of Things, 22, 100776.

[22] Hu, L., Han, C., Wang, X., Zhu, H., & Ouyang, J. (2024). Security enhancement for deep reinforcement learning-based strategy in energy-efficient wireless sensor networks. Sensors, 24(6), 1993.

[23] Priyadarshi, R., Bagri, S. S., & Ranjan, R. (2024, December). Transfer learning for efficient node placement in dynamic wireless sensor networks. In 2024 IEEE 16th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 147-152). IEEE.

[24] Sachithanandam, V., Jessintha, D., Balaji, V. S., & Manoharan, M. (2025). Deep reinforcement learning and enhanced optimization for real-time energy management in wireless sensor networks. Sustainable Computing: Informatics and Systems, 45, 101071.

Research Article | Volume 13, Issue 3 | Pages 447-456 | e-ISSN: 2347-470X

[25] Prabhu, D., Alageswaran, R., & Miruna Joe Amali, S. (2023). Multiple agent-based reinforcement learning for energy efficient routing in WSN. Wireless Networks, 29(4), 1787-1797.