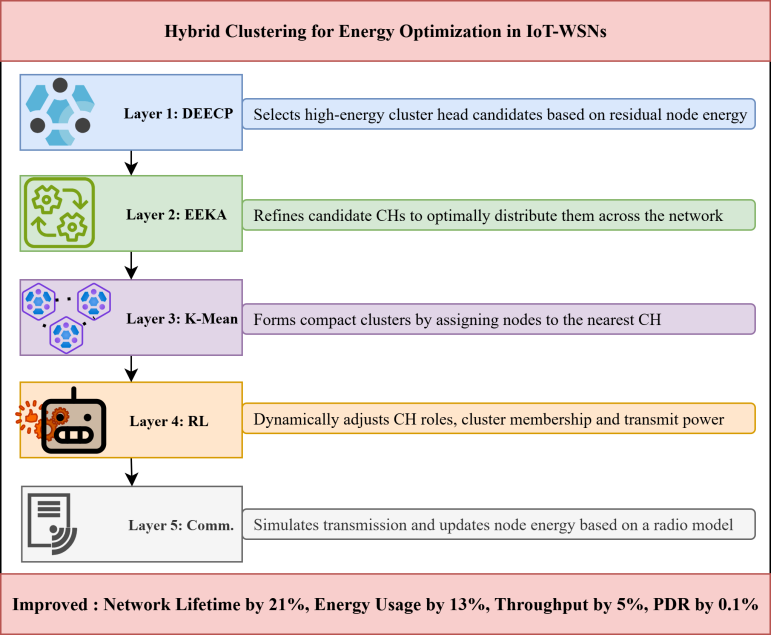
Reinforcement Learning Driven Hybrid Clustering for Energy Optimization in IoT-WSNs

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

***Abstract—*In the realm of the Internet of Things (IoT), Wireless Sensor Networks (WSNs) serve as a foundational technology, enabling diverse applications such as urban infrastructure management, industrial automation and environmental monitoring. Despite their widespread adoption, achieving energy efficiency and adaptive clustering remains a major challenge in prolonging the operational lifetime of WSNs. Traditional clustering algorithms such as Low-Energy Adaptive Clustering Hierarchy (LEACH) and Distributed Energy-Efficient Clustering (DEEC) often suffer from uneven energy consumption and static decision-making, limiting their scalability under dynamic network conditions. To address these limitations, this paper proposes a Reinforcement Learning-Driven Hybrid Clustering (RLHC) framework that integrates DEEC, the Energy-Efficient Knapsack Algorithm (EEKA) and K-Means with Q-learning-based adaptive optimization. In the proposed method, DEEC identifies high-energy cluster head (CH) candidates, EEKA ensures energy balance and uniform spatial distribution of CHs, K-Means forms compact clusters to minimize intra-cluster distances and the Q-learning agent dynamically learns optimal adjustment strategies by observing network states defined by residual energy, cluster load and packet delivery ratio (PDR) and executes actions such as CH switching, member reassignment and transmission power tuning. Through continuous interaction with the environment, the agent converges toward energy-optimal configurations. Simulation results demonstrate that the proposed RLHC method significantly enhances network lifetime, PDR and energy balance compared to optimized algorithms built on top of LEACH and DEEC. The improvements include a 21% increase in network lifetime, 13% reduction in energy consumption, 5% higher throughput and 0.1% improvement in PDR. This hybrid intelligence approach provides a scalable and adaptive solution for next-generation IoT-based WSN applications.**

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

1. **INTRODUCTION**

Wireless Sensor Networks (WSNs) represent one of the foundational technologies driving the rapid evolution of the Internet of Things (IoT) paradigm. These networks consist of spatially distributed sensor nodes that cooperatively monitor and record environmental conditions and transmit the collected information to a central Base Station (BS) or sink node for further processing and analysis. The ability of WSNs to provide real-time, context-aware data makes them indispensable in modern intelligent systems. They find extensive applications in smart city governance, industrial automation, military surveillance etc [1], [2].

Through efficient data acquisition and communication, WSNs serve as the backbone of IoT ecosystems, enabling data-driven decision-making, predictive analytics and autonomous operations. Despite their vast potential, WSNs face critical constraints in terms of energy efficiency, communication reliability and network longevity. The sensor nodes are typically powered by small, non-rechargeable batteries and deployed in remote or harsh environments, where human intervention for maintenance or battery replacement is highly impractical. Consequently, energy depletion of individual nodes leads to node death, which in turn causes network partitioning, packet loss and degradation in overall system performance. This problem is particularly exist in applications requiring continuous and long-term monitoring such as health diagnostics, disaster prediction and precision agriculture, where uninterrupted operation and consistent data flow are imperative.

To mitigate such challenges, researchers have proposed numerous energy aware routing and clustering protocols aimed at optimizing energy utilization and balancing power consumption among nodes. Among these, clustering based routing has emerged as one of the most effective strategies for achieving energy efficiency. In this approach, sensor nodes are grouped into clusters, each governed by a Cluster Head (CH) that aggregates data from member nodes and transmits it to the BS. Representative clustering protocols include the Low-Energy Adaptive Clustering Hierarchy (LEACH), Stable Election Protocol (SEP) and Distributed Energy-Efficient Clustering (DEEC) [3]–[5]. These protocols primarily aim to reduce communication overhead and distribute the energy load more evenly across 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.

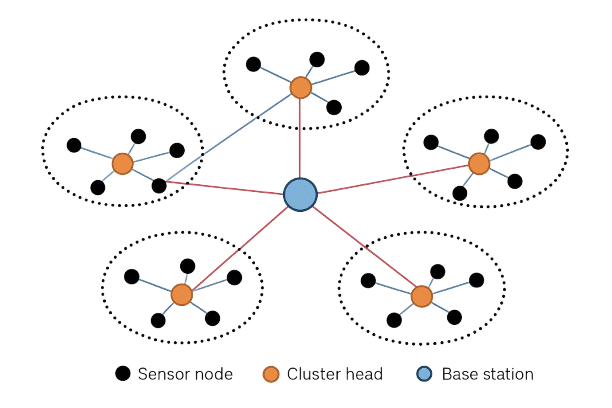
Figure 1: Wireless Sensor Network in Mines

Figure 1, illustrates a typical communication model of a WSN, where sensor nodes are randomly deployed across the sensing region and relay their sensed data to the base station through CH.

However, traditional clustering protocols still exhibit limitations in adaptability, scalability and dynamic optimization. Most 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, energy dynamics, node heterogeneity, communication interference and spatial distribution can vary significantly 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. Additionally, existing methods rarely incorporate adaptive learning or feedback mechanisms that can respond to dynamic environmental conditions or operational uncertainties. Thus the design of an intelligent and adaptive clustering mechanism becomes crucial for achieving long-term sustainability.

To overcome the aforementioned limitations, this paper introduces a Reinforcement Learning-Driven Hybrid Clustering (RLHC) approach that fuses intelligent optimization techniques across multiple layers of the clustering process. The proposed RLHC framework integrates DEEC, EEKA, K-Means clustering and RL to form an adaptive and energy-aware communication architecture. 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, determining the optimal number of CHs and selecting the best CHs to improve spatial balance to reduce inter-cluster communication overhead.. The 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, node density, cluster load and communication delay. 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. This includes the free-space (d²) and multipath (d⁴) channel models for short and long distances, respectively.

By integrating these five layers, the RLHC framework achieves adaptive energy balancing, enhanced throughput, improved PDR and prolonged network lifetime. This hybrid intelligent model demonstrates the potential to bridge the gap between traditional deterministic routing methods and modern ML-driven adaptive optimization techniques. It provides a robust, scalable and self-evolving solution for the future of IoT-enabled WSNs, paving the way for more reliable and sustainable sensor network deployments in real-world environments.

Our contributions can be summarized as follows:

1. A five-layer RLHC model integrating DEEC, EEKA, K-Means, RL and a communication layer for intelligent energy management in WSNs.
2. Energy-aware optimal clustering and RL based adaptive cluster refinement mechanism that dynamically optimizes cluster configurations in response to real-time variations in node energy and network conditions.
3. Performance enhancement as compared to the optimized algorithms built on top of traditional clustering protocols in terms of network lifetime, throughput, energy consumption and packet delivery ratio.

The remainder of this paper is organized as follows:

Section II presents a comprehensive literature review of existing clustering protocols and reinforcement learning approaches in WSNs. Section III outlines the theoretical and technical background relevant to energy-efficient communication and clustering mechanisms. Section IV describes the methodology of proposed Reinforcement Learning-Driven Hybrid Clustering (RLHC) framework in detail which highlighting its five-layer architecture. Section V discusses the simulation environment, performance evaluation metrics and comparative analysis of the proposed RLHC against existing protocols. Finally, Section VI concludes the paper and outlines potential directions for future research and real-world implementation.

1. **RELATED WORKS**

WSNs have become an integral component of modern IoT systems, enabling applications ranging from environmental monitoring and industrial automation to smart cities. A critical challenge in WSNs is energy

efficiency, as sensor nodes typically operate on limited battery power and network longevity directly depends on efficient energy utilization. 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 clustering and energy-efficient WSNs, highlighting their strengths, limitations and the evolution toward hybrid RL-based approaches.

1. *Classical Clustering Protocols*

Clustering is widely recognized as a foundational approach for energy-efficient WSNs. The LEACH protocol is among 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 base station and organizes nodes into energy-efficient clusters. However, it has several limitations, including uneven CH distribution, lack of consideration for residual energy and poor performance in heterogeneous or large-scale networks. These drawbacks motivated the development of variants that incorporate energy-awareness and residual energy into CH selection. Protocols such as LEACH-RLC and ReLeC enhance LEACH by integrating RL for adaptive CH selection. These methods learn network states and make dynamic decisions about CH election to improve energy efficiency and prolong network lifetime. Despite promising results, RL-based variants can impose computational and communication overhead that may be unsuitable for resource-constrained IoT nodes. Additionally, RL models like ReLeC may face convergence issues in large-scale or highly dynamic WSNs, limiting their practical applicability.

1. *Heterogeneity-Aware Protocols*

Recognizing the limitations of classical homogeneous protocols, researchers have introduced heterogeneity-aware approaches such as SEP (Stable Election Protocol) and DEEC. These protocols consider nodes with different initial energy levels and select CHs based on residual energy and network wide energy distribution. This ensures that higher energy nodes are more likely to assume CH roles, balancing the energy load across the network and preventing premature node failures. DEEC, in particular, extends LEACH by incorporating energy heterogeneity and energy based CH selection, improving network stability and prolonging lifetime. Enhancements such as EEKA further optimize CH selection by considering node centrality and spatial distribution, aiming for a uniform CH spread that reduces intra-cluster communication distances and overall energy consumption. These protocols effectively address energy imbalance but often rely on pre-defined thresholds or heuristic rules, limiting their adaptability to dynamic network topologies or sudden changes in energy states.

1. *Metaheuristic and Optimization-Based Approaches*

To achieve better CH selection and network optimization, metaheuristic algorithms such as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO) and Genetic Algorithms (GA) have been extensively applied in WSNs. PSO minimizes intra-cluster distances and identifies optimal CH positions by simulating social behavior among particles. ACO leverages pheromone based path selection to optimize cluster formation and routing, while GA evolves a population of CH candidates to optimize energy efficiency and load balancing. These approaches have demonstrated significant improvements in network lifetime and communication efficiency. However, metaheuristic based methods typically operate offline or rely on iterative convergence, which makes them less suitable for networks with dynamic topologies, mobile nodes, or rapidly changing energy states.

1. *Fuzzy Logic and Multi-Criteria Clustering*

Fuzzy logic-based clustering introduces multi-criteria decision-making for CH selection, considering parameters such as residual energy, node density, distance to the cluster center and network traffic. Protocols like MRCH (Modified RCH-LEACH) utilize fuzzy rules to determine CH candidacy, improving stability, packet delivery ratio and energy efficiency. Fuzzy-based approaches provide a flexible framework for handling uncertainties in sensor networks, enabling adaptive cluster formation under varying network conditions. Moreover, these methods can involve computationally intensive calculations, limiting their deployment on low-power sensor nodes.

1. *Reinforcement Learning in WSNs*

RL has emerged as a powerful tool for dynamic and adaptive WSN management. RL-based approaches model the network as an environment, where nodes or CHs act as agents that learn optimal actions to maximize long-term rewards, such as energy efficiency or network lifetime. Q-learning, SARSA and Deep RL have been explored to optimize CH selection, cluster reorganization and routing decisions. For example, EER-RL improves energy efficiency and prolongs network lifetime by dynamically adjusting CH roles based on learned energy patterns. Similarly, Q-learning LEACH models enhance adaptability to changing network topologies. Despite their effectiveness, RL-based methods often require centralized training, global knowledge, or extensive exploration, which may limit scalability in large or highly dynamic WSN deployments.

1. *Hybrid Clustering Approaches*

Hybrid clustering methods aim to combine the strengths of classical, metaheuristic, fuzzy and RL approaches. These methods address multiple challenges simultaneously, including CH optimization, energy balancing, spatial uniformity and adaptability. Multi-layered hybrid models often integrate energy-aware CH selection, K-means or fuzzy-based spatial clustering and RL-based adaptive decision-making to maximize network efficiency. Recent work has highlighted the efficacy of such hybrid frameworks. Protocols like EOCGS determine the optimal number of cluster and grid heads to balance energy consumption. These studies show that intelligent hybrid methods can outperform traditional approaches in terms of network lifetime, energy balance and adaptability.

Despite substantial progress in WSN clustering, existing methods continue to face several challenges. RL based and other meta heuristic approaches often introduce significant computational overhead, making them unsuitable for low-power IoT nodes. Many protocols also struggle to maintain efficiency in large-scale or highly heterogeneous networks. Classical methods frequently fail to prevent early node depletion, resulting in network partitioning. Furthermore, most existing solutions rely on static or pre-defined strategies, which are ill-equipped to handle dynamic topologies, node mobility, or sudden energy fluctuations. Excessive intra-cluster communication further accelerate energy depletion, limiting overall network lifetime.

**TABLE 1**. Summary of energy efficiency studies in WSN

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Paper & Year** | **CH Selection Criteria** | **Optimization Technique Used** | **Energy Efficiency Improvement (%)** | **Network Lifetime Improvement (%)** | **Limitations (Research gap)** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |