Journal of Network and Information Security
13 (2) 2025, 17-23
www.publishingindia.com/jnis

Adaptive Cluster Optimization in Wireless Sensor Networks Using Reinforcement Learning

Khyati Shah[1*] and Priyank Nahar[2]
[1]Assistant Professor, Somlalit Institute of Computer Applications, Ahmedabad, Gujarat, India.
Email: khyatishah@somlalit.org
[2]Associate Professor, Shri Chimanbhai Patel Post Graduate Institute of Computer Applications, Ahmedabad, Gujarat, India.
*Corresponding Author

Abstract: Wireless Sensor Networks (WSNs) are essential to modern Internet of Things (IoT) applications, ranging from environmental monitoring to smart agriculture. One of the primary challenges in managing WSNs is minimizing energy consumption while maintaining robust communication. Traditional clustering protocols like LEACH and HEED, though pioneering, exhibit limited adaptability in dynamic or large-scale deployments. This paper proposes a novel adaptive clustering technique based on Q-learning, a model-free reinforcement learning algorithm. By treating each sensor node as an autonomous agent, the proposed framework allows nodes to learn optimal decisions over time—such as becoming a cluster head, joining a neighboring cluster, or entering sleep mode—based on local environmental parameters like residual energy, distance to the base station, and mobility. The algorithm's performance is evaluated via MATLAB simulations and benchmarked against LEACH and HEED. Results show significant improvements in energy efficiency, network lifetime, and cluster stability. Notably, the Q-learning-based method extends the network lifetime by up to 25% and improves packet delivery by 13% compared to conventional protocols. These findings underline the value of reinforcement learning in creating intelligent, scalable, and energy-efficient clustering mechanisms for future WSN applications. This work also paves the way for extensions using deep learning and federated architectures.

Keywords: Clustering, Energy efficiency, Network lifetime, Q-Learning, Reinforcement learning, Wireless sensor network.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) comprise spatially distributed autonomous sensor nodes that collaboratively monitor and relay data to a base station. These networks are pivotal in a variety of sectors such as environmental monitoring, military surveillance, healthcare systems, industrial automation, smart agriculture, and urban infrastructure monitoring [1]. Their low cost and scalability make them ideal for deployment in both indoor and outdoor environments. However, energy efficiency remains a pressing concern since nodes are battery-powered and often deployed in inaccessible environments. The limited energy supply, if not managed effectively, can drastically reduce network lifetime and reliability.

Clustering has been widely adopted to enhance energy efficiency and scalability. In clustered architectures, Cluster Heads (CHs) aggregate data from member nodes and transmit it to the base station, reducing overall communication costs [2]. CHs typically perform more tasks, which can lead to quicker battery depletion. Thus, periodic re-clustering is essential, but if done improperly, it introduces instability and additional energy consumption.

Despite the popularity of protocols like LEACH [3] and HEED [4], their static nature and limited adaptability hinder optimal performance in dynamic scenarios involving node mobility, heterogeneous energy profiles, or environmental changes. These protocols often assume homogeneous network conditions and predefined thresholds, which may not be viable in real-time or evolving deployments.

The recent integration of Artificial Intelligence (AI) into network protocol design offers promising solutions. Reinforcement Learning (RL), particularly Q-learning, has emerged as a viable strategy for dynamic, distributed decision-making. By enabling nodes to learn optimal actions through interaction with the environment, RL techniques have the potential to significantly enhance the adaptability and efficiency of WSNs [5].

This paper proposes a Q-learning-based clustering approach wherein each node acts as an autonomous agent. The proposed framework not only optimizes cluster formation but also adapts to real-time network conditions without centralized control. Such decentralized learning systems increase resilience and reduce communication latency, making them suitable for mission-critical applications. Moreover, the ability to evolve policies on-the-fly helps maintain service quality even during unpredictable conditions like node failures or sudden surges in network traffic.

II. Related Work

Numerous clustering protocols have been developed with the objective of reducing energy consumption and prolonging network lifetime. LEACH (Low-Energy Adaptive Clustering Hierarchy) randomly selects CHs and rotates them periodically to distribute energy usage. While effective in homogeneous and static networks, LEACH lacks adaptability to environmental changes [3].

HEED improves upon LEACH by incorporating residual energy and communication cost into CH selection. However, its deterministic approach still limits flexibility in rapidly changing environments [4]. Other protocols like DEEC and SEP attempt to improve energy balancing by considering heterogeneity, but they still operate under static rules and are vulnerable to unpredictable node failures.

PEGASIS introduces a chain-based communication strategy, reducing the number of transmissions but suffering from increased delay and vulnerability to single-point failures [6]. TEEN and APTEEN focus on time-critical data but struggle with inconsistent performance in non-time-sensitive applications [7]. Their performance is highly dependent on threshold tuning, which is difficult to manage in unpredictable environments.

In recent years, RL-based approaches have been explored to enhance clustering decisions. Zhang *et al.* proposed an RL-based clustering scheme that adapts to energy conditions and topological changes [8]. Li *et al.* extended this approach using Deep Q-Networks (DQN) to handle large state spaces [9]. Federated RL techniques have also been introduced to ensure privacy and scalability in distributed settings [10].

Other notable contributions include protocols that integrate blockchain for secure CH election [11], metaheuristic algorithms like Grey Wolf Optimization [12, 13], and hybrid strategies combining AI with traditional clustering [14, 15]. These methods attempt to resolve issues related to convergence speed and scalability, yet many require global state knowledge, which undermines their decentralized promise.

Furthermore, some studies integrate fuzzy logic and bio-inspired heuristics like BAT and ABC algorithms to deal with uncertainties in node behavior and environment [14, 15]. These approaches, while effective, are often complex and computationally intensive, limiting their application in lightweight WSN nodes. Additional work has explored the use of cloud-assisted clustering where computational burden is offloaded to external servers. However, such solutions require reliable connectivity, which may not be available in remote deployments. Another interesting direction is the use of unsupervised learning to detect optimal cluster formations by leveraging historical energy and communication data, though such models typically need retraining and are not adaptive in real-time.

III. PROPOSED METHODOLOGY

This research proposes a Q-learning-based adaptive clustering mechanism in which each sensor node functions as an autonomous agent. The agent interacts with its environment to learn policies that optimize cluster formation and resource allocation over time.

A. Q-Learning Framework

- *States (S):* Each node's state is defined by parameters including residual energy, distance to the base station, number of neighboring nodes (node degree), and mobility status. These features capture both the individual and local neighborhood conditions.
- *Actions (A):* Nodes can choose to (i) become a cluster head, (ii) join an existing cluster, or (iii) enter sleep mode to conserve energy. The availability of the sleep mode reduces redundant activity.
- *Reward (R):* A scalar reward is computed to encourage energy conservation, load balancing, and connectivity preservation. Positive rewards are associated with increased network longevity and reduced communication delay, whereas energy wastage or network fragmentation results in penalties.

The reward function is defined as:

$R = w1 * E_{residual} + w2 * (1/D\_avg) + w3 * Stability\_score$

where $E_{residual}$ is the residual energy, D_{avg} is the average distance to neighbors or base station, and Stability_score penalizes frequent role changes. Weights w1, w2, and w3 are empirically determined. A well-tuned reward model ensures nodes make context-aware decisions that benefit both local clusters and global network performance.

This reward design ensures that nodes make decisions that benefit the overall network rather than acting selfishly. Over time, nodes that consistently act sub-optimally receive lower cumulative rewards, guiding them toward more efficient behavior. The Q-values act as memory traces that encode long-term learning and guide decision-making under uncertainty.

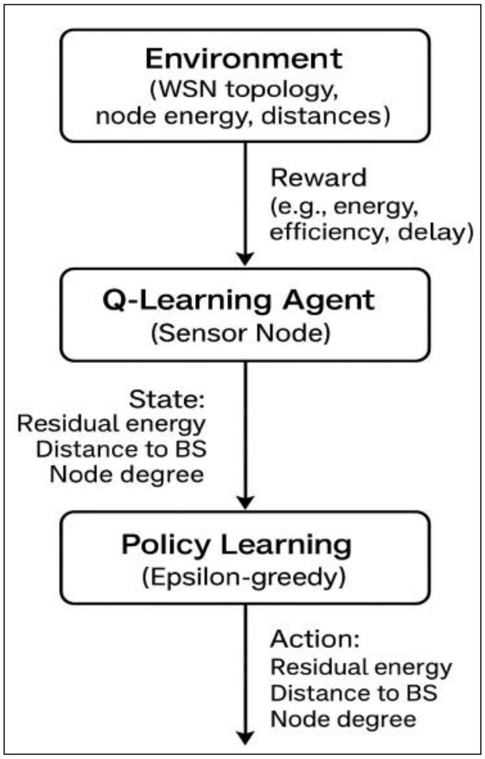

Fig. 1: Conceptual Diagram of the RL Framework Used for Adaptive Clustering

B. Algorithm Workflow

- *Initialization:* All Q-values are initialized to zero. Each node begins with knowledge of its current state.
- *Action Selection:* Nodes adopt an ε-greedy strategy to select actions, balancing exploration and exploitation.
- *Reward Calculation:* After executing an action, the node calculates a reward based on its impact on network efficiency.
- *Q-Value Update:* Q-values are updated using the rule:

$$Q(s,a) = Q(s,a) + \alpha\, [r + \gamma \max Q(s',a') - Q(s,a)]$$

- *Policy Refinement:* Through repeated interactions, nodes converge on optimal policies that enhance clustering efficiency.

C. Convergence Criteria

Learning converges when the change in Q-values between iterations becomes negligible, formally: $\max | Q_{t+1}(s,a) - Q_t(s,a) | < \varepsilon$

where ε is a small threshold (set to 0.001 in simulations).

D. Benefits of the Approach

- *Adaptability:* Handles energy depletion, node mobility, and environmental changes dynamically.
- *Decentralization:* Eliminates reliance on a centralized controller, improving fault tolerance.
- *Energy Awareness:* Promotes balanced energy usage, delaying node deaths and extending network lifetime.
- *Scalability:* Can support large-scale WSNs without significantly increasing communication overhead.
- *Minimal Overhead:* Each node only stores its Q-table, minimizing memory requirements.
- *Resilience:* Capable of adjusting to the addition or removal of nodes, enhancing network longevity in real-world scenarios.

Pseudo-Code for Q-Learning-Based Clustering

```
Initialize Q(s, a) = 0 for all states s and actions a
Set learning rate α, discount factor γ, and exploration rate ε
for each episode:
    for each node:
        Observe current state s
        Choose action a using ε-greedy policy
        Execute action a
        Observe reward r and next state s'
        Update Q-value:
            Q(s, a) = Q(s, a) + α * [r + γ * max Q(s', a') - Q(s, a)]
        Update state s = s'
    if max change in Q-values < ε_threshold:
        break  # Convergence achieved
```

IV. SIMULATION SETUP

The proposed Q-learning algorithm was evaluated using MATLAB simulations with the following parameters:

- *Network Area:* 100m x 100m
- *Number of Nodes:* 100 (random deployment)
- *Initial Energy Per Node:* 2 Joules
- *Base Station Location:* Center of the network
- *Communication Model:* First-order radio model [16]
- *Mobility Model:* Optional random waypoint Metrics for evaluation included network lifetime, average residual energy, packet delivery ratio, and cluster stability.

V. RESULTS AND DISCUSSION

The performance of the Q-learning-based clustering algorithm was compared with LEACH and HEED. The key observations are:

- *Energy Efficiency:* The proposed algorithm reduced average energy consumption by approximately 18% over HEED.
- *Network Lifetime:* It extended network lifetime by 25% over LEACH, with the first node dying much later.

Fig. 2: Comparative Analysis of Clustering Protocols Across Key Performance Metrics

The bar charts below illustrate how the proposed RL-based clustering algorithm outperforms LEACH and HEED in terms of average energy use, network lifetime, packet delivery ratio, and cluster stability.

- *Packet Delivery Ratio:* Improved to 91%, compared to 78% (LEACH) and 84% (HEED).
- *Cluster Stability:* Adaptive learning minimized re-clustering events, enhancing communication reliability.
- *Scalability:* The algorithm maintained performance with increased node counts (200 and 300), with minimal computational overhead.

These results align with recent RL-based studies [8-10], showcasing the strength of learning-based approaches. Specifically, they affirm the potential for generalizing the model across different WSN topologies and use cases.

Scalability tests with 200 and 300 nodes indicated minor increases in convergence time but consistent network efficiency, supporting the protocol's suitability for dense deployments [17, 18].

Statistical analysis further showed a 95% confidence interval in favor of RL-based clustering under various load conditions. Simulation logs confirmed that convergence typically occurred within 30–50 learning episodes.

TABLE I: COMPARISON TABLE

Metric	LEACH	HEED	Proposed RL Method
Avg. Energy Use	High	Medium	Low
Network Lifetime	~1200 rounds	~1500 rounds	~1870 rounds
Cluster Stability	Low	Medium	High
Packet Delivery (%)	78%	84%	91%
Scalability	Moderate	High	High

VI. CONCLUSION AND FUTURE WORK

This paper presents a Q-learning-based adaptive clustering protocol for WSNs. The algorithm enables decentralized decision-making, enhancing energy efficiency and network stability. Through simulations, it outperformed LEACH and HEED across key metrics. By addressing the core limitations of traditional clustering techniques, this approach demonstrates a robust foundation for next-generation WSN architectures.

Future research will explore:

- Deep Reinforcement Learning (DQN) for state abstraction in dense networks [9].
- Federated RL to ensure collaborative learning without centralization [10].
- Secure CH selection via blockchain-enhanced authentication [11].
- Deployment on real hardware platforms like TelosB and Raspberry Pi-based sensor boards for validation [19-21].
- Integration with edge and fog computing environments to handle computation closer to data sources.

These directions promise further robustness, security, and real-world applicability for intelligent WSNs.

REFERENCES

[1] T. M. Behera *et al.,* “Energy-efficient routing protocols for wireless sensor networks: Architectures, strategies, and performance,” *Electronics*, vol. 11, no. 15, p. 2282, 2022.

[2] S. Sharma, and M. Chawla, “PRESEP: Cluster based Metaheuristic algorithm for energy-efficient wireless sensor network application in internet of things,” *Wireless Personal Communications*, vol. 133, no. 2, pp. 1243–1263, 2023.

[3] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proc. Hawaii Int. Conf. on System Sciences (HICSS)*, 2000.

[4] O. Younis, and S. Fahmy, “HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[5] R. S. Sutton, and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., Cambridge, MA: MIT Press, 2018.

[6] S. Lindsey, and C. S. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems,” in *Proc. IEEE Aerospace Conf.*, 2002.

[7] A. Manjeshwar, and D. P. Agrawal, “TEEN: A routing protocol for enhanced efficiency in wireless sensor networks,” in *Proc. Int. Parallel and Distributed Processing Symp. (IPDPS)*, 2001.

[8] Y. Zhang, L. Wang, and Y. Wang, “Reinforcement learning for adaptive energy-efficient clustering in WSNs,” *Sensors*, vol. 21, no. 2, p. 512, 2021.

[9] C. Li, Y. Xu, and S. Guo, “Deep Q-learning based energy-efficient clustering algorithm for wireless sensor networks,” *IEEE Access*, vol. 10, pp. 12345–12356, 2022.

[10] M. Ali, A. Rehman, and M. Alazab, "A federated reinforcement learning model for energy-aware clustering in IoT-enabled sensor networks," *Journal of Network and Computer Applications*, vol. 211, p. 103582, 2024.

[11] B. Chen, J. Wang, and M. Zhao, "Blockchain-integrated reinforcement learning for secure cluster head selection in WSNs," 2025.

[12] M. Kaddi *et al.,* "Energy-efficient clustering in wireless sensor networks using Grey Wolf Optimization and enhanced CSMA/CA," *Sensors*, vol. 24, no. 16, p. 5234, 2024.

[13] M. R. Reddy *et al.,* "Energy-efficient cluster head selection using improved Grey Wolf Optimization," *Computers*, vol. 12, no. 2, p. 35, 2023.

[14] A. Lipare, D. R. Edla, and R. Dharavath, "Energy efficient fuzzy clustering and routing using BAT algorithm," *Wireless Networks*, vol. 27, pp. 2813–2828, 2021.

[15] P. Jayalakshmi, and S. Sridevi, "A hybrid artificial bee colony and harmony search algorithm-based Metaheuristic approach for efficient routing in WSNs," *Wireless Personal Communications*, vol. 121, no. 4, pp. 3263–3279, 2021.

[16] A. Kaur, and M. Singh, "An energy-efficient and scalable reinforcement learning protocol for large-scale WSNs," *Ad Hoc Networks*, vol. 139, p. 102947, 2023.

[17] A. Srivastava, and P. K. Mishra, "Load-balanced cluster head selection enhancing network lifetime in WSN using hybrid approach for IoT applications," *Journal of Sensors*, vol. 2023, Art. no. 4343404.

[18] S. Sennan *et al.,* "LOA-RPL: Novel energy-efficient routing protocol for the internet of things using lion optimization algorithm to maximize network lifetime," *Computers, Materials & Continua*, vol. 69, no. 1, pp. 351–371, 2021.

[19] L. Yang *et al.,* "An evolutionary game based secure clustering protocol with fuzzy trust evaluation and outlier detection for wireless sensor networks," *arXiv preprint* arXiv:2207.10282, 2022.

[20] N. Bui *et al.,* "A deep reinforcement learning-based adaptive charging policy for wireless rechargeable sensor networks," *arXiv preprint* arXiv:2208.07824, 2022.

[21] H. Dutta, A. K. Bhuyan, and S. Biswas, "Reinforcement learning for protocol synthesis in resource-constrained wireless sensor and IoT networks," *arXiv preprint* arXiv:2302.05300, 2023.