Reinforcement Learning Driven Hybrid Clustering for Energy Optimization in Agri-IoT WSNs

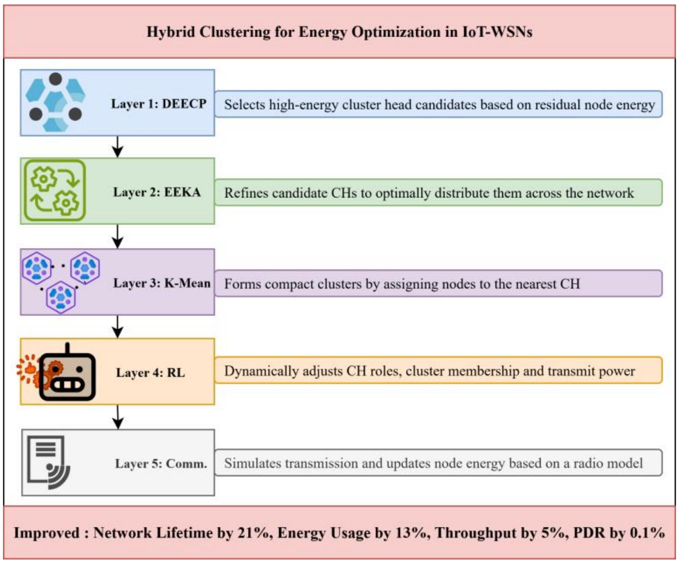
Shubham Kumar
Department of Electronics and Communication
National Institute of Technology
Patna, India
shubhamk.pg24.ec@nitp.ac.in

Bharat Gupta
Department of Electronics and Communication
National Institute of Technology
Patna, India
bharat@nitp.ac.in

Rakesh Ranjan
Department of Electronics and Communication
National Institute of Technology
Patna, India
rr@nitp.ac.in

Abhilekh Prakash
Department of Electronics and Communication
National Institute of Technology
Patna, India
abhilekhp.phd24.ec@nitp.ac.in

***Abstract*—The Agricultural Internet of Things (Agri-IoT) has rapidly evolved as a cornerstone of modern precision farming, supporting soil monitoring, irrigation control and real time environmental analysis. However, the performance of Wireless Sensor Networks (WSNs) that enable these applications often suffers due to limited node energy, uneven energy consumption and static clustering approaches. These factors collectively reduce the overall network lifetime. Conventional clustering techniques such as Low Energy Adaptive Clustering Hierarchy (LEACH) and Distributed Energy Efficient Clustering (DEEC) also struggle to adapt to the dynamic and heterogeneous nature of agricultural environments. To address these limitations, our work proposes a Reinforcement Learning Driven Hybrid Clustering (RLHC) framework that brings together DEEC, the Energy Efficient Knapsack Algorithm (EEKA) and K-Means clustering under a Q-learning based adaptive optimization strategy. In our design, DEEC is used to identify high energy Cluster Head (CH) candidates, EEKA helps to maintain spatial balance among CHs and K-Means organizes compact clusters to reduce intra cluster communication overhead. On top of that a Q-learning agent continuously observes network parameters such as residual energy, cluster load and packet delivery ratio (PDR) and takes adaptive actions including CH switching, node reassignment and transmission power adjustment. Through iterative learning, the agent gradually converges towards energy efficient configurations suited for real agricultural deployments. Simulation results indicate that the proposed RLHC framework outperforms both traditional and enhanced LEACH and DEEC variants, achieving approximately 21% longer network lifetime, 13% lower energy consumption, 5% higher throughput and 0.1% improvement in PDR. Overall, RLHC provides a scalable and energy aware clustering solution for modern Agri-IoT systems, contributing to more sustainable and data backed agricultural monitoring.**



***Keywords*—Wireless Sensor Networks (WSNs), Reinforcement Learning (RL), Q-learning, Distributed energy-efficient clustering (DEEC), Energy Efficient Knapsack Algorithm (EEKA), Energy Optimization, Internet of Things (IoT), K-Means Clustering.**

I. Introduction

The Internet of Things is expanding rapidly and one of the key technologies causing this growth is Wireless Sensor Networks (WSNs). Spatially dispersed sensor nodes make up these networks, which helps to monitor the environment and send the gathered data to a central Base Station (BS) or sink node for additional processing and analysis. Contemporary IoT ecosystems take advantage of their ability to deliver context aware, real time information. Examples of applications for WSNs are numerous, including military surveillance, smart agriculture and industrial automation [1], [2]. WSNs are the backbone of IoT ecosystems leveraging efficient data collection and communication, enabling autonomous operations, predictive analytics and data driven decision making. However, despite their great potential, WSNs encounter significant limitations in field of the network lifetime, reliability of communication and energy efficiency. The sensor nodes in WSNs have small, non-rechargeable batteries. Generally these nodes are deployed in remote or hostile locations which make accessing individual nodes to perform maintenance or replace batteries very difficult [3]. As a result, eventually, the battery depletes energy in an individual node, partitions the network, potentially loses packets and decreases the overall performance of the system. This scenario is particularly prevalent in applications which require long term continuous monitoring, such as precision agriculture, disaster prediction and health diagnostics, in which data has to flow reliably and continuous operation is a necessity. To mitigate these issues, numerous routing and clustering protocols focused on energy optimization and energy efficiencies. Researchers developed varieties of protocols on balancing power consumption among sensor nodes. Of these protocols, clustering based routing has shown to be one of the most effective means for energy efficiency [4]. The renowned clustering protocols are LEACH, Stable Election Protocol (SEP) and DEEC [3]–[5]. These protocols are mainly designed to minimize communication overhead and ensure a balanced distribution of energy consumption throughout the network. DEEC, in particular, selects CHs based on residual energy and average network energy, thereby extending the network lifetime more effectively than random or static CH selection methods such as LEACH.

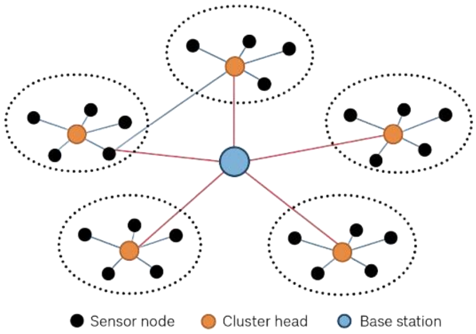


Figure 1: Wireless Sensor Network

Here figure 1, represents a typical communication model of a WSN, where sensor nodes are randomly deployed across the sensing region and relay their sensed data to the BS through CH.

However, traditional clustering protocols still have limitations in adaptability, scalability and dynamic optimization. Most of the existing schemes assume homogeneous or quasi-static environments, where node energy dissipation, communication distance and network topology remain relatively stable. In real world IoT deployments, the energy dynamics, node heterogeneity and communication interference can vary dramatically over time. As a result, static or semi-static CH selection often leads to sub optimal cluster configurations, premature energy depletion of certain nodes and overall performance degradation [5]. Also, existing methods rarely incorporate adaptive learning or feedback mechanisms that can respond to constant changing environmental conditions or operational uncertainties. Thus the design of an intelligent and adaptive clustering mechanism becomes crucial for achieving long term sustainability [6].

So to mitigate the above mentioned limitations, our study introduces a Reinforcement Learning Driven Hybrid Clustering (RLHC) approach that fuses intelligent optimization techniques across multiple layers of the clustering process [7]. The proposed RLHC framework integrates DEEC, EEKA, K-Means clustering and RL to form an adaptive and energy aware communication architecture where DEEC serves as the foundational layer, prioritizing nodes with higher residual energy for CH selection to ensure balanced energy consumption, EEKA filters and ranks candidate CH based on centrality and spatial distribution which determines the optimal number of CHs and select the best CHs to improve spatial balance and reduce inter cluster communication overhead, K-Means algorithm refines the cluster formation by minimizing intra cluster distances and balancing the spatial distribution of nodes, thereby further reducing communication cost. The RL layer introduces a self-learning capability that dynamically adjusts CH selection, transmission parameters and cluster reconfiguration based on real time feedback such as residual energy, cluster load, communication delay and PDR. The final layer models the radio energy consumption during data transmission and reception between nodes, CHs and the BS using the first order radio model which includes the free space (d^2) and multipath (d^4) channel models for short and long distances, respectively [8]. The RLHC framework accomplishes adaptive energy balancing, increased throughput, improved PDR and extended network lifetime by combining these five layers. The potential to close the gap between contemporary ML integrated adaptive optimization techniques and conventional deterministic routing methods is demonstrated by this hybrid intelligent model. It opens the door for more dependable and sustainable sensor network deployments in real world settings by offering a solid, scalable and self-evolving solution for the future of IoT-enabled WSNs.

Our contributions can be summarized as follows:

i. A five-layer RLHC model integrating DEEC (layer 1), EEKA (layer 2), K-Means (layer 3), RL (layer 4) and Communication layer (layer 5) for intelligent energy management in WSNs.

ii. Energy aware optimal clustering and RL based adaptive cluster refinement mechanism that dynamically optimizes cluster configurations in real time network conditions.

iii. Performance enhancement as compared to the optimized algorithms built on top of traditional clustering protocols in terms of network lifetime, throughput, energy consumption and PDR.

The remaining part of this paper is organized as follows:

Section II represents a comprehensive literature review of existing clustering protocols and RL approaches in WSNs. Section III describe the theoretical and technical background related to energy efficient communication and clustering mechanisms. Section IV focus on the methodology of proposed RLHC framework in details. Section V presents the simulation setup, performance evaluation parameters and comparative analysis of the proposed RLHC protocol with existing approaches. Finally, Section VI provides the conclusion and highlights possible future research directions.

II. Related Works

WSNs have become an integral component of modern IoT systems, which enables various applications in domain such as healthcare, agriculture etc. However, they still face challenge regarding energy efficiency. Since sensor nodes are battery powered, the overall network performance heavily relies on the effective management and utilization of this energy. Over the years, researchers have proposed a variety of clustering protocols, optimization algorithms and adaptive learning strategies to address this challenge. This section reviews existing work in this field [9].

A. Classical Clustering Protocols

Clustering is well recognized as a foundational approach for energy efficient WSNs. The LEACH protocol is the most seminal contributions, introducing randomized rotation of CHs to evenly distribute energy consumption among nodes. It reduces the number of direct transmissions to the BS and organizes nodes into energy efficient clusters [10]. But still it has several limitations, including uneven CH distribution, lack of consideration for residual energy and poor performance in heterogeneous or large scale networks. These limitations motivated the development of variants that incorporate energy awareness and residual energy into CH selection [9]. DEEC protocol is also one of the renowned protocol for achieving energy efficiency in WSNs, which take residual energy in account to find the suitable CH which balance the energy of network and improve the network lifetime [11]. Few protocol has been developed by integrating RL such as LEACH RLC and ReLeC (enhance LEACH) for adaptive CH selection. These methods learn network states and make dynamic decisions about CH election to improve energy efficiency and extend the network lifetime. While the results are encouraging, RL based variants can add computational and communication overhead. The practical applicability of RL models such as ReLeC may be limited in large scale or highly dynamic WSNs due to convergence issues [12]. It is evident that RL based clustering protocols have demonstrated significant gains in throughput, network stability and energy efficiency. However, these methods have certain drawbacks. For IoT nodes, the learning stage of RL models frequently adds computational and communication overhead. It may depletes the sensor node's energy, which is already scarce in terms of processing power.

B. Heterogeneity-Aware Protocols

In order to address the limitations of conventional homogeneous protocols, researchers have developed methods that acknowledge a heterogeneous community of nodes. SEP (Stable Election Protocol) and DEEC are the two most well-known protocols that take into account the fact that nodes have different initial energy states and CH selection is based on residual energy states as well as energy distribution throughout the network. This ensures that nodes that are more energetic are more likely to accept the responsibilities of being a CH [13]. DEEC, in particular, increases functionality and network lifetime by extending LEACH to include energy heterogeneity and energy based CH selection methods. Other enhancements to the CH selection process such as EEKA reads the CH role allocation undergoing based on both node centrality and spatial distribution to promote more uniform CH allocation to really achieve reduced distances in intra cluster communication and energy consumption.

C. Metaheuristic and Optimization-Based Approaches

Several metaheuristic algorithms, such as Particle Swarm Optimisation (PSO), Genetic Algorithms (GA) and Ant Colony Optimisation (ACO) have been extensively investigated in WSNs to support CH selection and network performance. Based on simulating the behavior of a swarm of like particles, PSO associates the best CH locations and reduces intra cluster distances [14]. Conversely, ACO uses a pheromone based path selection process to increase routing effectiveness and cluster formation. GA evolve the population of CH candidates through time, improving load balancing and energy efficiency. However, once an algorithm generates its optimal solution, like a GA, it typically does so offline, while other metaheuristic approaches often rely on iterative based processes [15]. These techniques are less effective in networks that are large, mobile, have changing topologies, or quickly fluctuating energy states.

D. Fuzzy Logic and Multi-Criteria Clustering

Clustering based on fuzzy logic provides a multi criteria decision making technique for CH selection, considering factors like residual energy, nodes density, distance to the cluster center and network traffic. MRCH (Modified RCH-LEACH) and others use fuzzy rule sets to choose eligible CH candidates. The method promotes a higher packet delivery ratio, provides improved network stability and increases overall energy efficiency. This approach contributes to clusters exhibiting dynamically adaptive characteristics in response to dynamic changes in the network by providing a flexible method for managing uncertainty in sensor networks [16]. However, some low power sensor nodes may restrict the use of fuzzy based computations due to their computations complexity.

E. Reinforcement Learning in WSNs

RL has become a powerful technique for control of dynamic and adaptive WSNs. An environment is developed using RL based notation of the network. Nodes or CHs are agents learning the optimal action to maximize long term utility, such as lifetime and energy efficiency. Q-learning, SARSA and DQN have been implemented to improve CH selection, cluster re-organisation and routing. Techniques such as EER-RL increase energy efficiency and lifetime by modifying CH roles in response to learnt energy patterns [17]. Similarly, Q-learning integrated with LEACH increase the protocol's adaptability to variations in network topology. While these RL based approaches work well, RL often requires global knowledge, involves training in a centralized scenario, which may limit the scalability of RL based techniques for large scale dynamic deployments of WSNs.

F. Hybrid Clustering Approaches

Hybrid clustering methods aim to combine the strengths of classical, meta heuristic, fuzzy and RL approaches. These methods address multiple challenges simultaneously. This covers energy balancing, spatial uniformity, CH optimisation and adaptability. In order to maximise network efficiency, we discovered that multi layered hybrid models frequently incorporate energy aware CH selection, K-means or fuzzy based spatial clustering and RL based adaptive decision making. The effectiveness of such hybrid frameworks has been demonstrated in recent work. The ideal number of cluster and grid heads to balance energy consumption is determined by protocols such as EOCGS [18]. According to these studies, intelligent hybrid techniques can perform better than conventional methods in terms of network lifetime, energy balance and adaptability.

Even with notable improvements in WSN clustering, existing methods continue to face several challenges. Metaheuristic approaches often needs heavy computation. Classical methods frequently fail to prevent early node depletion, resulting in network partitioning. In contrast, Q-Learning introduces minimal computational overhead and enables the system to adapt effectively to dynamic environments. Moreover, most existing solutions rely on static or pre defined strategies, which are not enough to handle dynamic topologies or sudden energy fluctuations.

TABLE 1. Summary of energy efficiency studies in WSN

SL. No.	PAPER & YEAR	TECHNIQUE USED	LIMITATIONS (RESEARCH GAP)
1	Farahzadi et al. "An Improved Cluster Formation Process in Wireless Sensor Networks to Decrease Energy Consumption" (2021) [19]	BS selects CH based on node's residual energy and proximity within predefined zones	Assumes ideal energy estimation and homogeneous energy levels BS assumes global knowledge of all node's energy
2	Panchal et al. "EEHCHR: Energy Efficient Hybrid Clustering and Hierarchical Routing for Wireless Sensor Networks" (2021) [20]	Hybrid clustering with hierarchical routing using FCM	Increased routing complexity Static clustering radius and considered homogeneous WSN scenarios.
3	Al-Kaseem et al. "Optimized Energy-Efficient Path Planning with Multiple Mobile Sinks" (2021) [21]	Stable Election Algorithm (SEA) and residual energy for CH selection	Scalability issues beyond 100 nodes Simulation-only, no deployment
4	Prajapati et al. "Performance Analysis of LEACH with Deep Learning in Wireless Sensor Networks" (2022) [22]	CNN-based CH selection using LEACH algorithm	High computational overhead Limited scalability.
5	Mohapatra et al. "Mobility Induced Multi-Hop LEACH Protocol in Heterogeneous Mobile Network" (2022) [23]	Introduces a multi-hop extension of LEACH for heterogeneous mobile sensor networks	Assumes uniform mobility Limited scalability and fails in highly dynamic or large scale network conditions.
6	Gamal et al. "Enhancing Lifetime of WSNs Using Fuzzy Logic LEACH and PSO" (2022) [24]	Hybrid clustering algorithm K-Mean and PSO for cluster formation	Assumes static deployment and homogeneous node capabilities, limiting its applicability in highly dynamic or heterogeneous WSN
7	Bhatia et al. "Cluster Based Energy Efficient Routing Protocol using SA-LEACH to Wireless Sensor Networks" (2023) [25]	LEACH and a Simulated Annealing (SA) meta-heuristic to pick a better (neighboring) node if available.	High computational cost Considers static and homogeneous networks, limiting adaptability
8	Abose et al. "Improving Wireless Sensor Network Lifespan with Optimized Energy-Conscious Routing" (2024) [26]	Optimized energy conscious routing with a corner positioned base station	Assumes ideal energy estimation Does not consider node location or distance factor
9	El Khediri et al. "Energy-Efficient Cluster Routing Protocol for Wireless Sensor Networks" (2024) [27]	Cluster based routing using hybrid meta heuristic algorithms	Designed for homogeneous, static node deployments Limited adaptability
10	Zhu et al. "Improved Soft-k-Means Clustering Algorithm for Balancing Energy Consumption in Wireless Sensor Networks" (2024) [28]	Soft-k-means clustering with multi CH	High computational cost Limited adaptability
11	Tabatabaei et al. "New Energy Efficient Management Approach for Wireless Sensor Networks" (2025) [29]	Hierarchical clustering approach with meta heuristic algorithms	Assumes ideal energy estimation Does not consider node location or distance factors.

III. Technical Background

A. Wireless Sensor Networks and Energy Constraints

WSNs comprise a collection of sensor nodes that are spatially distributed to observe various physical or environmental parameters. Then it transmit the collected data to a central sink. Each sensor node typcally has limited energy, computation capability and communication range. As a result, operating the network in an energy-efficient manner becomes crucial to prolong its lifetime and ensure consistent data delivery [2].

The network operates in rounds, each consisting of cluster formation, data aggregation and transmission. Efficient energy management is critical because sensor nodes are battery powered and recharging may be impractical. Key challenges in WSNs include:

a) Limited energy resources.

b) Uneven energy depletion due to repeated CH selection.

c) Scalability for large networks.

d) Adaptability to dynamic scenarios including node failures, mobility and environmental variations.

To mitigate these challenges, cluster based routing is commonly employed, complemented by optimization techniques that efficiently manage intra cluster communication and CH selection, thereby enhancing the overall network lifetime [4].

B. Particle Swarm Optimization (PSO)

To mitigate the above mentioned limitations various algorithms are developed in which POS is very renowned. This algorithm is population-based and models the cooperative movement patterns observed in bird flocking and fish schooling [24]. Each member of the population, known as a particle, represents a possible solution. These particles move within the search space, updating their positions according to their own best performance and the best performance observed among all particles. Over iterations, particles converge toward the best solution.

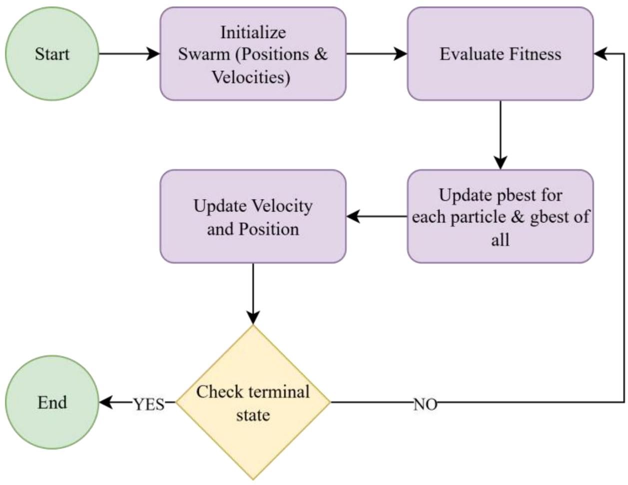


Figure 2: Flow chart of the PSO algorithm

Figure 2 represents the working principle PSO algorithm. Initially, a group of particles is generated, each symbolizing a candidate solution with randomly assigned positions and velocities. At each iteration, the fitness of all particles is evaluated using the objective function. Each particle then updates its personal best position (pbest) and the global best position (gbest) found by the entire swarm [30]. The velocity and position of each particle are updated according to the Equation 1 and Equation 2 respectively.

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (pbest_i - x_i(t)) + c_2 \cdot r_2 \cdot (gbest - x_i(t)) \tag{1}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{2}$$

where 'w' denotes the inertia weight, 'c1' and 'c2' are acceleration coefficients and 'r1', 'r2' belongs to [0,1] are random numbers. These Equations together balance the inertia (momentum), cognitive (self-learning) and social (swarm cooperation) components of each particle. This process iteratively continues until the swarm converges toward an optimal or near-optimal solution based on the defined fitness function.

C. Ant Colony Optimization (ACO)

One the well known algorithm ACO which is a metaheuristic algorithm introduced by Marco Dorigo. The algorithm is inspired by the imagining the behavior of real ants [31]. In nature, ants discover the shortest route between their nest and a food source by laying down pheromones along their paths. Then the other ants sense this pheromone trail and are more likely to follow stronger trails. Over time, shorter paths accumulate more pheromones which leads to the optimal path.

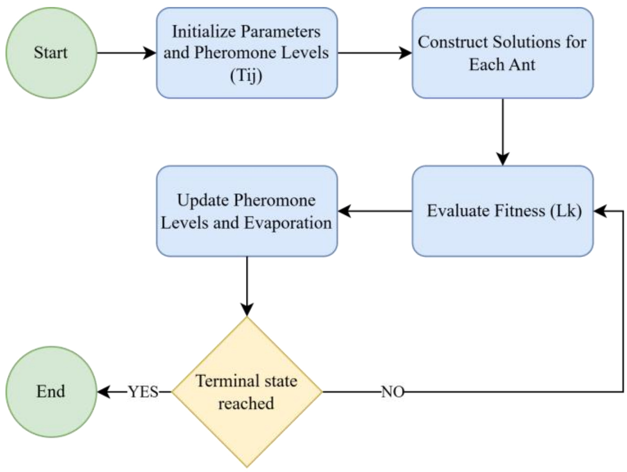


Figure 3: Flow chart of the ACO algorithm

Figure 3 show the working principle of Ant Colony Optimization (ACO) algorithm. It simulates artificial ants that iteratively construct solutions using pheromone trails and heuristic information. The process begins with the initialization of parameters and pheromone levels on all paths. During solution construction, each ant builds a solution based on the probability of selecting the next path. After all ants complete their paths, the pheromone update phase reinforces paths used by better solutions while allowing pheromone evaporation to prevent premature convergence and find optimal path. The algorithm proceeds in an iterative manner until a stopping criterion such as reaching the maximum number of iterations or achieving convergence is met [31]. The core Equations of ACO are the path selection probability probability (Equation 3), pheromone update (Equation 4) and pheromone deposition (Equation 5).

$$P_{ij}^{k}(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum\limits_{l \in N_i^k} [\tau_{il}(t)]^{\alpha} \cdot [\eta_{il}]^{\beta}}, & \text{if } j \in N_i^k \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where P_{ij} is probability that ant k moves from node i to node j, T_{ij} is pheromone concentration on edge (i,j) at time t. ηij is equals to 1/dij that is heuristic information (inverse of distance or cost). α, β are control parameters for pheromone and heuristic influence.

$$\tau_{ij}(t+1) = (1-\rho) \cdot \tau_{ij}(t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t) \tag{4}$$

Equation 4 represents pheromone update rule, where ρ represents pheromone evaporation rate (0<ρ<1), m is total number of ants and ΔT_{ij} amount of pheromone deposited by ant k on edge (i,j).

$$\Delta\tau_{ij}^{k}(t) = \begin{cases} \dfrac{Q}{L_k}, & \text{if ant } k \text{ uses edge } (i,j) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Equation 5 represents pheromone deposition amount, where Q is pheromone constant and L_k is total cost or path length of the tour constructed by ant k.

D. Fuzzy C-Means (FCM)

The FCM algorithm is a well known clustering technique where each data point belongs to a cluster with a degree of membership. This algorithm is fundamentally an optimization based clustering method which seeks to minimize an objective function that quantifies the total weighted distance between data points and cluster centers [20]. Each data point is assigned a membership degree to all clusters, allowing for soft clustering where points can partially belong to multiple clusters. The optimization process aims to find the optimal cluster centers that minimize the objective function by updating the membership degrees for each data point.

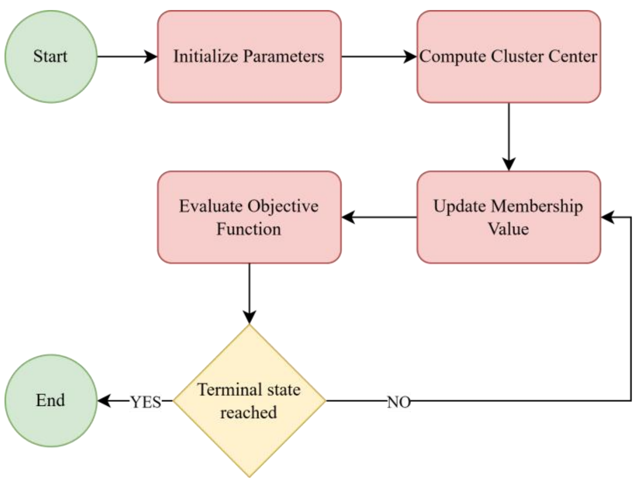


Figure 4: Flow chart of the FCM algorithm

Figure 4 illustrates the working principle of the FCM algorithm. Here the process begin with the initialization of the membership matrix U. It assigns random membership values to each data point for all clusters. Next, cluster centers are computed based on the current membership values. Then the algorithm updates the membership degrees for each data point using the updated cluster centers. Later on the objective function is evaluated to measure clustering performance. If the change in membership values or cluster centers between iterations is smaller than a predefined threshold, the process is said to have converged, otherwise, the algorithm repeats the update steps. Once convergence is achieved, the final cluster assignments are generated, representing the optimal fuzzy partition of the dataset [32].

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \tag{6}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \tag{7}$$

$$u_{ij} = \frac{1}{\sum\limits_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \tag{8}$$

Equation 6 represents the objective function that FCM seeks to minimize. It measures the total weighted distance between all data points xi and the cluster centers cj. The weight is given by the membership value Uij, m which indicates how strongly a data point belongs to a particular cluster. Equation 7 is used to update the cluster center. Each cluster center is calculated as the weighted mean of all data points, where the weights are the membership degrees raised to the power m. Equation 8 defines how the membership value Uij of each data point to each cluster is updated. From equation we can conclude that the membership is inversely related to the distance between the data point and the cluster center. The ratio term ensures that all membership values for a data point sum to 1 across all clusters for maintaining normalization.

IV. The Methodology

This paper presents a Reinforcement Learning Driven Hybrid Clustering (RLHC) framework. It improve energy efficiency and operational lifetime in WSNs. The whole framework divided into a five-layer architecture, in which each layer contribute its own advantage. The whole framework collectively optimizes CH roles and network management through real time feedback. From simulation results we can confirm that the integrated RLHC framework delivers improved throughput, higher packet delivery ratio (PDR) and extends network lifetime. It merges the deterministic optimization with machine learning guided adaptation. This is making whole system robust across diverse IoT enabled WSN scenarios. The RLHC framework composed of five integrated layers designed to improve energy efficiency and clustering in WSNs.

A. Layer 1: Distributed Energy Efficient Clustering Protocol (DEECP)

DEECP act as the basic fundamental layer of the proposed RLHC framework. DEECP employs an energy aware mechanism to identify the most suitable CH candidates. This algorithm compares the average energy of the network to the residual energy of each node. With this flexible CH selection method, nodes with a higher probability of becoming CHs is associated with higher remaining energy and causes energy consumption to be balanced throughout the network [4]. By alternating CH roles among qualified nodes on a regular basis, it effectively reduces premature node death and improves network stability. In heterogeneous IoT-enabled WSN environments, it not only maximises the initial CH selection but also works in concert with higher level adaptive modules to achieve balanced load distribution, extended network lifetime and improved overall energy efficiency.

$$T(i) = \begin{cases} \frac{P_i}{1-P_i\cdot\left(r \bmod \left(\frac{1}{P_i}\right)\right)}, & \text{if } i \in G \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

$$P_i = prop \times \left(\frac{E_i}{\bar{E}(r)}\right) \tag{10}$$

$$\bar{E}(r) = \frac{E_{\text{total}} \cdot \left(1 - \frac{r}{R}\right)}{N} \tag{11}$$

Equation 9, represents T(n) which is threshold for node n to become CH, Pi is probability of node i becoming a CH, r is current round number and G is set of nodes that have not been CHs in the last 1/Pi rounds. In Equation 10, the Pi is proportional to the Ei which represents residual energy of node i, E(r) which represent average residual energy of the network at round r, 'prop' is optimal CH probability, N represents total numbers of nodes, R represents total numbers of rounds and Etotal represents total energy of the network.

B. Layer 2: Energy Efficient Knapsack Algorithm (EEKA)

Following the DEECP layer, the EEKA operates as the second optimization layer within the RLHC framework. While DEECP ensures energy aware CH selection based on residual energy and average network energy, EEKA refines this process by optimizing the spatial distribution and number of CHs using a constrained knapsack formulation. In this layer, each potential CH candidate identified by DEECP is evaluated as an item in the knapsack problem, where parameters such as node centrality, residual energy and communication distance to the BS act as utility factors. It focus on maximizing the overall CH efficiency by considering minimum use of total energy consumption. The selection of the CH set is spatially balanced and energy efficient [33]. It guarantees uniformly coverage of the sensing field and limits overlong intra cluster communication distances. We can infer this selection eliminates energy hotspots and promotes uniform energy dissipation across the entire network's sensing field. The EEKA approach also resolves areas of spatial irregularity typically common in DEECP or LEACH based clustering approaches. As an additional benefit, the final CH arrangement reduces redundant transmissions, increasing communication reliability and network throughput. Furthermore, the optimal CH arrangement developed within EEKA provides a better input to the K-Means clustering layer, which refines the cluster boundaries to ensure even less intra cluster distances and transmission costs. Additionally, the layered optimization approach keeps transmission paths short and more evenly distributed energy consumption over the sensing field.

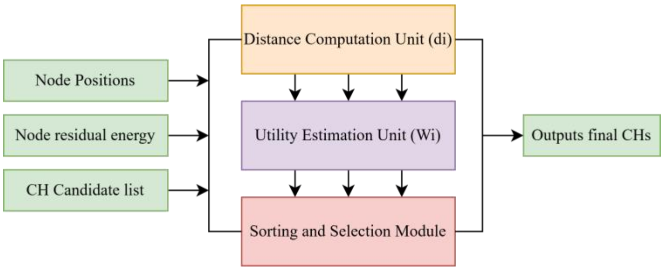


Figure 5: Internal block diagram of the EEKA algorithm

The internal design of EEKA is depicted in Figure 5. It works by obtaining candidate CH, node information about energy and position. It then calculates average inter node distances and determines node centrality using a combined utility function (Equation 12) to rank the candidates. The top K candidates are selected as CH to achieve the most balanced energy consumption and an even cluster distribution throughout the network.

$$W_i = E_i + \frac{1}{\bar{d}_i + \epsilon} \tag{12}$$

Here the utility function Wi represents the energy centrality score of each node. It determines node suitability to become a CH. It integrates two crucial parameters, residual energy and spatial centrality into a single quantitative metric.

C. Layer 3: K-Means Algorithm

In the third layer of the RLHC framework, K-Mean is used as a clustering algorithm. It refine the cluster formation process based on the CHs selected by the EEKA layer. While DEECP and EEKA togther determine the most energy efficient and spatially balanced CHs, the K-Means algorithm ensures that the remaining sensor nodes are optimally associated with these CHs to minimize intra cluster communication cost.

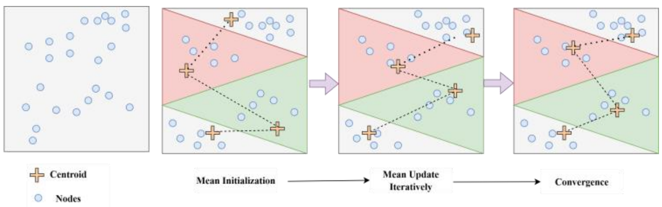


Figure 6: Process flow chart of K-Mean algorithm

The K-Means algorithm partitions a set of N sensor nodes into K clusters by iteratively minimizing the sum of squared Euclidean distances between each node and its assigned CH [34]. The process involves three main steps: centroid initialization, mean (centroid) update and convergence check.

$$J = \sum_{i=1}^{K} \sum_{x \in C_i} \|x - \mu_i\|^2 \tag{13}$$

The objective function of K-Mean clustering process is given in Equation 13, where Ci represents the set of nodes in cluster i, μi is the centroid of cluster i and sq(x−μi) denotes the squared Euclidean distance between node x and its cluster centroid. J is non-convex, meaning K-Means may converge to a local minimum, not necessarily the global one.

D. Layer 4: Reinforcement Learning (RL)

Layer 4 incorporates a RL agent to provide self-learning, adaptive control for the WSN. By utilizing real-time feedback from the network itself, it optimally adjusts and changes the CH selection, transmission parameters and cluster configurations dynamically. The RL agent uses the Q-Learning algorithm to obtain improved, optimal CH selection and transmission decisions/parameters. The learning agent focuses on maximizing long-term objectives, such as network lifetime and energy efficiency, rather than optimizing short term goals [34]. The RL agent is able to continuously improve and update its strategies over time and avoid some of the limitations associated with static, rule based systems. The RL agent also works in conjunction with the previous layers to improve the overall decision making process, while maintaining all capabilities for scalable, efficient and adaptive clustering. Although the RL agent does involve a small amount of computational overhead but it is crucial for long term dynamic system.

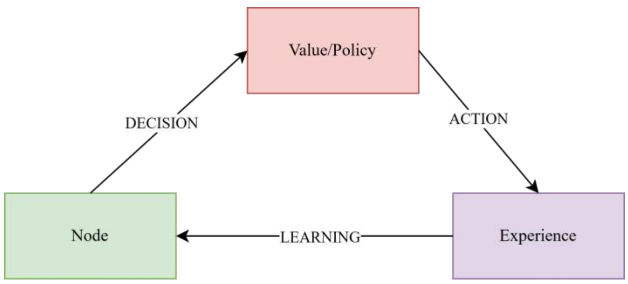


Figure 7. RL based Decision-Making Framework

In the diagram above shows that, with RL the system is able to flexibly adjust to new conditions and continually improve its decisions in a more optimal way as time stoops. This is what makes it very applicable to real-time IoT applications.

In Layer 4, a Q-learning agent is implemented with the sole objective of observing the current network state, taking actions (e.g., CH re-assignment, switching CHs and changing transmission power) and receiving rewards as experience. After executing an action, it receives feedback from the environment, including updated node energies, cluster configurations and network metrics (e.g., PDR, throughput) and computes a reward. This experience is then used to update the Q-table, refining future decision making. By introducing adaptive learning, this layer transforms static cluster management into a self-learning mechanism, enabling the network to optimize long term objectives such as energy efficiency and network lifetime. This layer is very crucial and enables the system to adapt effectively to dynamic changing environments.

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] \quad (14)$$

Equation 14 is used to update the Q-table, where α is the learning rate that determines how much new information overrides old knowledge, γ is the discount factor that defines the importance of future rewards and 'a' represents the action taken by the agent. This update allows the agent to iteratively improve its policy by balancing immediate rewards with long-term benefits, enabling more optimal decision making over time.

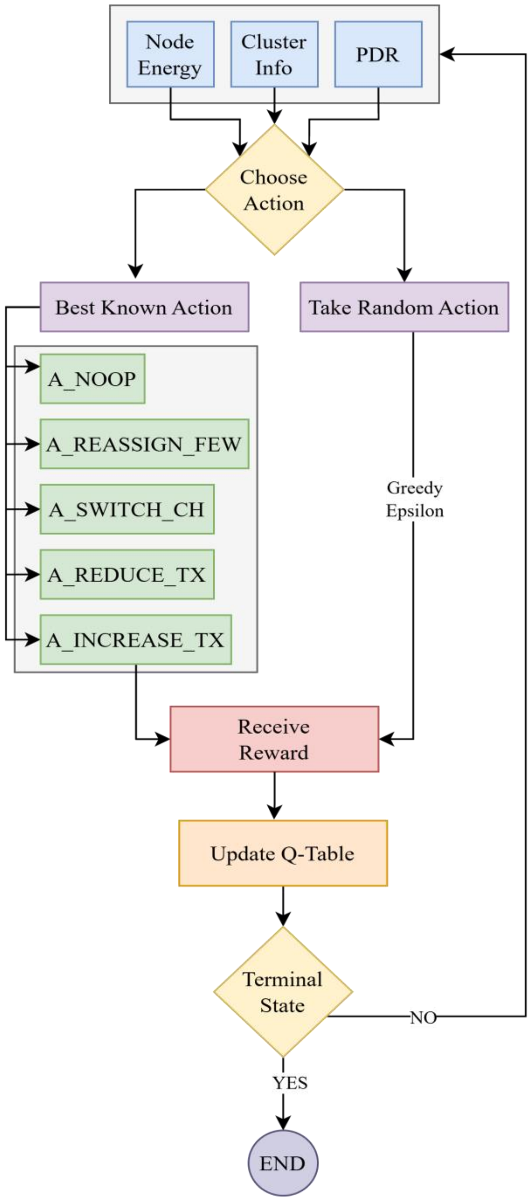


Figure 8. Q-Learning agenrt decision making framework

Figure 8 illustrates how the agent interacts with the environment to make adaptive chnages and fine tune various parameters.

E. Layer 5: Communication Layer

Layer 5 is responsible for wireless communication and energy modeling during data transmission between nodes, CHs and the BS. This layer simulates the radio energy dissipation model, which estimates the transmission and reception energy costs based on the distance between nodes and the amount of data transmitted. In each communication round, member nodes transmit sensed data to their respective CHs, which then perform data aggregation and forward the aggregated packet to the BS using a single-hop communication mechanism [35]. In the proposed model, wireless communication is simulated using the first-order radio energy model that is Additive White Gaussian Noise

(AWGN) channel., which incorporates both free-space (d^2) and multipath (d^4) channel models. For short-range intra cluster communication ($d<d_0$), the free-space model is used, assuming line-of-sight propagation. For long-range transmissions ($d \geq d_0$), the multipath model is applied to account for non-line-of-sight energy loss. This dual channel approach effectively captures large-scale path loss while maintaining computational simplicity, providing a realistic approximation of WSN communication behavior.

$$E_{Tx}(k,d) = \begin{cases} k \cdot E_{elec} + k \cdot \varepsilon_{fs} \cdot d^2, & \text{if } d < d_0 \\ k \cdot E_{elec} + k \cdot \varepsilon_{mp} \cdot d^4, & \text{if } d \geq d_0 \end{cases} \quad (15)$$

Equation 15, computes the energy needed to transmit a k-bit packet over a distance d, Longer distances ($d \geq d_0$) incur significantly higher energy costs due to multi path fading, where k is size of the data packet in bits, E_elec is energy required per bit for transmission/reception circuitry, ε_fs is amplifier energy for free-space model, ε_mp is amplifier energy for multipath model, d is distance between sender and receiver, d_0 is threshold distance to switch between free-space and multi path.

$$E_{Rx}(k) = k \cdot E_{elec} \quad (16)$$

$$E_{DA}(k) = k \cdot E_{DA} \quad (17)$$

Equation 16, calculates the energy required to receive k bits, receiving costs are independent of distance, as no amplification is needed. Equation 17, E_DA is energy required for data aggregation for k bits before transmission.

$$E_{CH} = \left(\frac{N}{k} - 1\right) kE_{elec} + N \cdot E_{DA} + kE_{elec} + k\varepsilon_{fs} d_{BS}^2 \quad (18)$$

$$E_{nonCH} = k \cdot E_{elec} + k \cdot \varepsilon_{fs} \cdot d_{CH}^2 \quad (19)$$

Equation 18 and 19 calculates the total energy consume by the CHs and CMs, when considering no multi path fading.

The proposed architecture intends to raise the energy efficiency and reliability of communication for WSNs in Agri-IoT contexts. Features the multi-layer architecture that effectively blends conventional clustering approaches and smart reinforcement learning for optimization. In this system, sensor nodes deployed randomly throughout the agricultural field. The sensor nodes will continuously collect environmental data such as soil moisture, humidity and temperature. The data collected from the sensor nodes will then be sent to a base station, where the data will be processed and decisions made with the results. The architecture consists of five functional layers. In the first layer, the nodes are deployed and run DEECP, to identify the best options to serve as CH. In the second layer, the remaining energy and degree of centrality of the candidates are identified and ranked. In the third layer, the K-Means clustering algorithm is performed, linking the non CH nodes to the closest CH will reduce communication distances between nodes in the cluster and distribute energy consumption. The fourth layer introduces a Q-learning agent that intelligently optimizes CH selection based on the ever changing state of the network. The agent observes the environment, selects actions such as reassigning CHs or adjusting transmission power and updates its Q-table based on the obtained rewards. Ultimately, the fifth layer is focused on the communication, where member nodes transmit their data to CHs and aggregated packets will be forwarded to the BS. The radio communication model used in this work incorporates free-space (d^2) and multipath (d^4) channel propagation characteristics which represent conditions of agricultural fields where line-of-sight and non-line-of-sight transmissions are expected. Due to this holistic design, extends network lifetime and ensure data reliability.

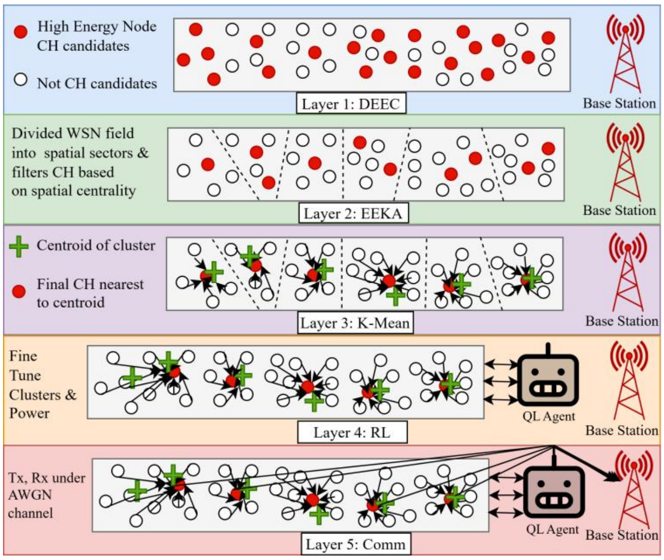


Figure 9. System architecture of the proposed algorithm

Algorithm 1: RHLC - Reinforcement Learning Driven Hybrid Clustering

---

Initialize:

Network parameters, node positions and energy.

Q-Table Parameters

Set round counter r = 0.

While (any node is alive):

Layer 1 (DEEC): *Estimate CH probability based on residual energy ratio.*

Layer 2 (EEKA): *Divide WSN field into spatial sectors and select energy-efficient CH candidates using node centrality and residual energy.*

Layer 3 (K-Means): *Form compact clusters by associating nodes with the nearest CH closest to the centroid.*

Layer 4 (Q-Learning): *Observe state (residual energy, distance, alive ratio, pdr, cluster info) Then choose action using ε-greedy policy (CH reassignment / maintain). Then update Q-table using Bellman update equation.*

Layer 5 (Communication): *Transmit data: Member nodes to CH (Free-space, d^2), CH to BS (Multipath, d^4) under AWGN channel. Then update node energies and performance metrics (PDR, throughput etc).*

Increment round counter r.

End While.

Table 2: WSN parameters for Agri-IoT WSNs

Network Parameters	Value
WSN area size	500m x 500m
Number of nodes	200
Initial energy	2 Joules
Packet size	512 bytes
Data aggregation energy	5n J/bit/signal
Transmit energy	50n J/bit
Receive energy	50n J/bit
Free space loss	10n J/bit
Multipath loss	0.0013p J/bit
Simulation Rounds	1000

The simulation will evaluate WSN performance by tracking the number of dead nodes, network lifetime, energy consumption, throughput and improvement in PDR.

V. Results and Discussion

This section presents the performance evaluation results of the optimized algorithms built on top of LEACH and DEEC and the proposed mechanism. The analysis focuses on the network lifetime, energy consumption, throughput and packet delivery ratio. The simulations assume that all nodes are either stationary or exhibit only micro-mobility, energy losses due to dynamic random channel conditions are neglected.

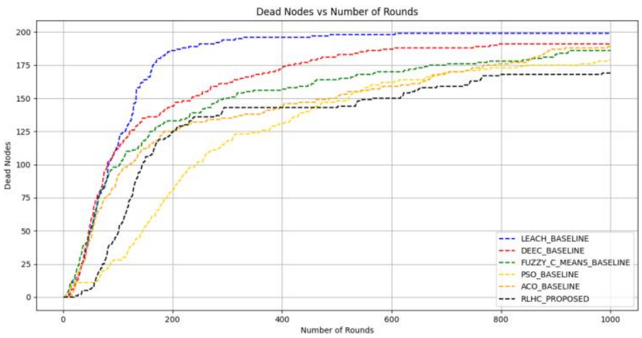


Figure 10. Number of dead nodes during network operation

The relative performance of different protocols in terms of network lifetime is shown in Figure 10. The first node dies around node 60 in the LEACH protocol, which exhibits the earliest network degradation. By extending the first node death to approximately 100, DEEC provides a minor improvement. Algorithms such as FCM and PSO improve the network's overall stability, ACO typically performs marginally better in most situations. In terms of protocol performance with respect to the first node failure time metric, RLHC, demonstrates that the first node failure occurs around 150 rounds, which is about 1.2 times later than ACO and 3.8 times later than LEACH. The reduced rate of nodes depleting energy suggest that RLHC's adaptive learning method provides a more equitable distribution of energy throughout the network.

Data throughput displays a similar trend depicted in Figure 11. LEACH shows the lowest total throughput at roughly 23,000 packets after 1000 rounds, primarily due to early node deaths. DEEC and Fuzzy C-Means bring throughput levels up to about 44,000 packets and 55,000 packets, due to their enhanced CH selection process. The same trend is continued further with PSO and ACO, which both maintain poor stability throughout the simulation, while also improving throughput, peaking at roughly 53,000 and 71,000 packets, respectively. The proposed RLHC protocol presents a dramatic reduction in throughput as the RLHC protocol and achieves the highest cumulative and final throughput at nearly 75,000 packets sent to the BS. This is a whopping +226% greater than LEACH and +5% more packets sent than ACO, thereby supporting the hypothesis that reinforcement learning maximizes the operation of the network while decreasing unnecessary communication losses.

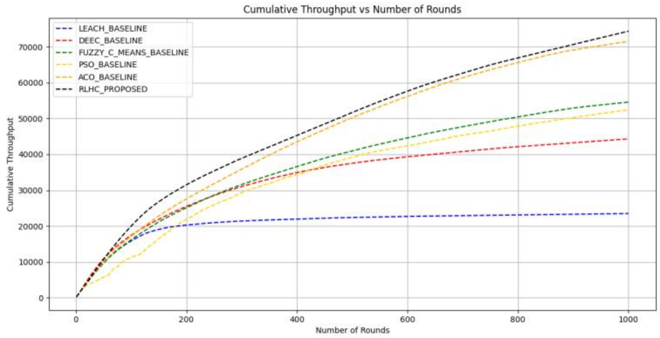


Figure 11. Cumulative throughput during network operation

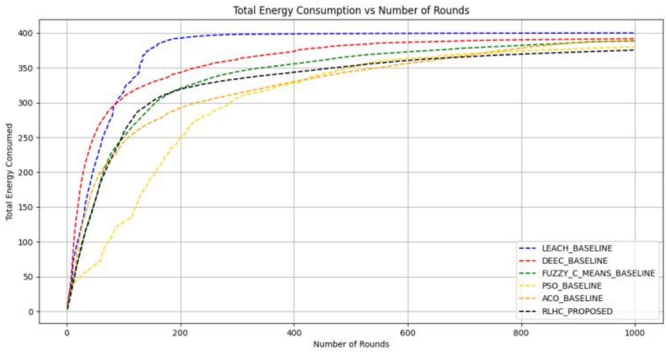


Figure 12. Total energy consumption during network operation

In Figure 12, we show the effectiveness of each protocol in utilizing the available energy resources. Here, we begin with the two protocols LEACH and DEEC, which consume energy quickly, showing total consumption of almost 400 J around the 300th round, leading to unanticipated failures of nodes. In contrast, the FCM and PSO protocols demonstrated slower and more controlled energy consumption towards the end of the simulation, with total consumption of around 370 J and 380 J. The ACO protocol showed even better energy use which around 365 J by round 1000. The protocol we proposed, RLHC consumed the least energy overall with consumption of around 360 J after the 1000th round which indicates a balanced energy utilization.

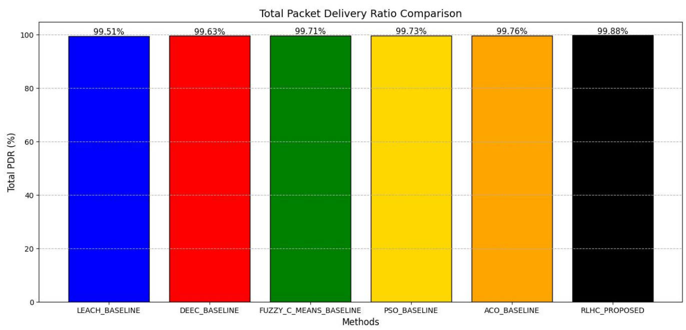


Figure 13. Total PDR during network operation

The comparative analysis of the Total Packet Delivery Ratio (PDR) provided in Figure 13 showcases the performance of different algorithms including LEACH, DEEC, Fuzzy C-Means, PSO, ACO and the new proposed RLHC method. Overall, the results indicate that all algorithms have a high PDR, which reflects that the capability of data delivery is reliable within the network. However, there is an increase in the proposed RLHC method which records the highest PDR of 99.88%, followed closely by ACO with 99.76%, PSO with 99.73%, Fuzzy C-Means with 99.71%, DEEC with 99.63% and lastly LEACH with the lowest recorded PDR of 99.51%. The improved PDR in the RLHC protocol signifies that it dynamically optimizes communication paths while minimizing packet loss caused by energy depletion or transmission failures.

Figure 14 compares the operational lifetime of the network for the same set of algorithms, measured in terms of the total number of rounds completed before the network becomes non-functional. The proposed RLHC approach achieves the longest operational duration of 7128 rounds outperforming all other algorithms.

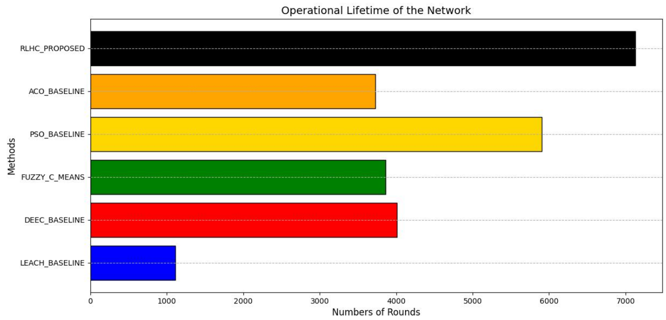


Figure 14. Operational lifetime of the network

Table 3: Comparison of overall performance

Method	LEACH	DEEC	FCM	PSO	ACO	RLHC
Network Lifetime (rounds)	1110	4008	3862	5904	3730	7128
Energy Consumption (J)	400	390	380	380	385	360
Packet Delivery Ratio (%)	99.51	99.63	99.73	99.71	99.76	99.88
Throughput(bits)	23000	44000	55000	53000	71000	75000
First Node Dead (FND)	8	12	4	9	14	21
Half Node Dead (HND)	83	81	103	257	120	145
Last Node Dead (LND)	1110	4008	3862	5904	3730	7128
Energy Model	Homogeneous	Homogeneous	Homogeneous	Hetrogeneous	Hetrogeneous	Hetrogeneous

As shown in Table 3, the RLHC protocol shows noticeable improvements over earlier clustering and optimization based routing techniques in nearly every important performance metric. RLHC achieves the highest network lifetime (7128 rounds), lowest energy consumption (360 J) superior packet delivery ratio (99.88%) and highest throughput (75000), which demonstrating its ability to maintain reliable communication and balanced energy utilization. Also, the delayed occurrence of the first and last node deaths highlights the protocol's enhanced stability and robustness in heterogeneous environments. These results collectively validate the effectiveness of the RLHC in WSNs.

VI. Conclusion

In this study, RLHC protocol was proposed to improve the energy efficiency and lifetime of heterogeneous WSNs. By integrating Q-learning with multi layered optimization mechanisms, RLHC dynamically adapts to network conditions and optimizes CH selection and data transmission paths. Simulation results clearly demonstrate that RLHC surpasses traditional protocols such as LEACH, DEEC as well as optimized protocol built on top of traditional algorithm such as FCM, PSO and ACO in terms of network longevity, throughput and energy conservation. The proposed approach effectively balances the energy load among nodes, reduces early node deaths and sustains overall network connectivity for extended periods.

Future work will focus on extending RLHC to mobile and large scale sensor networks, incorporating Simulated Annealing (SA) for further improvement and validating performance under real time deployment scenarios.

References

[1] S. M. Chowdhury and A. Hossain, ''Different energy saving schemes in wireless sensor networks: A survey,'' Wireless Pers. Commun., vol. 114, no. 3, pp. 2043–2062, Oct. 2020.

[2] F. M. Salman, A. A. Mohammedand A. F. Mutar, "Optimization of LEACH protocol for WSNs in terms of energy efficient and network lifetime," Journal of Cyber Security and Mobility, vol. 12, no. 3, pp. 275–296, May 2023, doi: 10.13052/jcsm2245-1439.1232.

[3] G. M. T. Tamilselvan and K. Gandhimathi, ''Network coding based energy efficient LEACH protocol for WSN,'' J. Appl. Res. Technol., vol. 17, no. 1, pp. 251–267, Jun. 2019

[4] W. Heinzelman, A. Chandrakasanand H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Hawaii Int. Conf. System Sciences*, 2000, pp. 1–10, doi: 10.1109/HICSS.2000.926982.

[5] A. Chehri, R. Saadane, N. Hakemand H. Chaibi, ''Enhancing energy efficiency of wireless sensor network for mining industry applications,'' Proc. Comput. Sci., vol. 176, pp. 261–270, Jan. 2020.

[6] P. Kathiroli and K. Selvadurai, ''Energy efficient cluster head selection using improved sparrow search algorithm in wireless sensor networks,'' J. King Saud Univ.-Comput. Inf. Sci., vol. 2021, pp. 1–12, Sep. 2021.

[7] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," MIT Press, 2nd ed., 2018. [Online]. Available: http://incompleteideas.net/book/the-book-2nd.html

[8] H. Zhang, G. Liuand T. Jiang, "Parameter configuration scheme for optimal energy efficiency in LoRa-based wireless underground sensor networks," IEEE Transactions on Vehicular Technology, vol. 74, no. 6, pp. 9961–9973, Jun. 2025, doi: 10.1109/TVT.2025.3351526.

[9] B. R. Al-Kaseem, Z. K. Taha, S. W. Abdulmajeedand H. S. Al-Raweshidy, "Optimized energy–efficient path planning strategy in WSN with multiple mobile sinks," IEEE Access, vol. 9, pp. 79994–80007, Jun. 2021, doi: 10.1109/ACCESS.2021.3087086.

[10] F. M. Salman, A. A. Mohammedand A. F. Mutar, "Optimization of LEACH protocol for WSNs in terms of energy efficiency and network lifetime," Journal of Cyber Security and Mobility, vol. 12, no. 3, pp. 275–296, May 2023, doi: 10.13052/jcsm2245-1439.1232.

[11] M. Gamal, N. E. Mekky, H. H. Solimanand N. A. Hikal, "Enhancing the lifetime of wireless sensor networks using fuzzy logic LEACH technique-based particle swarm optimization," IEEE Access, vol. 10, pp. 36935–36948, Mar. 2022, doi: 10.1109/ACCESS.2022.3163254.

[12] A. Panchal and R. K. Singh, "EOCGS: Energy efficient optimumnumber of cluster head and grid head selection in wireless sensor networks," Telecommun. Syst., vol. 78, no. 1, pp. 1–13, Apr. 2021, doi:10.1007/s11235-021-00806-w.

[13] R. K. Singh, S. Verma, A. Panchal and S. Dubey, "Modified RCHLEACH (MRCH) for wireless sensor networks (WSN)," in Proc. 9th Int. Congr. Inf. Commun. Technol. (ICICT). Singapore: Springer, Jan. 2024, pp. 331–340, doi: 10.1007/978-981-97-3559 4_26.

[14] K. Yilmaz, R. Kara and F. Katircioglu, "Energy-efficient hybrid adaptive clustering for dynamic MANETs," IEEE Access, vol. 13, pp. 51319–51331, 2025, doi: 10.1109/ACCESS.2025.3552232.

[15] S. Kaviarasan and R. Srinivasan, "A novel spider monkey optimized fuzzy C-means algorithm (SMOFCM) for energy-based cluster-head selection in WSNs," Int. J. Electr. Electron. Res., vol. 11, no. 1, pp. 169–175, Mar. 2023, doi: 10.37391/ijeer.110124.

[16] S. K. Chaurasiya, S. Mondal, A. Biswas, A. Nayyar, M. A. Shah and R. Banerjee, "An energy-efficient hybrid clustering technique (EEHCT) for IoT-based multilevel heterogeneous wireless sensor networks," IEEE Access, vol. 11, pp. 25941–25958, 2023, doi: 10.1109/ACCESS.2023.3254594.

[17] A. Aleem and R. Thumma, "Hybrid energy-efficient clustering with reinforcement learning for IoT-WSNs using knapsack and K-means," IEEE Sensors Journal, vol. 25, no. 15, pp. 30047–30058, Aug. 2025, doi: 10.1109/JSEN.2025..

[18] R. Alsaqour, E. S. Ali, R. A. Mokhtar, R. A. Saeed, H. Alhumyaniand M. Abdelhaq, "Efficient energy mechanism in heterogeneous WSNs for underground mining monitoring applications," IEEE Access, vol. 10, pp. 100123–100138, Jul. 2022, doi: 10.1109/ACCESS.2022.3188654.

[19] Farahzadi, H. R., Langarizadeh, M., Mirhosseini, M., & Fatemi Aghda, S. A. (2021). An improved cluster formation process in wireless sensor network to decrease energy consumption. Wireless Networks, 27(2), 1077–1087. https://doi.org/10.1007/s11276-020-02485-y.

[20] Panchal, A., & Singh, R. K. (2021). EEHCHR: Energy Efficient Hybrid Clustering and Hierarchical Routing for Wireless Sensor Networks. Ad Hoc Networks, 123, 102692. https://doi.org/10.1016/j.adhoc.2021.102692.

[21] Al-Kaseem, B. R., Taha, Z. K., Abdulmajeed, S. W., & Al-Raweshidy, H. S. "Optimized Energy-Efficient Path Planning Strategy in WSN with Multiple Mobile Sinks." IEEE Access, vol. 9, 2021, pp. 82833-82847. https://doi.org/10.1109/ACCESS.2021.3087086

[22] Prajapati, H. K., & Joshi, R. (2022). Performance Analysis of LEACH with Deep Learning in Wireless Sensor Networks. International Journal of Electronics and Telecommunications, 68(4), 799–805. doi: 10.24425/ijet.2022.143888.

[23] Mohapatra, S., Behera, P. K., Sahoo, P. K., Bisoy, S. K., Hui, K. L., & Sain, M. (2022). Mobility Induced Multi-Hop LEACH Protocol in Heterogeneous Mobile Network. IEEE Access, 10, 132895–132907. https://doi.org/10.1109/ACCESS.2022.3228576.

[24] Gamal, M., Mekky, N. E., Soliman, H. H., & Hikal, N. A. (2022). Enhancing the Lifetime of Wireless Sensor Networks Using Fuzzy Logic LEACH Technique-Based Particle Swarm Optimization. IEEE Access, 10, 36935-36948. https://doi.org/10.1109 /ACCESS.2022.3163254.

[25] Chekuri, S. & Bhandari, B. N. (2023). Cluster Based Energy Efficient Routing Protocol using SA-LEACH to Wireless Sensor Networks. Indian Journal of Science and Technology, 16(7): 492-500. DOI: 10.17485/IJST/v16i7.425.

[26] Abose, T. A., Tekulapally, V., Megersa, K. T., Kejela, D. C., Daka, S. T., & Jember, K. A. (2024). Improving wireless sensor network lifespan with optimized clustering probabilities, improved residual energy LEACH and energy efficient LEACH for corner-positioned base stations. Heliyon, 10(14), e34382. https://doi.org/10.1016/j.heliyon.2024.e34382.

[27] El Khediri, S., Selmi, A., Khan, R. U., Moulahi, T., & Lorenz, P. (2024). Energy-Efficient Cluster Routing Protocol for Wireless Sensor Networks Using Hybrid Metaheuristic Approaches. Ad Hoc Networks, 158, 103473. https://doi.org/10.1016/j.adhoc.2024.103473.

[28] Zhu, B., Bedeer, E., Nguyen, H. H., Barton, R., & Henry, J. (2024). Improved Soft-k-Means Clustering Algorithm for Balancing Energy Consumption in Wireless Sensor Networks. arXiv preprint arXiv:2403.15700.

[29] Tabatabaei, S. (2025). New Energy Efficient Management Approach for Wireless Sensor Networks in Target Tracking Using Vortex Search Algorithm. Heliyon, 11(5), e42867. https://doi.org/10.1016/j.heliyon.2025.e42867

[30] H. H. El-Sayed and Z. M. Hashem, "Comparison of the new versionof DEEC protocol to extend WSN lifetime," EURASIP J. WirelessCommun. Netw., vol. 2023, no. 1, p. 56, Jul. 2023, doi: 10.1186/s13638-023-02265-0.

[31] B. Zeng, J. Deng, Y. Dong, X. Yang, L. Huang and Z. Xiao,"A whale swarm-based energy efficient routing algorithm for wireless sensor networks," IEEE Sensors J., vol. 24, no. 12, pp. 19964–19981, Jun. 2024, doi: 10.1109/JSEN.2024.3390424.

[32] A. Verma, S. Kumar, P. R. Gautam, T. Rashid and A. Kumar, "Fuzzy logic based effective clustering of homogeneous wireless sensor networks for mobile sink," IEEE Sensors J., vol. 20, no. 10, pp. 5615–5623, May 2020, doi: 10.1109/JSEN.2020.2969697.

[33] M. U. Younus, M. K. Khan and A. R. Bhatti, "Improving the softwaredefined wireless sensor networks routing performance using reinforcement learning," IEEE Internet Things J., vol. 9, no. 5, pp. 3495–3508, Mar. 2022, doi: 10.1109/JIOT.2021.3102130.

[34] H. Dutta, A. K. Bhuyan and S. Biswas, "Contextual deep reinforcementlearning for flow and energy management in wireless sensor andIoT networks," IEEE Trans. Green Commun. Netw., vol. 8, no. 3, pp. 1233–1244, Sep. 2024, doi: 10.1109/TGCN.2024.3358230.

[35] S. Mohapatra, P. K. Behera, P. K. Sahoo, S. K. Bisoy, K. L. Huiand M. Sain, "Mobility induced multi-hop LEACH protocol in heterogeneous mobile network," *IEEE Access*, vol. 10, pp. 132895–132907, Dec. 2022, doi: 10.1109/ACCESS.2022.3228576.

[36] Q. Zhu, X. Wu, J. Wangand Z. Xu, "Deep reinforcement learning-based UAV trajectory planning for energy-efficient wireless sensor networks," IEEE Transactions on Vehicular Technology, vol. 69, no. 12, pp. 15646–15655, Dec. 2020, doi: 10.1109/TVT.2020.3035872.