Comparative Analysis of DRL Techniques for Energy-Aware Clustering in IoT-WSN

Dr. S. Rizwana[1], M. Sangeetha[2]

[1]Research Supervisor & Head, Department of Computer Science, Erode Arts and Science College (Autonomous), Erode

[2]Research Scholar, Department of Computer Science, Erode Arts and Science College (Autonomous), Erode

Abstract: ***The rapid expansion of Internet of Things (IoT)-based Wireless Sensor Networks (WSN) has introduced significant challenges in energy management, routing efficiency, scalability, and network reliability. Traditional optimization techniques often fail to adapt to the dynamic and resource-constrained nature of large-scale IoT environments. To address these limitations, recent advancements in Deep Reinforcement Learning have emerged as promising solutions for intelligent and energy-aware network optimization. This paper presents a unified comparative framework for evaluating multiple Deep Reinforcement Learning (DRL) algorithms in IoT-based WSN environments. The key performance metrics such as Energy Consumption, Packet Delivery Ratio, Latency, Throughput, and Network Lifetime are systematically evaluated. The results clearly demonstrate that federated DRL frameworks provide enhanced scalability, privacy preservation, and optimization capability, making them the most effective solution for IoT-based WSN environments.***

Keywords: Wireless Sensor Networks, Deep Reinforcement Learning, Energy-Efficient Routing, Intelligent Clustering, Resource Optimization

1. Introduction

The rapid growth of the Internet of Things has significantly increased the deployment of Wireless Sensor Networks in applications such as smart cities, environmental monitoring, industrial automation, healthcare, and agriculture [1]. In these IoT-enabled WSNs, sensor nodes are typically battery-powered and deployed in large numbers across wide geographical areas. The physical objects can be seamlessly integrated and operated through the Internet of Things. As a result, millions of devices equipped with sensors and actuators are interconnected via wired or wireless communication channels to enable data exchange and system automation. These IoT devices continuously generate massive volumes of heterogeneous data with varying formats and quality levels.

Machine Learning (ML) in IoT - Machine Learning enables IoT-based Wireless Sensor Networks to analyze sensed data and make intelligent decisions without explicit programming. In IoT-WSNs, ML techniques are widely used for tasks such as energy prediction, anomaly detection, fault diagnosis, data aggregation, and traffic classification. By learning patterns from historical sensor data, ML models can optimize routing decisions, improve cluster head selection, and enhance network lifetime. Since many IoT devices operate under resource constraints, lightweight ML algorithms are often preferred to ensure minimal computational overhead while improving system efficiency.

Deep Learning (DL) - A subset of ML [2] represented in Fig1. is based on artificial neural networks with multiple hidden layers, provides advanced capability to process large-scale and high-dimensional IoT sensing data. In IoT-based WSNs, DL models are applied for complex tasks such as traffic prediction, intrusion detection, energy consumption forecasting, and data quality enhancement. Due to their ability to automatically extract features from raw sensor data, DL techniques improve accuracy compared to traditional ML methods. However, because of computational and energy limitations in sensor nodes, DL models are typically implemented at the base station or edge server rather than directly on sensor devices. Reinforcement Learning is a decision-making approach in which an agent learns optimal actions through interaction with the environment by maximizing cumulative rewards. In IoT-based WSNs, RL is particularly useful for dynamic problems such as cluster head selection, routing optimization, transmission scheduling, and energy management. Unlike supervised learning, RL does not require labeled data and can adapt to changing network conditions. This makes RL highly suitable for energy-efficient and adaptive routing in IoT-WSNs, where network topology, traffic patterns, and node energy levels vary over time.Reinforcement learning has proved strong performance in a wide range of application areas with different levels of complexity and difficulty. However, these achievements are typically obtained in structured or well-defined environments, and the results may not always transfer directly to unpredictable or real-world situations.

Cluster-based routing [3] has been widely adopted to reduce communication overhead and improve scalability. In this approach, certain nodes are selected as Cluster Heads (CHs), which aggregate and forward data to the base station. However, improper selection of CHs leads to uneven energy consumption, early node death, network instability, and reduced overall lifetime. Traditional clustering protocols such as LEACH rely on probabilistic or threshold-based CH selection, which does not adapt effectively to dynamic IoT environments characterized by heterogeneous devices, varying traffic loads, and changing network conditions.

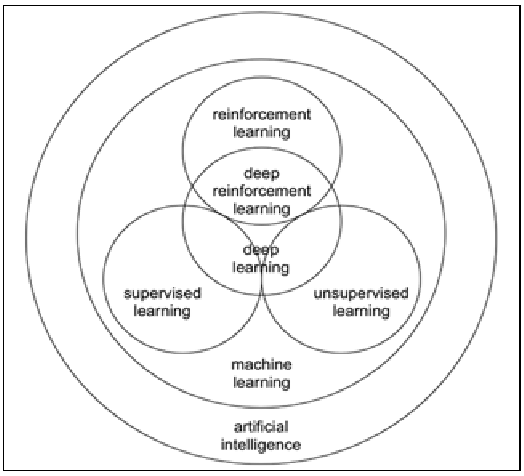

Figure 1: Conceptual Hierarchy of Artificial Intelligence, Machine Learning, and Deep Reinforcement Learning

As IoT networks scale from tens to hundreds or thousands of nodes, static and rule-based clustering mechanisms fail to provide optimal energy balancing, highlighting the need for intelligent and adaptive solutions. Deep Reinforcement Learning (DRL) has recently emerged as a powerful technique for decision-making in dynamic and complex environments. By combining reinforcement learning with deep neural networks [4], DRL enables agents to learn optimal policies through interaction with the environment without requiring explicit mathematical modelling. In IoT-based WSNs, DRL can model cluster head selection as a sequential decision-making problem. A centralized agent, typically located at the base station, observes the network state (such as residual energy, node density, distance to base station, and traffic conditions) and learns to select optimal CHs that maximize network lifetime while minimizing energy variance and communication delay.

A. Recent Developments in Deep Reinforcement Learning for IoT-WSN

Deep Reinforcement Learning (DRL) has emerged as a transformative approach for solving complex decision-making problems in dynamic and resource-constrained environments. By integrating reinforcement learning with deep neural networks, DRL enables agents to approximate optimal policies in high-dimensional state spaces, making it highly suitable for IoT-based Wireless Sensor Networks. Unlike traditional optimization algorithms, DRL can continuously learn from network interactions and adapt to changing topologies, traffic patterns, and energy conditions. Recent advancements in Deep Reinforcement Learning have significantly enhanced its applicability in complex and dynamic IoT-enabled wireless sensor networks [5]. By combining deep neural networks with reinforcement learning principles, DRL efficiently handles uncertain environments. Unlike conventional optimization techniques, DRL continuously interacts with the environment, learns optimal policies through reward feedback, and adapts its decisions based on evolving network conditions. The following highlights the major advancements and benefits observed in IoT-based applications.

- **Adaptive Long-Term Policy Learning:** DRL focuses on cumulative rewards rather than short-term gains, enabling the system to derive stable long-term optimization policies. This is particularly useful in IoT-WSNs where network conditions such as node energy levels, traffic patterns, and link quality vary over time.
- **Fast Convergence with Deep Function Approximation:** The integration of deep neural networks allows DRL to approximate complex value functions efficiently. This reduces computational complexity compared to traditional tabular reinforcement learning and accelerates convergence in high-dimensional environments.
- **Real-Time Autonomous Control:** Once trained, DRL agents can make decisions in real time with minimal latency. This enables rapid adaptation to topology changes, node failures, congestion, and environmental disturbances in IoT deployments.
- **Model-Free Learning Capability:** DRL does not require explicit mathematical modeling of the network environment. It learns optimal behavior directly through interaction, making it suitable for highly dynamic and unpredictable IoT scenarios.
- **Scalable to Large-Scale Networks:** By leveraging deep architectures, DRL overcomes the curse of dimensionality and supports large-scale IoT systems with hundreds or thousands of interconnected nodes.
- **Multi-Metric Optimization:** Modern DRL frameworks allow the design of reward functions that simultaneously consider energy efficiency, latency, throughput, reliability, and security, thereby achieving balanced network performance.
- **Integration with Emerging AI Paradigms:** Recent research explores combining DRL with federated learning, graph neural networks, and attention mechanisms to enhance distributed intelligence, privacy preservation, and contextual awareness in IoT networks.

Despite these improvements, challenges remain in reducing computational complexity, ensuring real-time implementation, and balancing exploration–exploitation trade-offs in energy-constrained sensor networks. Unlike metaheuristic algorithms that re-optimize the network at every round without learning from past experiences, DRL continuously improves its decision policy using accumulated knowledge. Despite the growing interest in applying DRL to IoT-based WSN optimization, there remains a lack of systematic comparison among different DRL techniques under consistent evaluation criteria. This paper addresses this gap by analyzing and benchmarking various DRL algorithms for cluster head selection, providing insights into their strengths, limitations, and suitability for dynamic IoT environments.

2. Related Works

The challenge of uneven energy depletion caused by inefficient cluster head selection has motivated significant research in IoT-based Wireless Sensor Networks. Conventional clustering protocols primarily rely on random or threshold-based mechanisms, which often fail to adapt to dynamic network conditions. To address these limitations, researchers have explored optimization algorithms and intelligent learning techniques for improving CH selection and routing efficiency.

In computational intelligence have emphasized the integration of advanced learning and optimization techniques to improve the balance between exploration and exploitation in reinforcement learning. In IoT-based systems, deep reinforcement learning methods have been widely adopted to enhance decision-making and generalization capability. To address this limitation, Pradeep Kumar Tiwari *et al* [6] proposed a modelBayesian Bootstrap Deep Q-Network (BBDQN)combining bootstrapped DQN with Bayesian inference techniques. By extending Bayesian linear regression concepts to nonlinear neural network models, the approach improves posterior parameter estimation and enhances exploration efficiency. Experimental results demonstrate that BBDQN achieves superior exploration performance compared to conventional DQN and bootstrapped DQN methods, particularly in complex decision-making environments.

To address the growing demand for energy-efficient Wireless Sensor Networks (WSNs) within the expanding Internet of Things ecosystem, S. Regilan *et al* [7] introduces a novel protocol named ReLeC-MEO which integrates reinforcement learning-based clustering with multi-objective optimization to enhance overall network performance. The reinforcement learning component dynamically optimizes cluster formation, while the multi-objective optimization framework identifies Pareto-optimal solutions to balance energy usage, transmission reliability, and network longevity. By ensuring a fair trade-off among these competing metrics, the protocol significantly improves operational efficiency. Simulation results demonstrate that ReLeC-MEO outperforms conventional methods, achieving substantial reductions in latency as 42.9% and drop in energy consumption as 51.6 % while enhancing throughput as 35% compared to baseline protocols.

Chilamkurthy *et al* [8] proposed a Synchronized Weight Control (SWC) framework aimed at optimizing routing in Low-Power Wide-Area Networks (LPWANs). The proposed method leverages reinforcement learning to adaptively modify routing decisions by considering link quality and the residual energy of nodes. Experimental evaluations conducted in LoRa-based LPWAN scenarios show that the SWC approach achieves up to 28% improvement in packet delivery rate and 35% enhancement in energy efficiency compared to standard protocols such as RPL. Despite these advancements, the study does not provide comprehensive analysis for large-scale multi-hop networks and overlooks the impact of interference in densely deployed IoT environments. The authors conclude that SWC offers a strong basis for developing scalable and energy-aware routing solutions in LPWAN systems.

Gupta *et al* [9] introduce an energy-efficient routing optimization method for underwater IoT (UIoT) networks by integrating Q-learning with predictive learning techniques. In this hybrid framework, Q-learning is utilized for adaptive path selection, while predictive learning anticipates environmental changes to proactively adjust routing decisions. The proposed approach was evaluated under varying node densities in a simulated UIoT environment, achieving a 32% improvement in energy efficiency and a 40% reduction in packetloss compared to conventional routing protocols. Despite these promising results, the study does not thoroughly address real-time adaptation under extreme and highly dynamic underwater channel conditions, indicating the need for further validation in real-world deployments. Overall, the findings demonstrate that hybrid learning strategies can significantly enhance UIoT network performance, although practical implementation remains a key area for future research.

Sattibabu [10] presents a Federated Reinforcement Learning (FRL) framework tailored for IoT-enabled Wireless Sensor Networks, enabling decentralized training across sensor nodes without exchanging raw data, thus maintaining privacy. The framework incorporates adaptive model updates, efficient aggregation of heterogeneous data, and an energy-aware federated averaging mechanism suited for resource-constrained environments. Simulation results indicate that the FRL approach improves packet delivery by 13% compared to DQN and 30% over RL-based routing, while also enhancing energy efficiency by 15% and 24%, respectively. These outcomes demonstrate the capability of FRL to outperform conventional methods and improve the overall performance and lifespan of IoT-based WSNs.

The fast expansion of IoT-based Wireless Sensor Networks (IoT-WSNs) has introduced significant security concerns, especially in resource-constrained environments where devices are unable to support computationally intensive or complex security mechanisms. Chaurasia *et al* [11] proposes CREN-RLC, a lightweight and adaptive security framework designed for resource-constrained IoT-based Wireless Sensor Networks. The approach combines energy-aware clustering, based on residual energy and neighbor communication patterns (CREN), with a regression learning classifier (RLC) for real-time intrusion detection. While the clustering mechanism ensures balanced energy consumption, the classifier utilizes historical data to detect evolving attack patterns. Experimental evaluation shows high detection performance, achieving over 94% classification accuracy and strong precision–recall metrics, while maintaining low packet drop rates and high forwarding efficiency even under heavy attack scenarios. The results indicate that CREN-RLC provides an energy-efficient and scalable security solution suitable for practical IoT-WSN deployments.

Chowdhuri and Deb Barma [12] propose a hybrid deep reinforcement learning approach for node localization and coverage hole detection in Wireless Sensor Networks. The framework combines clustering techniques with reinforcement learning to enhance position estimation accuracy. Simulation results show a 34% increase in coverage area and a 28% reduction in localization error compared to conventional clustering methods. Although the model demonstrates improved performance in simulated environments, it has not been validated in large-scale real-world deployments. The authors suggest extending the work to support real-time mobility for broader practical applications.

Kalyana Sundari *et al* [13] addresses efficient data collection in IoT-based Wireless Sensor Networks using Mobile Data Collectors (MDCs) as an alternative to static sink-based approaches. To overcome energy constraints and data redundancy issues, the authors propose a tour scheduling framework based on Deep Reinforcement Learning (DRL). The DRL model determines optimal visiting sequences for MDCs by considering sensor types and data generation rates. To enhance convergence speed, Sunflower Optimization (SFO) is incorporated, while the Capuchin Search Algorithm (CapSA) is applied to identify stable and low-latency collection paths. Simulation results demonstrate reduced data collection delay and packet loss, along with improved packet delivery ratio and residual energy compared to conventional methods.

In IoT applications, these sensors are prone to faults, which heightens the risk of failures. To address issues such as faulty nodes, broken links, and increased communication overhead, Aarthi *et al* [14]proposed an approach integrates multi-objective deep reinforcement learning (MODRL) for intelligent fault identification and K-means clustering for efficient cluster formation. A mobile sink is employed to improve energy-efficient data collection and extend network lifetime. Simulation results indicate that the proposed method achieves 92% fault detection accuracy, while reducing the false alarm rate to 3.5% and false-positive rate to 2.5%. It improves network lifetime by 42% and overall throughput by 37% compared to existing techniques.

Conventional energy optimization techniques are often insufficient to manage the growing complexity of IoT-integrated wireless network systems. Kagi, S. *et al* [15] proposes an N-Federated Deep Reinforcement Learning (FDRL) framework to address energy management challenges in large-scale IoT-enabled Wireless Sensor Networks. The model integrates reinforcement learning with federated learning to optimize resource utilization while preserving data privacy. After federated aggregation, Pelican Optimization is applied for centralized performance enhancement. Simulation results demonstrate significant improvements, including up to 98% energy efficiency, 96% packet delivery ratio, 97% network lifetime extension, and 92% latency reduction compared to existing approaches. The study highlights the potential of combining deep learning and federated techniques for efficient and sustainable IoT-WSN optimization.

3. Comparative Analysis of Existing DRL-Based IoT–WSN Approaches

The Internet of Things-based Wireless Sensor Network represents a transformative integration of sensing, communication, and intelligent data processing technologies. In IoT-based WSNs, numerous sensor nodes equipped with sensing, computation, and wireless communication capabilities are deployed to monitor physical or environmental conditions such as temperature, humidity, pressure, motion, and energy consumption. These nodes collaboratively collect and transmit data to centralized or distributed processing units for analysis and decision-making. By enabling real-time monitoring, automation, and intelligent control, IoT-based WSNs play a crucial role in applications such as smart cities, healthcare monitoring, industrial automation, agriculture, and environmental surveillance.

The comparative performance analysis presented in Table 1. provides a comprehensive evaluation of the existing deep reinforcement learning–based for IoT-enabled Wireless Sensor Networks. The comparison focuses on key performance metrics such as energy efficiency, packet delivery ratio, latency reduction, network lifetime, throughput, fault detection accuracy, and overall scalability. By examining the quantitative improvements reported in recent studies, this analysis highlights the relative strengths and limitations of each approach in addressing energy optimization, reliability, security, and adaptability challenges in resource-constrained IoT-WSN environments. The table facilitates a clear understanding of how advanced DRL and federated techniques outperform traditional routing and optimization methods across multiple performance dimensions.

Table 1. Comparative Analysis of DRL Approaches for Energy-Efficient IoT-WSN

Method	Key Objective	Energy Improvement	PDR / Throughput	Latency Reduction	Other Key Metrics
ReLeC-MEO[7]	RL-based clustering andmulti-objective optimization	51.6% reduction of energy consumption	35% increase in throughput	↓42.9% latency	Balanced Pareto trade-off
SWC (LPWAN)[8]	RL-based adaptive routing	↑ 35% energy efficiency	↑ 28% PDR	Optimized routing stability	Link-quality aware routing
Q-learning and Predictive Learning (UIoT)[9]	Energy-efficient underwater routing	↑ 32% energy efficiency	↓ 40% packet loss	Improved link reliability	Adaptive path selection
Federated RL (FRL)[10]	Privacy-preserving routing	↑ 15–24% energy efficiency	↑ 13–30% PDR	Real-time adaptability	Decentralized training
CREN-RLC[11]	Energy-aware secure clustering	Balanced energy usage	High forwarding efficiency	Low packet drop	94% intrusion detection accuracy
Hybrid DRL (Localization)[12]	Coverage & localization	Extended network lifetime	Efficient load balancing	Lower end-to-end latency	↑ 34% coverage, ↓ 28% localization error
DRL–SFO–CapSA[13]	MDC tour optimization	Improved residual energy	↑ PDR	Reduced delay	Optimized path stability
MODRL and K-means[14]	Fault detection & tolerance	↑ 42% network lifetime	↑ 37% throughput	Robust performance under dense deployment	92% FDA, 3.5% FAR
N-Federated DRL (FDRL)[15]	Large-scale energy optimization	↑ 98% energy efficiency	↑ 96% PDR	↓92% latency	↑ 97% network lifetime

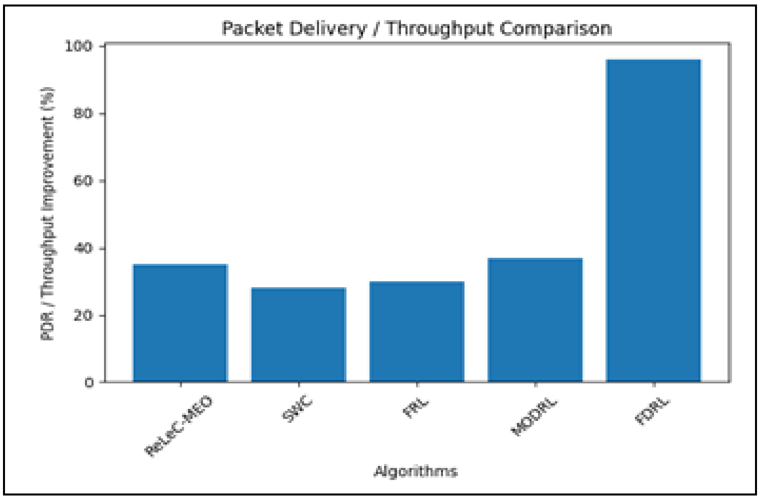

Figure 2: Comparative Analysis of Energy Consumption Reduction in IoT-WSN Algorithms

A comparative evaluation of recent DRL-based IoT–WSN optimization frameworks reveals significant advancements in energy efficiency, latency reduction, throughput enhancement, and network longevity. Among the surveyed approaches, the N-Federated DRL (FDRL) framework [15] demonstrates the highest overall improvement, achieving 98% energy efficiency, 96% packet delivery ratio, 97% network lifetime extension, and 92% latency reduction, indicating superior scalability and optimization capability in large-scale IoT environments. The ReLeC-MEO protocol [7] also shows strong performance by integrating reinforcement learning with multi-objective optimization, reducing energy consumption by 51.6% and latency by 42.9% while improving throughput by 35%. Similarly, the MODRL-based fault detection approach [14] enhances network lifetime by 42% and throughput by 37%, demonstrating the effectiveness of multi-objective learning in reliability-focused applications. Fig 2 illustrates the percentage reduction in energy consumption achieved by different learning-based routing protocols. The comparison clearly shows that advanced DRL and federated learning approaches significantly outperform traditional methods in minimizing energy usage. Protocols integrating multi-objective optimization and federated learning demonstrate the highest energy savings. Improved energy efficiency directly contributes to prolonged network lifetime in IoT-WSNs.

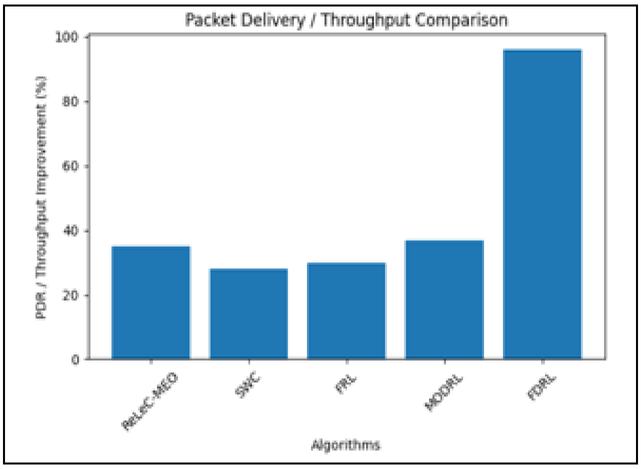

Figure 3: Performance Comparison of Packet Delivery Ratio in IoT-WSN Algorithms

In routing optimization, SWC [8] and Q-learning-based underwater routing [9] provide moderate yet consistent improvements in packet delivery (28%) and energy efficiency (32–35%), though their scalability in large multi-hop or dynamic environments remains limited. The Federated Reinforcement Learning framework [10] emphasizes privacy-preserving decentralized optimization, improving packet delivery by up to 30% and energy efficiency by 24%, making it suitable for distributed IoT scenarios. Fig 3 compares the enhancement in packet delivery ratio and throughput across various intelligent routing frameworks. Learning-based approaches achieve higher data reliability by optimizing routing decisions dynamically. The results indicate that federated and multi-objective DRL models provide better data transmission stability compared to conventional DQN-based methods. Higher PDR ensures reliable communication in dense IoT environments.

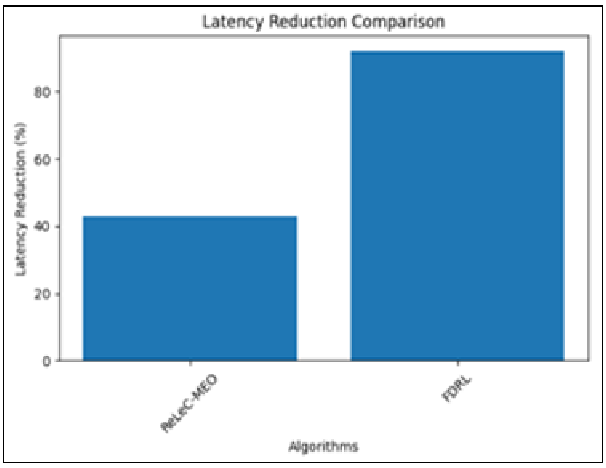

Figure 4: Performance Comparison of Latency in IoT-WSN Algorithms

Security-focused work such as CREN-RLC [11] achieves over 94% intrusion detection accuracy while maintaining low packet drop rates, highlighting the integration of learning-based security with energy-aware clustering. Meanwhile, the hybrid DRL localization model [12] improves coverage by 34% and reduces localization error by 28%, addressing spatial efficiency rather than routing performance. The DRL–SFO–CapSA framework [13] enhances mobile data collection efficiency by reducing delay and packet loss while improving residual energy levels.Fig 4 presents the percentage decrease in end-to-end delay achieved by different optimization algorithms. DRL-based scheduling and clustering mechanisms effectively reduce communication latency by selecting optimal transmission paths. Multi-objective models particularly show significant delay minimization. Lower latency is critical for real-time IoT applications such as healthcare monitoring and industrial automation.

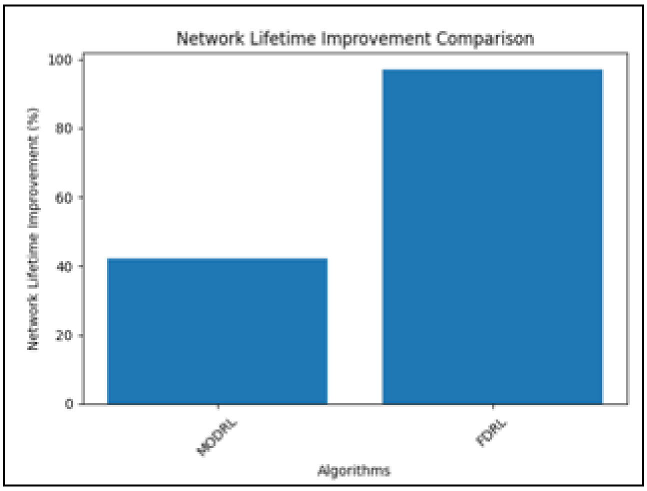

Figure 5: Performance Comparison of Network Lifetime in IoT-WSN Algorithms

Fig 5 highlights improvements in overall network lifetime due to intelligent energy-aware routing strategies. By balancing residual energy and avoiding early node depletion, advanced DRL and federated frameworks extend operational duration. Multi-objective and federated models demonstrate superior performance in sustaining long-term network functionality. Enhanced lifetime ensures stable and cost-effective IoT-WSN deployments.

4. Conclusion

With the increasing deployment of IoT-based WSNs across diverse real-world applications, efficient resource utilization and reliable communication have become critical challenges. As IoT-based WSNs continue to expand in scale and application scope, the need for scalable, energy-efficient, and intelligent routing mechanisms becomes increasingly important. This study presented a comprehensive comparative analysis of advanced deep reinforcement learning for IoT-enabled Wireless Sensor Networks. The results clearly demonstrate that Deep Reinforcement Learning and Federated Deep Reinforcement Learning approaches significantly outperform conventional routing and energy optimization techniques. In particular, intelligent learning-based protocols effectively reduce energy consumption, improve packet delivery ratio, minimize latency, and extend overall network lifetime. Overall, while most approaches achieve improvements in specific metrics, federated DRL models demonstrate superior holistic optimization. However, several studies lack large-scale real-world validation, dynamic interference handling, and extreme environment adaptability, indicating future research opportunities in scalable, secure, and real-time IoT–WSN deployments.

References

[1] Ladosz, Pawel, Lilian Weng, Minwoo Kim, and HyondongOh.. "Exploration in deep reinforcement learning: A survey." *Information Fusion* 85, 1-22, 2022.

[2] Verma, Vanita, and Vijay Kumar Jha. "Secure and energy-aware data transmission for IoT-WSNs with the help of cluster-based secure optimal routing." *Wireless Personal Communications* 134.3 (2024): 1665-1686.

[3] Braiber, Rafa Sami, and Mussab Riyadh. "IoT Applications Using Clustering Protocols in Wireless Sensor Networks." *Journal of Al-Qadisiyah for Computer Science and Mathematics* 17.1, Pg No: 240-251, 2025.

[4] Rajput, Meenal, and Ram Narayan Yadav. "Machine and deep learning driven energy efficient clustering in IoT-WSNs: A Review." *IEEE Sensors Journal*, 2025.

[5] Suresh, S. Sebastin, et al. "Intelligent data routing strategy based on federated deep reinforcement learning for IOT-enabled wireless sensor networks." Measurement: Sensors 31, Volume 6,2024. https://doi.org/10.1016/j.prime.2023.100298.

[6] Tiwari, PradeepKumar, Singh, Pooja, Rajagopal, Nava neethaKrishnan, Deepa, K., Gulavani, Sampada, Verm a, Amit, Kumar, Yekula Prasanna, "IoT-Based Reinforcement Learning Using Probabilistic Model for Determining Extensive Exploration through Computational Intelligence for Next-Generation Techniques", *Computational Intelligence and Neuroscience*, 2023, 5113417, 13 pages, 2023. https://doi.org/10.1155/2023/5113417.

[7] S. Regilan, L. K. Hema, J. Jenitha, "ReLeC-MEO: Reinforcement Learning-Based Clustering with Multi-Objective Efficient Optimization for Energy-Efficient IoT Networks", International journal of communication system, Volume38, Issue15, October 2025, https://doi.org/10.1002/dac.70239.

[8] N.S. Chilamkurthy, S.A. Hakeem, S.K. Tiwari, et al. Optimal Routing Protocol in LPWAN Using SWC: A Novel Reinforcement Learning Framework. *IEEE Sensors Journal* 2024, 24 (9), 15607–15619

[9] D. Gupta, A. Bajpai, N.K. Tiwari, S. Yadav. Energy-Efficient Routing Optimization for Underwater Internet of Things using Hybrid Q-Learning and Predictive Learning Approach. *Procedia Computer Science*, 235, 45–55, 2024.

[10] Sattibabu, G., Ganesan, N. & Kumaran, R.S. IoT-enabled wireless sensor networks optimization based on federated reinforcement learning for enhanced performance. *Peer-to-Peer Netw. Appl.* 18, 75 (2025). https://doi.org/10.1007/s12083-024-01887-5.

[11] N. Chaurasia and P. Kumar, "CREN-RLC: Clustering-Based Adaptive Security with Regression Learning for IoT-WSNs," in *IEEE Sensors Journal*, vol. 25, no. 24, pp. 44984-44993, 15 Dec.15, 2025, doi: 10.1109/JSEN.2025.3620211.

[12] R. Chowdhuri, M.K.D. Barma. Node position estimation based on optimal clustering and detection of coverage hole in wireless sensor networks using

hybrid deep reinforcement learning; Springer US, Vol. 79, 2023.

[13] P. Kalyana Sundari, R. Vadivel, "Tour Path Scheduling Using Optimized Deep Reinforcement Learning for IoT Mobile Data Collectors", Volume39, Issue4, 2026. https://doi.org/10.1002/dac.70402.

[14] Aarthi, A., T. Abinaya, and D. Sridharan. "Intelligent fault tolerance data gathering in IoT enabled WSN." *2025 1st* International Conference on Radio Frequency Communication and Networks (RFCoN). IEEE, 2025.

[15] Kagi, S., M, A., Kousalya, C.G. *et al.* "Federated deep reinforcement learning (FDRL) framework using pelican optimization (PO) to achieve sustainable energy in IoT-integrated wireless networks". International. Journal of information. technology. 2025. https://doi.org/10.1007/s41870-025-02752-1.

[16] Le, H., Wang, Y., Gotmare, A. D., Savarese, S., & Hoi, S. C. H., "Coderl: Mastering code generation through pretrained models and deep reinforcement learning. Advances in Neural Information Processing Systems", *35*, 21314-21328, 2022.