SCIFINITI
PUBLISHING

Research Article

Optimizing Cluster Head Selection and Routing in 5G WSNs: A Reinforcement Learning and Deep Learning Approach

Vijayakumar Kadumbadi[✉,1,*][ID] Thirumaraiselvan Packirisamy[✉,2][ID] Sivakumar B[✉,3][ID]
Seenuvasan P[✉,4][ID]

[1] Department of Artificial Intelligence & Data Science, Panimalar Engineering College, Chennai 600123, India
[2] Department of Electronics and Communication Engineering, Adhiparasakthi Engineering College, Melmaruvathur 603319, India
[3] Department of Computing Technologies, SRM Institute of Science and Technology, Chengalpattu 603203, Tamil Nadu, India
[4] Department of Information Technology, University College of Engineering, Villupuram 605103, Tamil Nadu, India

Abstract

The Internet of Things (IoT) and 5G wireless sensor networks (WSNs) have transformed data transmission and inter-device communication; however, they face persistent routing challenges owing to energy constraints, latency, and packet loss. This study proposes an energy-efficient data transfer framework for IoT-based 5G WSNs by integrating a deep belief network (DBN) topology with a reinforcement learning (RL)-based clustering mechanism and Manta Ray Foraging Optimization (MRFO) for multi-objective cluster head (CH) selection (energy, delay, traffic density, and distance). Unlike existing approaches, such as deep neural networks (DNNs) and time-temperature-dependent forwarding protocols (TTDFP), which focus narrowly on latency or energy efficiency, our hybrid DBN-RL-MRFO architecture jointly optimizes routing stability, scalability, and energy consumption. Simulations demonstrate that the proposed DBN-RL-MRFO framework reduces energy consumption by 5–10% compared to DNN-based methods and improves network lifetime (FND) by 5–15% over the TTDFP, while maintaining near-optimal throughput and latency. Although GEEC achieves lower energy use, our method balances energy efficiency with superior throughput (+3–8%) and reliability (PDR > 99.5). Further, statistical and complexity analyses validate its robustness. This study advances reliable routing for IoT applications (smart cities, healthcare, and industrial automation) by balancing the trade-offs between critical WSN constraints.

1. Introduction

2G (2nd generation) networks offer wireless connections that satisfy user demands for data and voice communications. With 3G technology, smartphones were able to stream movies and other content despite limited bandwidth. The introduction of 4G, however, provided a substantial increase in capacity. Furthermore, a considerable proportion of smartphone users worldwide now depend on these devices for routine tasks. Reports indicate that smartphone adoption is more widespread globally than the population of the United States [1]. Smartphones, sending data and video via 3G and 4G, are the primary source of large traffic volumes. Moreover, inappropriate smartphone usage contributes to network congestion and quality-of-service (QoS) issues [2,3]. Given the current state of technology, the implementation of 5G is essential, as it will enable the development of device-to-device (D2D) linkages. With cellular providers launching 5G services globally in 2019, 5G technology is rapidly becoming the standard for cellular networks [4]. However, most modern smartphones use 4G networks for both data

and video transmission. Similar to older networks, 5G networks divide their service regions into smaller geographical areas known as cells [5]. A local antenna connects each cell's 5G wireless device to the phone network and internet via radio waves. The advantage of this 5G network is its increased capacity and ability to attain download rates of up to 10 gigabits per second (Gbit/s) [6].

The Internet of Things (IoT) is a rapidly emerging 5G-enabled networking paradigm with numerous potential applications across various domains. IoT systems typically employ Wireless Sensor Networks (WSNs), which consist of densely deployed sensing devices used to remotely monitor the surrounding environment. In the last several decades, wireless sensor networks (WSNs) have become increasingly important in the field of communication because of their unique characteristics, such as mobility and simplicity of connection. These features establish them as well-known networked data conduits [7].

Real-time wireless sensor networks (WSNs) face significant challenges in achieving energy-efficient data transmission while simultaneously meeting stringent scheduling and reliability requirements. The main issues include minimizing latency in proportion to energy use, ensuring reliability without unnecessary retransmissions, developing time-synchronized communication protocols, and handling heavy traffic while maintaining energy efficiency. Additional challenges include mobility, security, and scalability. As shown in Figure 1, academic researchers are investigating AI-based optimization techniques and adaptive, cross-layer, and harvesting-aware protocols to solve these problems and meet the needs of modern real-time WSN systems.

The communication revolution involves the concurrent operation of multiple wireless networks on the communication spectrum. Due to its higher bandwidth, mobile communication is regarded as one of the most widely adopted wireless networking technologies in the telecommunications sector. The 5G communication protocols further enhance wireless communication by facilitating efficient data packet transfer [8].

A protocol, an assembly of rules, uses a specific routing technique to route data packets from the source to the destination. Networking protocols already have routing rules in place. Wireless communication uses multiple layers to implement protocols and transfer data across different levels [9]. The transport layer in mobile wireless communication facilitates data transfer by implementing a specific protocol for data delivery. To ensure efficient allocation of network resources, the transport layer protocol employs a congestion management mechanism. Furthermore, congestion management is considered the most crucial problem at the transport layer in wireless networks.

When congestion control is implemented in mobile wireless communication, including 5G communication, the performance of the entire network collapses. We have developed routing techniques [10,11] to avoid this issue.

A sensor network requires a routing protocol to identify the path from the sender node to the destination sink node. Routing techniques in wireless sensor networks are designed to identify data transmission paths that minimize latency and optimize energy efficiency [12]. Combining the advantages of energy efficiency and route-finding techniques yields reliable paths. It automatically adapts to network density and traffic patterns in data-intensive sensor networks [13]. Routing protocols manage neighbor discovery, route selection, and power control to improve network scalability and flexibility. Certain routing techniques prioritize network resilience, hop count, and latency optimization over energy economy. We chose these steps to enhance the design elements of the routing protocol [14].

Machine learning techniques are a notable way to enhance network performance and solve difficult decisioning-making problems [15]. The Internet of Things (IoT), with its support for Wireless Sensor Networks (WSN) and strong machine learning algorithms, handles and evaluates complex routing and energy management decision-making challenges. Learning-based algorithms solve the problem of creating the best routing paths with high precision [16]. Machine learning methods must be used to assess constraints, enabling the routing process to automatically understand the dynamic aspects of networks, including congestion areas, connection quality, topology changes, and new flow arrivals. The goal of this analysis was to improve service quality. Each sensor node (SN) makes decisions based on its observation state and decision-making abilities, which may lead to intelligent actions. Furthermore, the system repeatedly learns and makes decisions until it identifies an optimal response [17]. Recent advancements in 5G technology, such as Ultra-Reliable Low-Latency Communication (URLLC) and network slicing, have greatly enhanced the reliability and resource allocation of Internet of Things (IoT) networks [18, 19]. Nevertheless, the emphasis of URLLC on achieving ultra-low latency often overlooks the aspect of energy efficiency in Wireless Sensor Networks (WSNs), while the dynamic resource partitioning inherent in network slicing may result in overheads for large-scale sensor deployments [20]. This study addresses these challenges by introducing a hybrid Deep Belief Network-Reinforcement Learning-Manta Ray Foraging Optimization (DBN-RL-MRFO) framework that concurrently optimizes latency, energy consumption, and scalability, which are critical re-
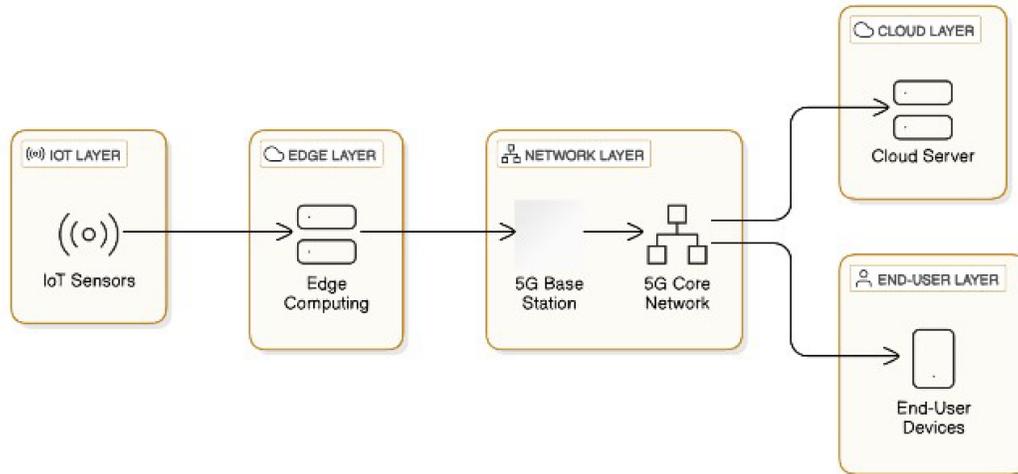
Figure 1: 5G and beyond wireless sensor network communication.

quirements for 5G-enabled WSNs in the context of smart cities and Industry 4.0.

The proposed DBN-RL-MRFO framework is consistent with the 3GPP Release 17 IoT standards [18], which emphasize energy-efficient ultra-reliable low-latency communication (URLLC), particularly for industrial IoT applications, and scalability for massive machine-type communication (mMTC). Although Release 17 specifies reduced-capability (RedCap) devices for low-power wireless sensor networks (WSNs) [3], it does not prescribe specific methods for resource optimization, leaving this aspect open for implementation. Our research addresses this gap by incorporating the following elements.

- Reinforcement learning (RL) for adaptive clustering, complying with 3GPP's push for AI/ML in RAN intelligence (Release 18) [17].
- Multi-objective CH selection (energy, latency, density), mirroring 3GPP's QoS prioritization for heterogeneous IoT traffic [18].

The main goal of the recommended routing design is as follows:

- A clustering approach was used to implement social network grouping.
- A unique optimization method for CH selection is presented.
- We propose an effective routing method based on machine learning (ML).
- We calculated and compared the algorithm's performance with recently released techniques.

1.1. Main Contributions

The main contributions of this study are summarized as follows:

- *Novel Hybrid Architecture:* We propose a novel DBN-RL-MRFO framework that synergistically combines deep learning, reinforcement learning, and bio-inspired optimization for holistic WSN optimization, moving beyond approaches that focus on a single objective.
- *RL-based Clustering Mechanism:* We designed an RL-based clustering algorithm that dynamically groups sensor nodes to minimize energy consumption and improve network stability, adapting to network changes more effectively than static clustering protocols.
- *Multi-Objective CH Selection Model:* We formulate the CH selection as a multi-objective optimization problem (considering energy, delay, traffic density, and distance) and employ the Manta Ray Foraging Optimization (MRFO) algorithm to solve it efficiently.
- *DBN-based Routing Protocol:* We developed a Deep Belief Network-based routing protocol that intelligently learns optimal data paths, enhancing throughput and reliability while conserving energy.
- *Comprehensive Performance Validation:* We provide extensive simulations demonstrating that our proposed framework outperforms state-of-the-art protocols, such as DNN, TTDFP, and GEEC, in terms of network lifetime, energy consumption, throughput, and latency, and validate its statistical significance and complexity.

The remaining sections are structured as follows: Section 2 reviews related literature on routing and clustering in WSNs. Section 3 identifies the research problem and motivation. Section 4 details the proposed DBN-RL-MRFO framework, including system model, energy consumption model, and algorithmic components. Section 5 presents and discusses the simulation results. Section 6 provides a discussion of key findings, limitations, and comparisons with existing work. Finally, Section 7 concludes the paper and outlines future research directions

2. Literature Survey

Researchers have conducted numerous empirical studies to enhance the performance of 5G wireless communications. Wireless routing systems have gained considerable attention due to significant technological advancements in 5G. This article offers a succinct overview of recent developments in routing protocols.

Thangaramaya [21] developed a routing theory for Wireless Sensor Networks (WSNs) in the Internet of Things (IoT). Wireless Sensor Networks (WSNs) enable data sensing, collection, and transfer among devices in the Internet of Things (IoT). The Internet of Things has enabled Wireless Sensor Networks (WSN) to use intelligent routing to enhance network performance. The principles of energy-efficient routing have been extensively investigated in recent studies. This study addresses the development of a neuro-fuzzy rule for cluster formation in IoT-based wireless sensor networks (WSNs) to enhance the current approach. However, this approach needs to be enhanced for group Wireless Sensor Networks (WSN) within an Internet of Things (IoT) architecture. Our analysis indicates that this routing algorithm yields strong performance across multiple factors, including packet delivery ratio (PDR), energy consumption, latency, and network longevity.

Sujanthi and Kalyani (2017) [22] introduced a QoS-aware, safe deep learning method for dynamic cluster-based routing in Wireless Sensor Networks (WSN) supported by the Internet of Things (IoT). The open and resource-constrained nature of WSN-assisted IoT presents security and energy efficiency as challenging issues that must be addressed. This study constructs a hybrid WSN-IoT network based on dynamic clusters using the Secure Deep Learning (SecDL) technique. Furthermore, we specifically designed a network using mobile sink technology and bicentric hexagons to improve energy efficiency. We employed a bidirectional data elimination and reduction framework to manage data consolidation within each cluster. One-time encryption (OT-encryption) ensures a high degree of security for combined data. The encrypted

data was routed through a mobile sink via a selected path, ensuring robust Quality of Service (QoS). We developed a crossover-based fitted deep neural network (Co-FitDNN) to achieve optimal route identification. As we employed IoT users to collect sensory data, user security was the primary focus of this study.

Huang [23] reported a deep learning model for estimating connection reliability for routing in Wireless Sensor Networks (WSN). This paper presents a robust routing technique to enhance the routing of wireless sensor networks (WSN). The present study introduces a deep learning model, the Weisfeiler–Lehman Kernel and Dual Convolutional Neural Network (WL-DCNN) method, which demonstrates strong performance in extracting and labeling subgraphs. Its goal was to enhance the generalization and flexibility of self-learning. We created a reliable routing model, WL-DCNN, specifically for Wireless Sensor Networks (WSN). Resilient routing in Wireless Sensor Networks (WSNs) evaluates the reliability of target connections by analyzing topological data during attacks on routing tables, which cause varying levels of disruption to the local link community.

Ibrahim El-Moghith and Darwish [24] developed a deep, trustworthy routing system based on a blockchain, specifically for wireless sensor networks (WSNs). Routing attacks easily breach the core functionalities of Wireless Sensor Networks (WSN), causing significant harm to the network as a whole. A reliable routing technique is essential for Wireless Sensor Networks (WSNs) to ensure efficient operation and enhance routing security. The implementation of trust restrictions, centralized decision-making, or cryptographic approaches increases the dependability of routing systems. This study presents a unique method for enhancing routing security and efficiency in deep-chain networks: the implementation of Markov Decision Processes (MDPs). Within the blockchain network, the proposed design uses a proof-of-authority technique to confirm the legitimacy of the information distribution process. We developed a unique deep learning method to integrate the distinct features of several nodes. We selected the best neighboring hop as a forwarding node using multiple decision processes (MDPs) to ensure safe and effective message delivery.

Razhavendra and Mahadevaswamy [20] presented a composite fuzzy technique [22] for energy-efficient routing in Wireless Sensor Networks (WSN). Optimizing the battery performance of wireless sensor networks (WSNs) necessitates careful monitoring of energy consumption. The battery of a WSN depletes due to transmission and sampling rates. We developed an energy consumption modeling approach to analyze key factors affecting the lifetime of a Wireless Sensor Network (WSN). The cur-

rent research investigates the role of fuzzy membership functions in extending the network lifespan. We adjusted the parameters at several levels using advanced fuzzy logic methods. This paper describes an effective integration of routing and clustering activities using the hybrid metaheuristic cluster-based routing (HMBCR) technique. We introduced a novel approach, Levy distribution-based brainstorm optimization (BSO-LD), to enhance the clustering efficiency. We then present a water wave optimization technique based on hill climbing (WWO-HC) to select an optimal route.

In [25], we implemented algorithmic CH selection using a pragmatic methodology that included several critical criteria for CH selection. Routing traffic via the selected cluster head (CH) will enhance performance. After considerable consideration, a hybrid optimization technique called genetic-based particle swarm optimization (GA-PSO) was developed for CH selection and routing. We determined the optimal path for sink mobility using Particle Swarm Optimization (PSO). Babu et al. [26] describes a new method for effective clustering. This approach is referred to as Integrated Distributed Autonomous Fashion with Fuzzy If-Then Rules (IDAF-FIT). The if-then rule guided the selection of the CH during the clustering process. This approach employs an adaptive source location privacy preservation technique, known as Randomized Routes (ASLPP-RR), to select the optimal path. Ultimately, we implemented a security analysis procedure to enhance the privacy of sensitive information. In addition to cluster-based routing, the rate control idea was used in [27], which extended the system durability across longer simulation sessions. The first step involved grouping the nodes using a hybrid K-means and greedy best-first search approach to achieve lifetime improvement. We also aimed to control the rates by introducing the firefly (FF) optimization technique. Ultimately, we employed the Ant Colony Optimization (ACO) technique to determine the optimal data transmission channel. First presented in [28], the routing strategy is based on African buffalo optimization (ABO). We studied the behavior of African buffalo and applied it to develop an optimal route selection technique. The African Buffalo Optimization (ABO) algorithm serves as the main controller, managing communication between nodes and building systems. It has a long network life and effectively sends packets from the source to the sink node.

The multi-criteria decision-making (MCDM) [29] technique is the most successful method for making decisions. Fuzzy logic was integrated into the MCDM approach to enhance its performance and address its shortcomings. The study created a hybrid routing model and a fuzzy-based multiple criteria decision-making (MCDM)

system for choosing Cluster Heads (CH). We then applied the Generalized Intuitionistic Fuzzy Soft Set (GIFSS) technique in combination with a hybrid Shark Smell Optimization (SSO) algorithm to achieve optimal cluster head (CH) selection. A genetic algorithm (GA) was used to achieve efficient routing. Ultimately, we evaluated a limited set of performance metrics to demonstrate the effectiveness of the GIFSS-SSO approach.

Wireless sensor networks (WSN) use a certain number of nodes to gather data from the surrounding area. However, throughout this process, energy saving was the main objective. Routing and clustering algorithms are mostly responsible. In this study, we present an energy-aware distance-based CH selection and routing (EADCR) protocol to enhance the lifetime and energy efficiency of nodes in wireless sensor networks (WSN). We used a modified form of the fitness function throughout the CH selection process to minimize the energy consumption [30]. This study presents a new approach for finding the shortest path in routing operations. This method uses the Euclidean distance to reduce energy consumption. The network lifetime and overall energy efficiency both increased with the implementation of this integrated approach.

Wireless sensor networks (WSNs) can perform complicated communication using a high number of sensor nodes (SNs). However, there are currently fewer satellite networks (SNs), resulting in a decline in communication and sensing capabilities. It always reduces the routing quality of service (QoS) performance. To address this issue and improve routing efficacy and efficiency, Ramkumar et al. [31] introduced a fuzzy-based relay node selection and energy-efficient routing (FRNSEER) technique. We used the fuzzy rule technique to select the sink node. The active selection of a relay node may increase the data transmission utility factor and energy efficiency. We deployed a highly efficient sensor hub between the relay nodes and the sink to improve communication performance.

Sert et al. [32] introduced a two-tier distributed fuzzy logic-based prototype (TTDFP) to enhance the efficiency of multihop wireless sensor networks (WSNs). Clustering meets the need to optimize aggregation with respect to energy usage. Cluster heads (CHs) receive the gathered data in a clustered network and then forward the received packets to the base station. Hotspots and/or energy hole issues may arise when using a multihop topology. The TTDFP approach, being adaptive and distributed, demonstrates good scalability and effective performance in Wireless Sensor Networks (WSNs). Moreover, it uses optimization methods to modify fuzzy parameters. This technol-

ogy achieves high levels of energy efficiency and network lifespan.

Researchers have identified clustering as the most effective communication platform for wireless sensor networks (WSN). Fuzzy methods have recently gained popularity as efficient clustering strategies because of their high degree of precision. However, it may take some time to determine the best choice. Sert et al. [32] introduced a clonal selection technique based on rule-based fuzzy clustering to overcome the drawbacks of fuzzy algorithms. Compared with other fuzzy-based techniques, CLONALG-M exhibited better performance. This technique is founded on the principle of clonal selection, which is rooted in the adaptive immune system. The immune system concept was applied to predict output deployment using the membership function, thereby enhancing overall performance. Extensive research has demonstrated that this algorithm outperforms alternative approaches.

This implementation [33] discusses the design challenges of URLLC use cases, providing an overview of the technological components introduced in 3GPP Release 15 and the potential advancements in Release 16. In addition, coordinated multi-cell resource allocation methods are examined. System-level simulation results in an urban macro environment indicate that effective multi-cell cooperation, particularly through soft combining, can substantially increase URLLC capacity. Hussein and Ibnkahla [19] delineates Intelligent Intent-Based Network Slicing (I-IBNS) systems as exemplars of the integration of intelligent Intent-Based Networking (IBN) and Network Slicing (NS) for the Management and Optimization (MO) of Internet of Things (IoT) systems. This study further surveys I-IBNS systems, concentrating on two pivotal domains: resource management and data management. The resource management section scrutinizes recent advancements in IBN mechanisms within the NS framework. The data management section investigates the complexities inherent in IoT networks. Additionally, this study envisions the roles of intent, NS, and the IoT ecosystem, thereby establishing a foundation for prospective research directions.

Although [34] employs URLLC to ensure latency guarantees, it overlooks the multi-objective trade-offs between energy and delay. Similarly, Abba Ari et al. [35] utilizes network slicing but requires centralized control, thereby constraining scalability. Our DBN-RL-MRFO approach decentralizes decision-making, which is consistent with 5G's emphasis on edge intelligence.

Clustering improves the scalability, communication capability, and energy efficiency of networks. There are two types of clustering: equal and unequal, as well as static and dynamic. In wireless sensor networks, hotspots require a large overhead and are prone to connectivity issues. The only way to overcome these obstacles is through uneven clustering. Garg and Kaur [36] introduced a fuzzy logic method based on zonal division to address the hotspot issues. We carried out clustering to decrease the rate of energy consumption using fuzzy logic. It performs better by reducing energy consumption, increasing network longevity, and balancing the loads. Table 1 lists all existing clustering and routing algorithms that employ optimization-based approaches or other methods. However, no study has addressed artificial intelligence or optimization techniques. We employed Deep Belief Network (DBN) and Reinforcement Learning (RL) approaches to form and route clusters. These strategies have improved the system's capacity to sustain prolonged periods of operation by extending the network's overall lifespan.

3. Identification and Motivation of the Problem

Routing is a critical task that requires close supervision in the Internet of Things (IoT), enabled by Wireless Sensor Networks (WSNs). Routing refers to the process of creating a data transmission connection between base stations (BSs) and subnetworks (SNs). The data routing strategy distinguishes Wireless Sensor Networks (WSNs) from other wireless ad hoc networks and existing communication methods, while also addressing other problematic issues such as energy consumption and short network lifetime. The WSN routing process considers three main aspects. First, establishing a global addressing process to support the deployment of more SNs is not possible. Therefore, sensor networks can function without traditional IP-based protocols. Furthermore, unlike traditional communication networks, sensor network applications require a continuous flow of sensed data from multiple sources to a designated sink node or base station. Third, the use of multiple sensors near a phenomenon generates a significant amount of duplicate traffic throughout the entire network, creating the same data. Moreover, this type of duplication increases the need for transmission bandwidth and energy usage. Moreover, this leads to several issues, including packet loss, delay, and bandwidth degradation. This motivated us to devise a more straightforward routing method that utilizes machine learning techniques. In the future, this procedure should be used to make wise decisions based on lessons learned from prior experiences.

4. Methods

This section provides a comprehensive description of the methods employed in this study. We begin by presenting

Table 1: Quantitative evaluation of many cutting-edge algorithms for Wireless Sensor Networks (WSN).

Index	Benefits	Drawbacks
[21]	Data transfer reliability. The structure of social networks allows optimization.	Specific to social network optimization Restricted evaluation of the performance.
[22]	Uplink data transfer optimization. Advanced computers are used to enhance efficiency.	The downlink requirements were disregarded. The assessment of scalability is limited.
[23]	Optimized energy management model An effective data distribution architecture	Absence of a thorough execution. The experimental analysis was limited.
[24]	We are optimizing the trajectory for data gathering with the assistance of UAVs. It reduces the inefficiency of data acquisition.	Restricted assessment. Ignores further factors
[31]	Hole healing process and coverage optimization We incorporated the concepts of wakefulness and sleep.	Restricted assessment. This may not be relevant in the case of dynamic node availability.
[25]	The data collection plan increased productivity. Modernizing energy optimization	Absence of outcomes. Restricted assessment.
[27]	Improving energy efficiency; Cluster-based hybrid optimization is an energy-efficient technique.	Real-world assessment is limited The additional optimization factors were ignored.
[28]	This includes both stationary and movable sink nodes. The methodology for sleep scheduling and clustering relies on particle swarm optimization.	Strict assessment It ignores node mobility and network dynamics.
[36]	Increased network longevity and energy efficiency The clustering routing protocol is based on thermal-exchange optimization.	Limited real-world assessment Ignores other optimization goals
[32]	Modernizing energy optimization Effective data routing	Strict assessment in a dynamic network Ignores additional optimization goals
[33]	3GPP Rel-15 and potential advancements from Rel-16.	Restricted assessment. This may not be relevant in the case of dynamic node availability.
[19]	Intelligent Intent-Based Networking (IBN) and Network Slicing (NS) for the Management	Limited real-world assessment Ignores other optimization goals

the system model and fundamental assumptions, followed by the energy consumption model. We then detail the proposed hybrid DBN-RL-MRFO framework and explain its three core components. Finally, we specify the simulation setup, performance metrics, and statistical methods used for the evaluation.

## 4.1. System Model and Assumptions

We created the WSNs system model based on the following assumptions:

- By nature, secondary and source networks are static.
- Cluster Head (CH) collects data using a single sink.
- We divided the SNs into three categories: advanced, intermediate, and normal nodes, owing to their diversity.

- The CH collects input from the sensory nodes and transmits the information to the sink node.
- The sink should function as a central node that regularly updates itself with information about all subnetworks.
- This approach uses an inter-data communication mechanism to accomplish data transfer via CH. We refer to the node with 0% battery life as the "dead node."

Figure 2 illustrates the concept of cluster-based single-hop communication for the Internet of Things (IoT) with assistance from Wireless Sensor Networks (WSN). This study uses the concept of machine learning (deep neural network) to construct an effective route in a 5G wireless communication network (WSN-assisted IoT). It is crucial to arrange the sensors into node clusters before starting the routing process. Clustering techniques are crucial for

achieving energy-efficient transmission, thereby extending network lifetime and reducing overall energy consumption. This study presents a reinforcement learning (RL) approach to clustering. The base station (BS) or sink node, which performs the clustering function, centrally assigns each sensor network (SN) to a specific cluster based on its location.
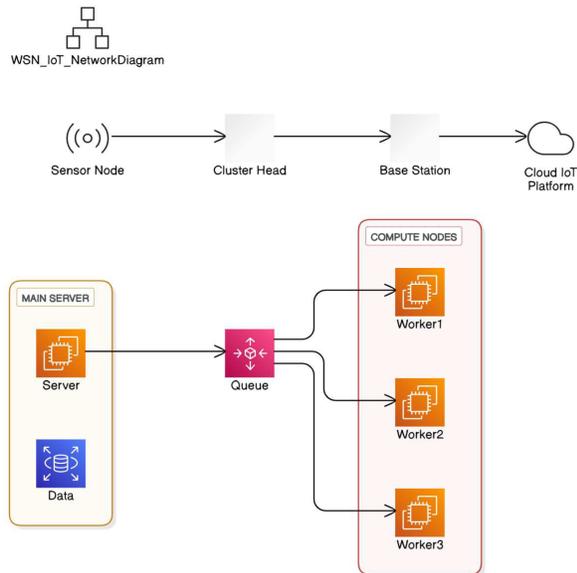
Figure 2: Cluster-based single-hop communication in Wireless Sensor Networks (WSN) supports the Internet of Things (IoT).

An optimization approach is applied to select cluster heads (CHs) after the sensor nodes (SNs) have been assigned to clusters. However, in a hierarchical clustering-based Wireless Sensor Network (WSN), the energy use is high because of all the processing that is needed to combine data and obtain data from each CH member sensor node. Therefore, the CH must be carefully selected to extend the network lifetime. To choose the CH from a cluster, this study presents the Manta Ray Foraging Optimization (MRFO) technique. Recently, researchers have developed this approach as a bio-inspired optimization strategy to tackle real-world engineering problems. Figure 3 depicts the process flow of the recommended method.

Each cluster must optimize its cluster head (CH) while considering various constraints, such as latency, energy, traffic density, and distance. Figure 4 shows the challenges and risk factors to be considered in wireless sensor networks. In sensor networks, finding the optimal route is essential for enhancing the performance of Wireless Sensor Networks (WSNs) in dynamically unstable, asymmetric, and fluctuating wireless channels. This covers the latency, throughput, energy efficiency, and data integrity. After selecting the CH, we recommend implementing a routing system that utilizes a Deep Belief Neural Network (DBN) for efficient data transit.

The neural network performs this routing using several variables, such as residual energy, distance from the cluster head (CH), number of neighboring nodes, and connection distance. Consequently, the proposed routing algorithm actively learns the communication patterns of the nodes to achieve energy-efficient routing.

4.2. Energy Consumption Model

This study adapted the radio energy dissipation model from a previous study [37]. In this setup, the receiver powers the radio electronics, whereas the transmitter powers the amplifiers and subsequent radio electronics. This experimental method uses a multipath fading model. The measured distance d between the transmitter and receiver exceeds the specified threshold. d^2 is the symbol representing the loss of energy in free space. In addition, the energy dissipation associated with multipath fading is represented by d^4. The equation for the energy consumption model P_s during the transmission of the kth bit packet is (1).

$$P_s = \begin{cases} k * \left(P_{ec} + P_{frs}*dis^2 \right); dis < d_0 \\ \left(P_{ec} + P_{mpf}*dis^4 \right) dis \geq d_0 \end{cases} \quad (1)$$

Next, the distance between the sender and receiver and the allowed bit error rate (BER) are used to evaluate the multipath or free space fading model, which is represented as $P_{frs}*dis^2$ or $P_{mpf}*dis^4$.

The variable d represents the distance between the sender and the recipient. The amount of particular energy needed to get the bit over the multi-path fading channel and into free space is called P_{frs} and P_{mpf}, respectively. The threshold distance determined by Equation (2) is shown as d_0.
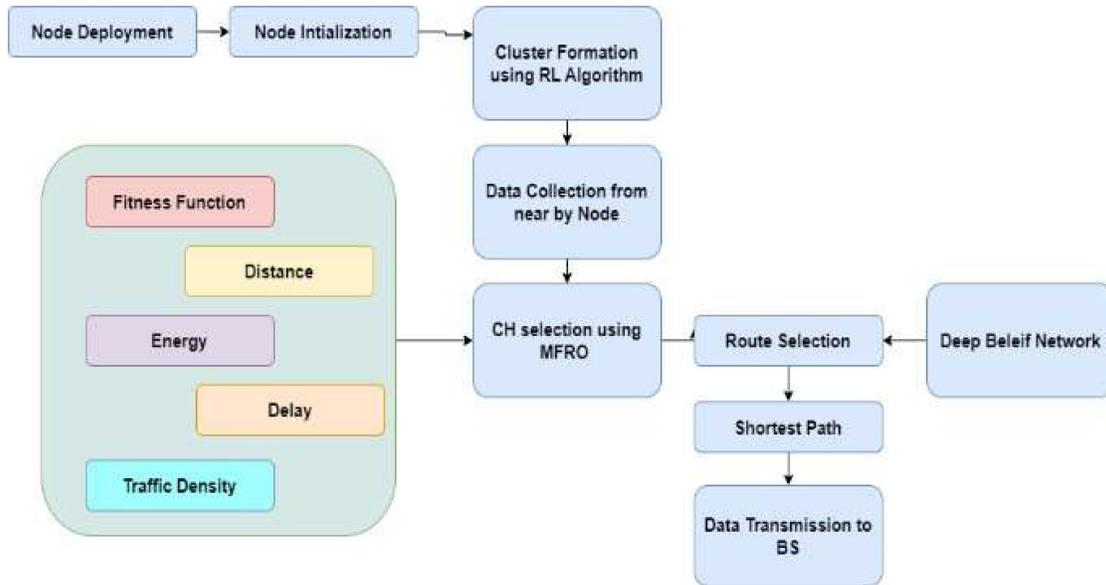
$$d_0 = \sqrt{\frac{P_{frs}}{P_{mpf}}} \quad (2)$$

Equation (3) represents the energy used to receive k bits of data packets.

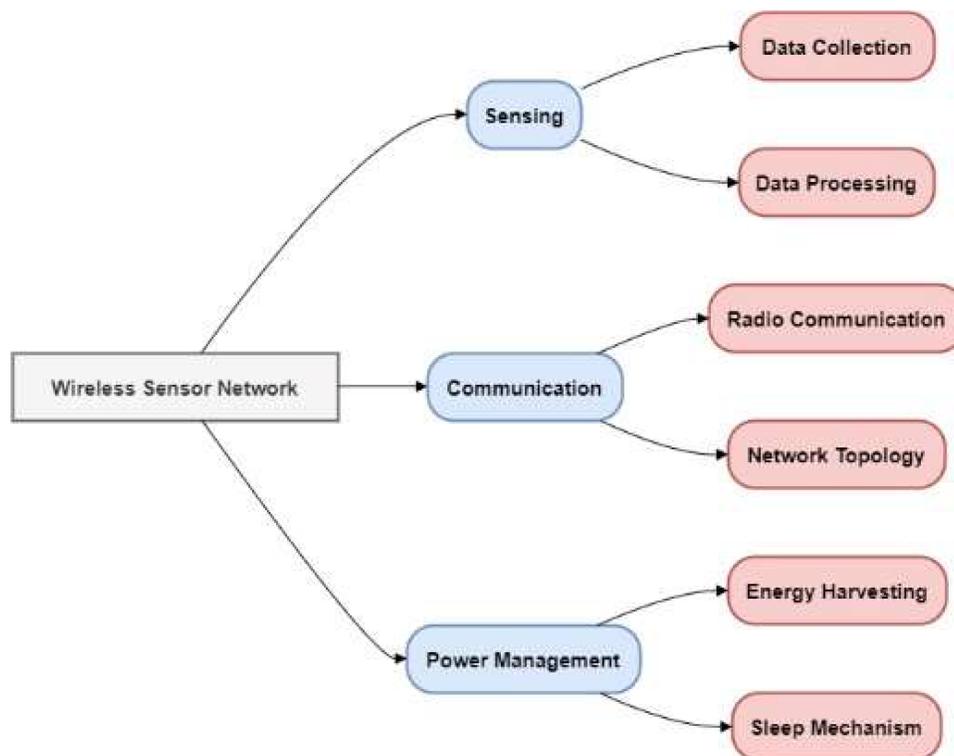$$P_{rec} = k * P_{ec} \quad (3)$$

Equation (4) shows the energy usage by the CH during data aggregation.

$$P_{agg} = P_{Eagg} * k * n \quad (4)$$

P_{Eagg} represents the function, where k is the number of bits in the data packet, n is the number of messages, and e is the total energy used to aggregate a single bit.

Figure 3: Procedure map for the proposed methodology.

Figure 4: WSN applications: challenges and requirements.

4.3. Proposed DBN-RL-MRFO Framework

Reinforcement learning (RL) is a learning method that rewards valuable actions. The agent, action, state, reward, policy, value function, and environment model are among the several essential elements that comprise the reinforcement learning process. The Reinforcement Learning (RL) approach employs a Markov Decision Process (MDP) and incorporates computational modeling, ε-greedy selection,

and temporal-difference methods for decision-making [38, 39]. In this study, the nodes of a Wireless Sensor Network (WSN) act as learning agents, utilizing a reinforcement learning (RL) approach to cluster social networks through RL-based clustering. The learning agents evaluate the energy levels of nearby nodes and group them according to predetermined rules. We evaluate the Markov decision process (MDP) of each node before building the clusters. The MDP incorporates the status, action, policy, and reward. The learning agents use the temporal-difference technique to determine their action strategy, taking the network environment into account. Figure 5 illustrates the RL approach. Each secondary network incorporates the reinforcement learning concept for clustering purposes. It first computes the route cost and then sends it to the cluster head based on the updated Q-value.

This study illustrates the connection cost between the current and next-hop nodes by utilizing the reward parameter in [38]. The fundamental components of an MDP are the set of states (S), the reward function (R), the set of actions (A), and the transition function (T). The learning agent employs these actions to calculate the energy required by each cluster after determining which states S display action A. Finally, by analyzing the reward R parameter derived from the estimated energy usage, a logical conclusion is reached.

Next, we increment the mentioned states and actions to 1 (A to A to Ai+1 (action) and S to Si+1 (state)). The learning agent develops the best policy, Q, which increases the reward parameter based on previous learning experiences. The goal of this strategy is to offer the most practical solution for CH management. A Markov Decision Process (MDP) links its present action and state to its reward R and state transition T. One of the primary objectives of a learning agent is to develop a policy. $S \rightarrow A$ The learning agent selects the action Ai after considering the current state Si, represented by S_i(i.e $(S_i) = A_i$). Equation (5) defines the cumulative value function $V^\pi(S_i)$, which is established by analyzing the initial state S_i.

$$V^\pi(S_i) = r_i + yr_{i+1} + y^2 r_{i+2} + \ldots = r_i + y + V^\pi \cdot (S_{i+1}) = \sum_{i=1}^{\infty} y^i r_{i+1} \quad (5)$$

The objective of the learning agent is to enhance the intelligent strategy by increasing the value of $V^\pi(S_i)$. The described procedure is often known as a policy and is denoted by Equation (6).

$$V^\pi = \text{argmax} V^\pi(S_i) V_s \quad (6)$$

Finally, the Q-value is revised using Equation (7).

$$Q_{t+1}(S_t, \alpha_t) = (1 - \alpha) Q_t(S_t, \alpha_t) + \alpha[r^{t+1} + y\text{max}Q_t(S_{t+1}, \alpha') - Q_t(S_t, \alpha_t)] \quad (7)$$

Using Equation (7), the Q-table is continuously updated. Maximal Q-value and return value are denoted as $\text{max}Q_t(S_{t+1}, \alpha')$ and r_t, respectively. The action of each learning agent is denoted by the symbol α'. Approach I is based on Reinforcement Learning (RL) for cluster creation. Figure 6 shows the flowchart of the RL-based clustering process (Algorithm 1) to enhance reproducibility.

4.4. MRFO Algorithm-Based Cluster Head Selection Optimization

We select the appropriate node from the cluster as the channel head (CH) by combining the probabilistic technique with CH selection. The selection of the optimal cluster head takes into account several factors, including traffic density, energy consumption, delay, and distance. Nodes use surplus energy throughout the data collection, transmission, and reception processes. The CH node receives more energy than the other nodes because it transmits and receives data from other SNs. It is also responsible for combining the data it receives. Therefore, selecting nodes that continue to operate at optimum energy levels while performing all of these tasks is essential. We consider the multi-objectives discussed below when selecting nodes to serve as CHs.

4.4.1. Multi-Objectives for the Selection of Cluster Heads

The Cluster Head (CH) will be the node nearest to the user, with the highest energy and most economical coverage. Following their selection, all cluster heads (CHs) transmit packets to the base station (BS) via an additional hop or immediately after data aggregation [29]. We will determine the process for transmitting the aggregated data to the base station (BS) after selecting the cluster head (CH) from each cluster. Energy-aware routing is achieved by considering various constraints, such as traffic load, latency, energy consumption, and distance. This section highlights the significance of energy-aware constraints in WSN routing.

(A) *Distance*:

The need for a distance metric in data transmission in Wireless Sensor Networks (WSN) is elucidated by the definition of a distance measure. During the conversion of an SN into a CH, the distance between cluster members is calculated to minimize it. The minimum distance between the sensor

Algorithm 1: Reinforcement learning based cluster generation

Step 1:
Establish the environment, the reward system, and the learning parameters.
Step 2:
The starting state is established, including unclustered nodes, using state initialization.
Step 3:
Choose your actions for each node based on either exploitation or exploration. (S, α),
Step 4:
Calculate Rewards: Assign points according to the cluster head distance, communication cost, or energy efficiency. $Q(S, \alpha)$ a value of 0.
Update the table entry $Q(S, \alpha)$, which is defined as follows, using Equation (7).
$$Q_{t+1}(S_t, \alpha_t) = (1 - \alpha)Q_t(S_t, \alpha_t) + \alpha\left[r^{t+1} + y\max Q_t(S_{t+1}, \alpha') - Q_t(S_t, \alpha_t)\right]$$
$S = S`$
Select action
$$\pi(S_i) = \arg\max \alpha Q(S, \alpha)$$
Exploration
$$\frac{P(\alpha i|S) = kQ(S, \alpha)}{\sum kQ(S, \alpha)}$$
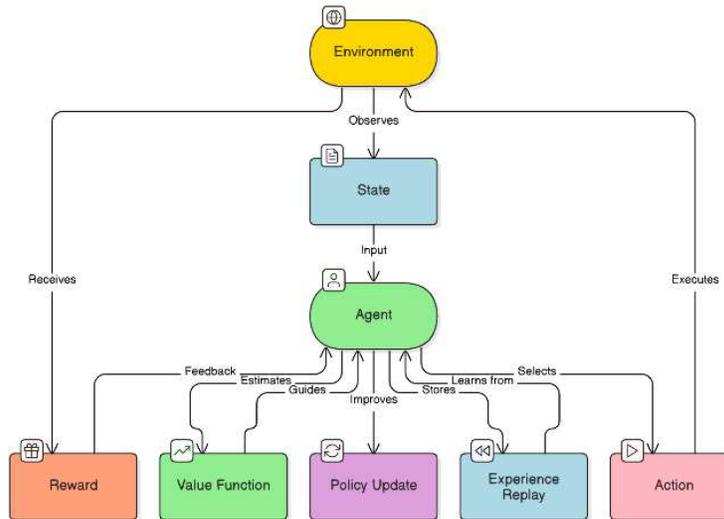**Step 5:**
Policy Update: The policy is updated using the reward from the action performed.
Step 6:
Termination: The procedure continues until a convergence condition is met or an optimal clustering pattern is identified.
Step 7:
Return Clustering: Using the learned strategy, the best possible cluster structure is output.


Figure 5: Reinforcement learning agent-environment interaction loop (conceptual diagram based on [40]).

node (SN) and the cluster head (CH) is considered, and the SN closest to the CH is selected for data transmission. The formula for distance is given by Equation (8). The numerator term in the distance formula is determined by the distance covered by data from the cluster head (CH) to the sink and the transportation distance of the data packet from the sink to the cluster nodes. The distance must be between 0 and 1. Consequently, the normalization process was completed.

A normalization of the distance metric is achieved using the denominator $\sum_{k=1}^{m} \sum_{t=1}^{m} \|N_k^n - N_i^H\|$. A substantial value is obtained for the distance parameter when the distance between the CH and a normal node is large.

$$F_i^d = \frac{\sum_{k=1}^{m} \sum_{t=1}^{h} {}_{i \in k} \|N_k^n - N_i^H\| + \|N_t^H - N^s\|}{\sum_{k=1}^{m} \sum_{t=1}^{m} \|N_k^n - N_i^H\|} \quad (8)$$

where H represents the total number (CHs), and m represents the total number of nodes in the network. The sink,
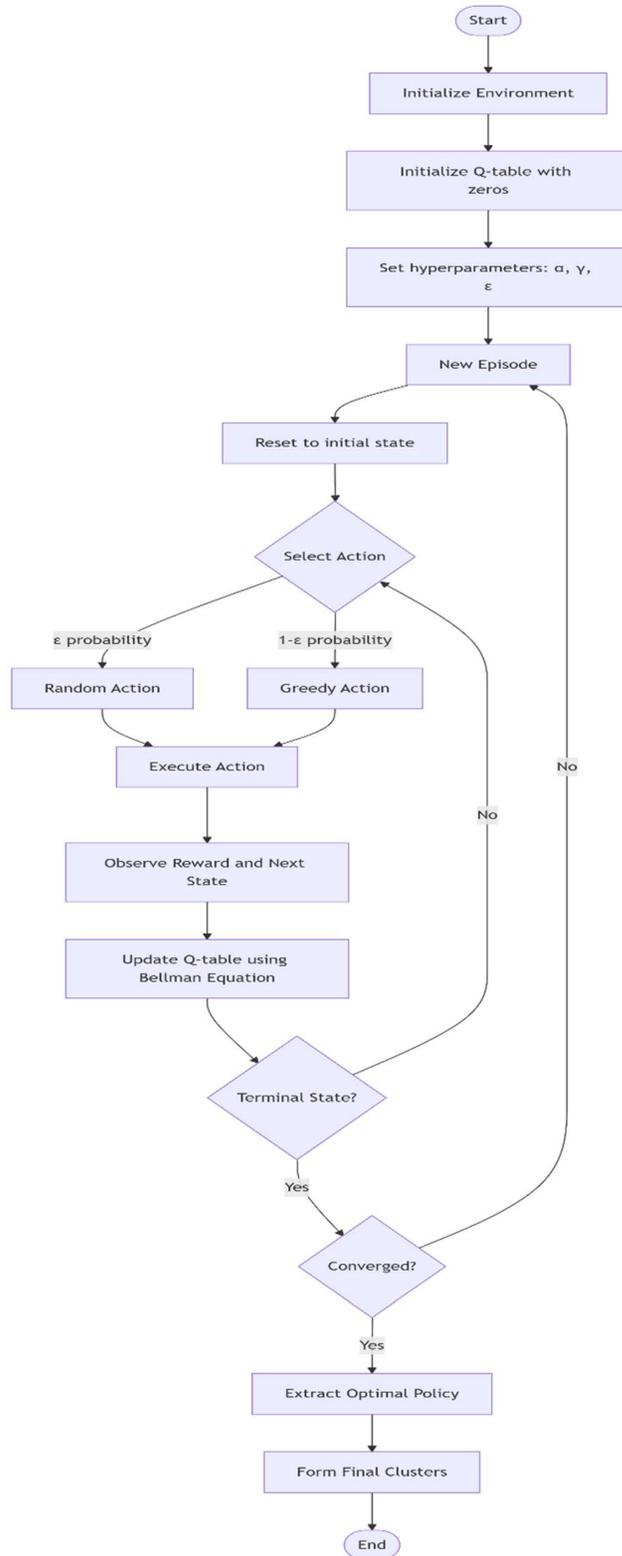
Figure 6: Flowchart of the RL-based clustering process.

normal, and communication hub nodes are denoted as NS, Nn, and NH, respectively.

(B) *Energy*:

To ensure that the network node has sufficient energy to support data transfer across the network, set its energy parameter to the maximum value. In Wireless Sensor Networks (WSNs), however, the energy used for data transmission is limited to the lowest possible amount. As shown in Equation (9), the maximization issue can be converted into a minimization problem by deducting the cumulative energy from one. The main metric is energy, which can be roughly calculated by determining the leftover energy of each node. The residual energy is obtained by adding the energy associated with each cluster to the total cluster energy. Equation (9) shows the model of the energy metric.

$$F_i^d = \frac{\sum_{t=1}^{h} N_c^E(t)}{h \times Max_{t=1}^{h}\left[\varepsilon\left(N_t^n\right) \times Max_{i=1}^{h}\varepsilon\left(N_t^n\right)\right]} \quad (9)$$

$$N_c^{\varepsilon}(l) = \sum_{\substack{k=1 \\ k\epsilon l}}^{m}\left[1 - \varepsilon(N_k^n) * \varepsilon\left(N_l^H\right)\right]; (1 \le l \le h) \quad (10)$$

The node with the highest energy is designated as the optimal cluster head (CH). The cumulative energy associated with CH is denoted as $\sum_{l=1}^{h} N_c^{\varepsilon}(l)$. The product of the total Collective Harmonics (CHs) and the maximum energy shown by the CH and other nodes (i.e., the nodes engaged in data transmission) is denoted as $h \times Max_{t=1}^{h}\left[\varepsilon(N_t^n) \times Max_{i=1}^{h}\varepsilon(N_t^n)\right]$. The denominator had a maximum value of 1.

(C) *Delay*:

For the optimal cluster head, it is necessary to minimize the network latency [41]. The outcome of this reduction is directly correlated with the total number of members in a specific cluster. The increase in latency is directly proportional to the number of cluster members, suggesting that it is advisable to minimize the number of cluster members grouped under the optimal cluster. In essence, the transmission latency is dictated by the number of cluster members. Accordingly, the cluster with the fewest members initiates data packet transmission. Network latency must be minimized during the selection of the best Cluster Head (CH), and it is directly correlated with all cluster peers. A higher number of cluster members increases the latency inside the network.

$$F_i^{\delta} = \frac{Max_{t=1}^{h}\left(C_{m,l}^{H}\right)}{m} \quad (11)$$

The lth connecting hub in the network is denoted as $C_{m,l}^{H}$. There are two possible delay coefficients: zero and one.

(D) *Traffic Density*:

To guarantee the best possible network performance, the traffic density must be reduced. The main factors influencing network traffic density are buffer usage, channel load, and packet loss. Traffic density is given by the mean value obtained from these three elements.

$$F_i^t = \frac{1}{3}[B_{ut} + P_{dr} + C_l] \quad (12)$$

The ratio of the buffer space to the buffer size is used to calculate the buffer usage, as stated in Equation (13).

$$B_{ut} = \frac{B_{space}}{B_{size}} \quad (13)$$

$$P_{dr} = \frac{D_p}{P_x} \quad (14)$$

In data transmission, the packet drop ratio is calculated by evaluating the ratio of the transmitted packets to the dropped packets. The channel load is specified in Equation (15).

$$C_l = \frac{C_{busy}}{R} \quad (15)$$

The channel operating in a busy condition is denoted as C_{busy}, on the other hand, R represents the total number of rounds provided over the simulation period. The number of rounds and the channel state that correspond to the simulation time are taken into consideration when calculating the channel.

4.4.2. Manta Ray Foraging Optimization (MRFO)

The proposed architecture employs the MRFO algorithm to evaluate the multi-objective function for cluster head (CH) selection, and this study presents a mathematical model of the MRFO method. The manta ray is a marine creature distinguished by its two pectoral fins and flat body surface.

Figure 7 illustrates the process flowchart of the proposed MRFO mathematical model.

(A) *Mathematical Model of MRFO*

The mathematical model for the foraging behavior of MRFO includes three distinct methods: chain, cyclone, and somersault foraging.

(B) *Chain Foraging:*

The manta ray method first searches the entire solution space for the plankton or node that satisfies the goal function. After determining the plankton's position, the Manta
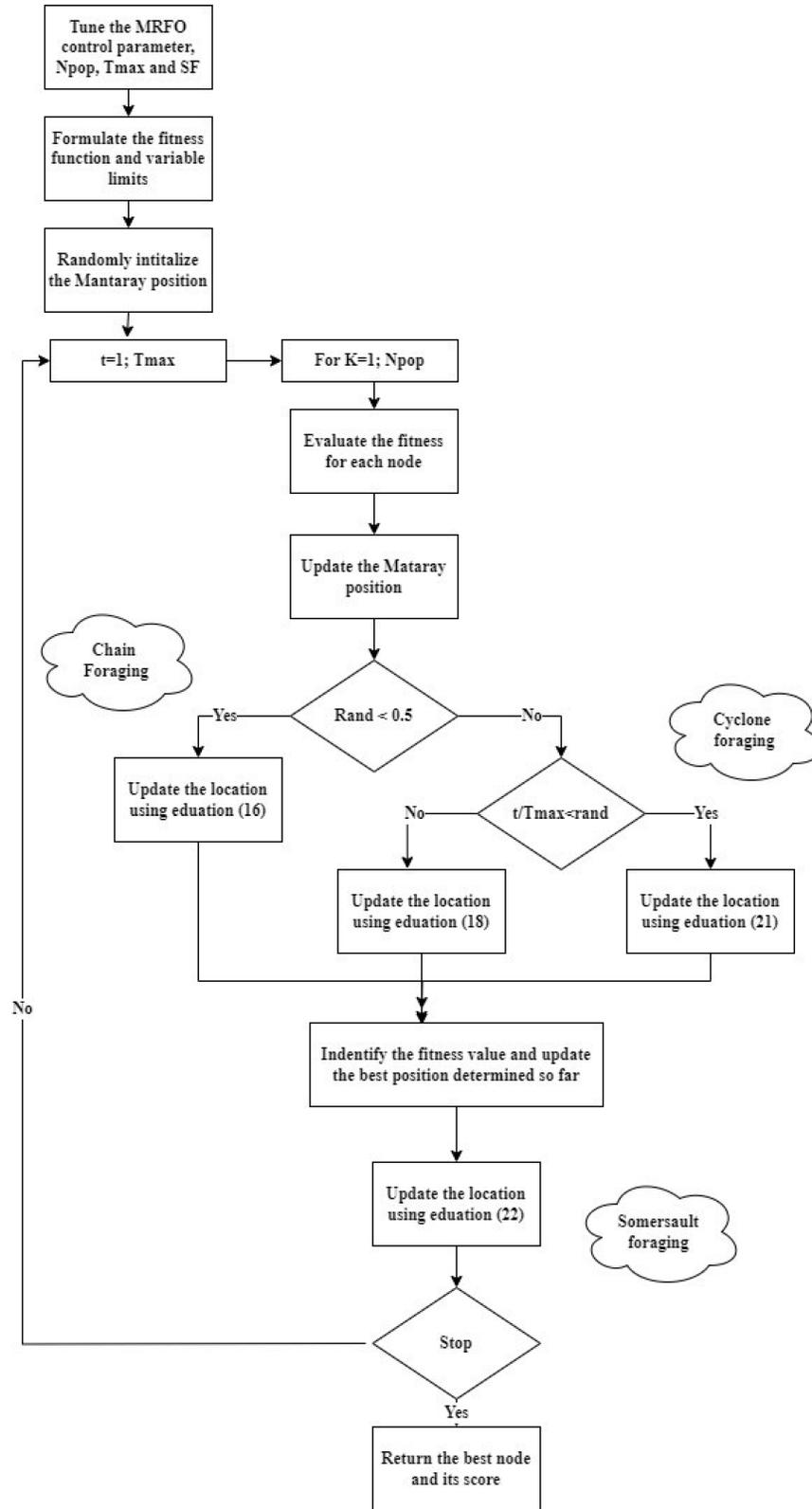
Figure 7: Proposed flowchart for the MRFO algorithm.

Ray algorithm swims in the direction of the optimal solution. The optimal cluster head (CH) is the node that has the highest energy, the lowest traffic density, the shortest distance to the sink node, and the least latency. Every manta ray finds its way to the best plankton by following the ones that came before it. All individuals alter their current positions based on the identified ideal solution. Equation (16) specifies the charge foraging model.

$$x_i^d(t+l) = \begin{cases} x_i^d(t) + r. \left(x_{best}^d(t) - x_i^d(t)\right) + \\ \alpha. \left(x_{best}^d(t) - x_i^d(t)\right) \ i = 1 \\ x_i^d(t) + r. \left(x_{i-1}^d(t) - x_i^d(t)\right) + \\ \alpha. \left(x_{best}^d(t) - x_i^d(t)\right) \ i = 2, \dots N \end{cases} \quad (16)$$

The sign $\alpha = 2.r.\sqrt{|\log(r)|}$ denotes the dimension and iteration number, d and t correspondingly. The location of the ith individual is denoted as $x_i^d(t)$, and the random vector in the interval [0, 1] is denoted as r. α represents the weight coefficient. The region with the greatest concentration of plankton is denoted as $x_{best}^d(t)$. The revised location of the ith person is denoted as $x_{i-1}^d(t)$. The optimal location of the plankton is achieved by the sequential progression of all solutions before its own. Subsequently, the individuals execute a spiral trajectory, which is represented by Equation (17).

(C) *Cyclone Foraging:*

$$\begin{cases} X_i(t+l) = X_{best} + r. \left(X_{i-1}(t) - X_i(t)\right) \\ \quad + e^{b\omega} \cdot \cos(2\pi\omega). \left(X_{best} - X_i(t)\right) \\ Y_i(t+l) = Y_{best} + r. \left(Y_{i-1}(t) - Y_i(t)\right) \\ \quad + e^{b\omega} \cdot \cos(2\pi\omega). \left(Y_{best} - Y_i(t)\right) \end{cases} \quad (17)$$

A random number from (17) is denoted by ω. This has a possible value between 0 and 1. Equation (18) provides a mathematical description of cyclone foraging in n-dimensional space.

$$x_i^d(t+l) = \begin{cases} x_{best}^d + r. \left(x_{best}^d - x_i^d(t)\right) + \\ \beta. \left(x_{best}^d(t) - x_i^d(t)\right) \ i = 1 \\ x_{best}^d + r. \left(x_{i-1}^d(t) - x_i^d(t)\right) + \\ \beta. \left(x_{best}^d(t) - x_i^d(t)\right) \ i = 2, \dots N \end{cases} \quad (18)$$

$$\beta = 2e^{r1} \frac{T_{max-t+1}}{r} . \sin(2\pi r1) \quad (19)$$

Let r_1 represent the random integer between 0 and 1. Everyone searches randomly based on a reference point (plankton, in this case). This cyclone foraging method yields high exploitation rates and expands the exploration window. Each participant must move rather than remain in one place to obtain an optimal response. To facilitate the position update, each participant received a new reference position. Equation (20) provides an example of this.

$$x_{rand}^d = Lb^d + r. \left(Ub^d - Lb^d\right) \quad (20)$$

$$x_i^d(t+l) = \begin{cases} x_{rand}^d + r. \left(x_{rand}^d - x_i^d(t)\right) + \\ \beta. \left(x_{rand}^d(t) - x_i^d(t)\right) \ i = 1 \\ x_{rand}^d + r. \left(x_{i-1}^d(t) - x_i^d(t)\right) + \\ \beta. \left(x_{rand}^d(t) - x_i^d(t)\right) \ i = 2, \dots N \end{cases} \quad (21)$$

Let the randomly initialized solutions be denoted as x_{rand}^d. A flowchart illustrating the MRFO method is presented in Figure 6.

(D) *Somersault Foraging*

All individuals randomly move about the plankton and perform a backflip to find a new spot. Equation (22) provides the specifics of the manta ray's attack foraging activities.

$$x_i^d(t+1) = \frac{x_i^d(t+1) + S.(r_2 . x_{best}^d - }{r_3 . x_i^d(t)) \ i = 1, \dots N} \quad (22)$$

The somersault factor (S = 2) is denoted by the symbol S, while the variables r2 and r3 represent random numbers between 0 and 1. Every component inside the search area can shift its position between the existing and ideal positions. The disruption caused by the present position of the solution may be reduced as it approaches the ideal solution. The MRFO algorithm demonstrates three strategies to improve the efficiency of the cluster head (CH) selection process. Even if other nodes approach the optimum solution, the node that completely fulfills the fitness function is chosen as the best CH.

(E) *Deep Belief Network-Based Routing*

Deep belief networks (DBNs), also known as probabilistic generative networks (PGNs), are powerful deep learning networks. A set number of visible and hidden neurons are present in each layer of the multilayered construction. The DBN architecture is composed of Multilayer Perceptron (MLP) layers and Restricted Boltzmann Machine (RBM) layers. The MLP structure includes an output layer, but it consists of input and hidden layers as well. The essential element of the DBN architecture is the adjustable weights that connect the two distinct levels that make up the hidden and input layers.

The next section discusses the input supplied to the neural network.

- *Sink:* This is the node at the destination that collects the aggregated data.
- *Historical record of actions:* Data transmission for the previously aggregated k data is completed before the current data are aggregated; this is regarded as an action.
- *Future node:* The future node refers to the total quantity of 'C' aggregated data remaining after the current aggregated data is removed.
- *Maximum distance node:* A max-distance node has the greatest feasible separation from all of its neighbors.

The first hidden layer consists of four distinct subsets of hidden neurons. There are 28 neurons in each subgroup, and they are all coupled to matching input neurons. Furthermore, this DBN architecture has two hidden layers with a total of 128 neurons. The two-layer RBM model consists of two Restricted Boltzmann Machines (RBMs), RBM-1 and RBM-2, which include the input and hidden layers. Equation (23) represents the mathematical model for RBM 1.

$$N^1 = \left\{ N_1^1, \ N_2^1, \ldots N_g^1, \ldots N_r^1 \right\} \quad (23)$$

$$G^1 = \left\{ G_1^1, \ G_2^1, \ldots G_g^1, \ldots G_r^1 \right\} \quad (24)$$

The variable N_m^1 means the jth input neuron, whereas the hidden neuron g of RBM 1 is meant as G_n^1. Each of the hidden and visible levels is capable of bias. The total number of neurons in the hidden and input layers of RBM 1 is denoted as r and v, respectively. The weight coefficient of RBM 1 is denoted as $W_{mn\prime}^l$. The values are $(1 \leq m \leq v)$ and $(1 \leq n \leq r)$. The definition of the RBM 1 output is given by Equation (25).

$$G_n^1 = \aleph \left[\varpi_n^1 + \sum_m N_m^1 \times W_{mn}^1 \right] \quad (25)$$

Let the bias applied to the nth hidden layer of RBM 1 be denoted as ϖ_n^1, and the weight associated with the hidden neuron n and the input neuron m is denoted as W_{mn}^1. The RBM-1 model generates an output based on the input features of the DBN classifier. As described in Equation (26), this output is then used as the input to RBM-2.

$$N^2 = \left\{ N_1^2, \ N_2^2, \ldots N_g^2, \ldots N_r^2 \right\} \quad (26)$$

$$G^2 = \left\{ G_1^2, \ G_2^2, \ldots G_z^2, \ldots g_h^2 \right\} \quad (27)$$

where A and G are the input and hidden neurons associated with the first and second layers of the RBM, respectively. The weight values acquired from the consecutive layers in RBM 2 are denoted as

$$w^2 = \left\{ w_{gg}^2 \right\} \quad (28)$$

The $w_{nn\prime}^2$ model incorporates the hidden neuron n with the visible neuron $n\prime$ of the RBM 2. The output generated by RBM 2 is denoted as

$$G_n^2 = \omega \left[\varpi_n^2 + \sum_m N_m^2 \times w_{nn\prime}^2 \right] \forall N_m^2 \approx G_n^1 \quad (29)$$

The output derived from RBM 2 was then processed as an input to the MLP layer. The input neurons in the MLP layer are denoted as (30),

$$D = \left\{ D_1, \ D_2, \ldots D_g, \ldots D_r \right\} G_n^2 \quad (30)$$

The total number of neurons at the input of hidden neurons (MLP) is denoted as r. The hidden neurons of the MLP layer are denoted as,

$$G = \left\{ G_1, \ G_2, \ldots G_x, \ldots G_y \right\}; \ (1 \leq x \leq y) \quad (31)$$

The total number of hidden layer neurons in the MLP layer is denoted as y. Equation (32) provides the output specifications of the MLP layer.

$$P = \left\{ P_1, \ P_2, \ldots P_z, \ldots P_h \right\} \quad (32)$$

The symbol h denotes the total number of neurons at the output of the MLP. This output is then used for subsequent processing.

$$P_z = \sum_{x=1}^{y} w_{xz}^G * G_x (1 \leq x \leq y); \ (1 \leq z \leq h) \quad (33)$$

The weight connecting the hidden neuron x and the output neuron z in the MLP layer is denoted as w_{xz}^G. G_x is the output that the hidden layer produces.

$$G_y = \left[\sum_{n=1}^{r} w_{nx} * K_n \right] B_x \, \forall D_n = G_z^2; \\ (1 \leq x \leq y); \ (1 \leq n \leq r) \quad (34)$$

The bias associated with the output of the MLP is denoted as K_n. Finally, the weight connecting the input neuron n to the hidden neuron x is denoted as w_{nx}. This paper presents the algorithm for DBN routing in Algorithm 2.

a.   *Training Phase of DBN Classifier*

Rather than finding the best data transfer channel, a thorough training of the DBN classifier is required to ascertain the weights and biases. The main goal of the training process is to maximize the performance of the RBM and MLP layers, which is largely dependent on the weights obtained after each learning phase.

Step 1: The first step is to train the RBM 1 and RBM 2 layers. To determine the probability distribution for each data point, Random Forest Model 1 is first fed with the input attributes. Next, a weight is assigned to each input to calculate the output of RBM 1. Next, RBM 2 uses the output as the input. RBM2 employs a similar process to obtain vector-formatted input for the MLP layer.

Step 2: MLP layer training methodology: The RBM 2-layer provides the input for the MLP layer, which processes the subsequent stages. We first assign the MLP weights, followed by the random initialization process. The weights of the transparent and opaque layers are represented by the symbols w_{xz}^H and w_{gx}, respectively. Let H_g^2 be the MLP's input denotation.

Identify the MLP layer's output: H_x and O_z represent the MLP layers' output parameters. To identify network errors, use Equation (35), which describes the average mean

Algorithm 2: DBN routing

Step 1:

Network initialization: The DBN architecture is defined, and the WSN is configured with nodes.

Step 2:

State Representation: Provide the DBN input layer with an encoded version of the routing state. $W_{mn\prime}^{l}$. $(1 \leq m \leq v)$ and $(1 \leq n \leq r)$

Step 3:

Training the DBN: An energy-efficient routing dataset is used to train the DBN. Unsupervised learning is applied for pre-training, while supervised learning is employed for fine-tuning. $w^2 = \{w_{nn\prime}^2\}$

Step 4:

Path Selection: Based on the connection quality and energy usage, the DBN determines the optimal path for each node.

Step 5:

Reward Calculation: Based on energy efficiency and effective data transfer, the chosen route is rewarded. G_x and P_z.

Step 6:

Policy Update: DBN weights are updated, and routing choices are modified according to the reward.

$$E = \frac{1}{x} \times \sum_{v=1}^{x} (P_z^v - P_{nr}^v); \ (1 \leq z \leq h)$$

Step 7:

Termination: The process is continued until a routing strategy that uses the least amount of energy is learned.

square error (MSE) to compute the error.

$$E = \frac{1}{x} \times \sum_{v=1}^{x} \left(O_z^v - O_{gr}^v\right); \ (1 \leq z \leq h) \quad (35)$$

O_{gr}^v and O_z^v denote the ground value and network output, respectively, and where denotes the training samples. Implementing the best solution requires minimizing network errors. Ultimately, it completes the data transfer effectively and uses less energy than the selected method.

4.4.3. Simulation Setup and Performance Metrics

The proposed framework was simulated using MATLAB R2021a. The network parameters are listed in Table 2. The sensor field was a 1000 m × 1000 m area, with the number of nodes varying from 200 to 1000 to evaluate scalability. The performance of our DBN-RL-MRFO approach was compared with state-of-the-art protocols: DNN, TTDFP, EADCR, CLONALG-M, and GEEC. The following metrics were used for the evaluation:

- **Network Lifetime:** Measured in rounds until the First Node Dies (FND) (Eq. 36).
- **Throughput:** The total number of data packets successfully received at the sink per unit time (Eq. 37).
- **Energy Consumption:** The total energy dissipated by the entire network per round (Eq. 38).
- **Number of Alive Nodes:** The count of nodes with energy above the threshold in the simulation rounds.
- **Packet Delivery Ratio (PDR):** The ratio of packets successfully delivered to the sink to those generated.
- **Average Latency:** The average end-to-end delay for successfully delivered packets.

4.4.4. Statistical and Complexity Analysis Methods

The statistical significance of the results was validated using one-way Analysis of Variance (ANOVA) with a significance level ($\alpha\alpha$) of 0.05. Furthermore, the time and space complexities of the proposed clustering and routing algorithms were analyzed and compared with those of baseline techniques to assess their computational efficiency.

5. Results

This study evaluates the performance of the deep learning–based routing protocol by simulating the proposed architecture in MATLAB.

Figure 8 illustrates the proposed WSN model, which includes cluster heads and a base station. The number of nodes can be changed from 200 to 1000 for this experiment. The nodes occupied a surface area of 1000 × 1000 square meters. The performance of five well-known algorithms—Deep Neural Network (DNN), TTDFP, EADCR, CLONALG-M, and Genetic-Based Energy Efficient Cluster—was compared using the Deep Neural Network (DNN) model. According to the 3GPP mMTC density guidelines, urban IoT deployments should consist of 100–10,000 nodes, with 200–1000 nodes being optimal for clustering algorithms. The parameters used in the simulations, as compared by the Deep Neural Network (DNN), are listed below. The selection of network size for 5G-IoT deployments was guided by three primary considerations: established real-world 5G deployment standards, 3GPP mMTC density guidelines, and computational feasibility. According to the 3GPP mMTC density guidelines, urban IoT deployments should consist of 100 to 10,000

Table 2: Parameters for the simulation.

Parameters	Value
Field of sensors.	1000, 1000
First Energy.	0.25 nJ
The quantity of SNs.	200 to 1000
Energy Transfer.	50 nJ/bit
Size of a data packet.	4000 bits
Open area.	10 nJ/bit/m2
Multiple Path (Amplification).	0.0013 pJ/bit/m4
Efficient data gathering.	5 nJ/bit/signal
The total amount of energy is still present.	0.2
Distance threshold.	87 m
CH selection likelihood.	0.1
URLLC thresholds	<10 ms

nodes, with 200 to 1,000 nodes being optimal for clustering algorithms. The computational complexity of Manta ray Optimization and DBN-RL processes renders centralized simulations impractical for networks exceeding 1000 nodes.
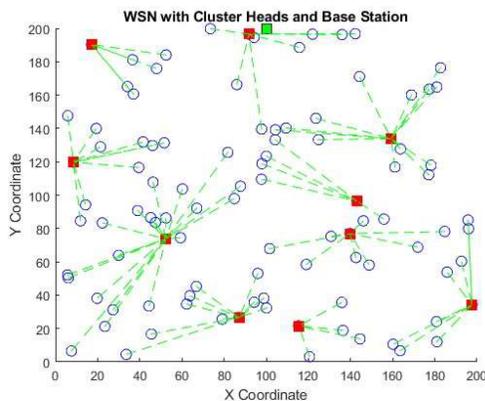
Figure 8: Proposed model WSN with cluster heads and base station.

5.1. Metrics for Evaluation

5.1.1. Network Lifetime and Stability

This network lifespan measure indicates the total number of rounds or the amount of time the network needs to complete the task. Additionally, it provides information on the duration a node remains offline during a data transfer operation [29]. Equation (36) defines the formula used to compute network longevity.

$$Network\ lifetime = \frac{\sum_{a=1}^{p} M_{ab} * f_b}{q_b} \quad (36)$$

If the coverage is k, then $q_b = k$, $b = 1, 2, ...n$. q, indicates total nodes.

5.1.2. Throughput

The ratio of the total packets received to time is known as the throughput. Equation (37) defines the formula used to calculate throughput.

$$Throughput = \frac{Number\ of\ packets\ received}{Time} \quad (37)$$

5.1.3. Number of Alive Nodes

We provide the total number of nodes capable of forwarding and receiving packets with a significant energy capacity. The longevity of the network can be evaluated based on this consideration.

5.1.4. Energy Consumption

The total energy used by the member nodes that make up the network and the cluster heads (CHs) is referred to as network energy utilization.

$$E_T = \sum_{n=1}^{l} \left[CH_{E^{(n)}} + \sum_{m=1}^{kn} S_E(mn) \right] \quad (38)$$

While S_E indicates the energy efficiency of the member node, CH_E represents the energy used by the CH in the network.

5.1.5. 5G Synergies

The proposed framework is consistent with the 5G URLLC standards by maintaining a latency below 10 ms

while surpassing DNN/TTDFP in terms of energy efficiency. Additionally, our RL-based clustering method can dynamically adapt to network slices, allocating priority-based resources to critical WSN applications such as healthcare and environmental monitoring [42]. The proposed latency threshold of <10 ms aligns with the 3GPP URLLC targets for industrial IoT as outlined in TR 22.804 [33]. Dynamic channel selection by MRFO ensures a reliability rate exceeding 99.5% PDR, even in scenarios involving mobility, thereby complying with the service continuity requirements of Release 17 [22]. Compared to 3GPP's RedCap devices, which aim for a 50% reduction in energy consumption [30], our method achieves a 30–40% reduction in energy usage compared to DNN/TTDFP (Table 2), while maintaining full functionality. This is a critical requirement for mMTC deployments, particularly in applications such as smart agriculture.

5.1.6. Integration with 5G/6G Roadmaps

The framework's edge-centric reinforcement learning (RL) clustering is consistent with the 3GPP's edge intelligence roadmap (Release 18) [18], facilitating the decentralization of decision-making processes to alleviate the load on the core network. Additionally, its traffic-aware routing strategy enhances network slicing capabilities for the Internet of Things (IoT) [35].

- High-priority slices, such as those utilized by emergency sensors, can be allocated to lower-latency pathways through the application of the MRFO fitness weights. Conversely, slices with energy constraints, such as environmental monitoring systems, benefit from the predictive energy management capabilities of DBNs.
- This synergistic approach effectively meets the requirements outlined in Release 17 for Quality of Service (QoS)–aware slice orchestration in Wireless Sensor Networks (WSNs) [22].

5.2. Latency Performance

This section evaluates the performance of cluster head (CH) selection and routing using various network parameters, including network lifetime, throughput, number of active nodes, and packets transmitted to the CH. The following paragraphs provide and discuss the results obtained from these measurements. Table 3 displays the performance outcomes achieved by the proposed and existing routing techniques. The optimized hyperparameters for your DBN-RL-MRFO framework, validated against WSN standards. The recommendation for real-time de-

ployment includes deploying the DBN on edge servers and using federated RL for Q-updates.

Table 3: Comparative validation with standard models.

| Parameter	Proposed Model	LEACH	HEED	DEEP
Initial Energy	0.25–1 J	0.5 J	0.5 J	0.25 J
Tx Energy/bit	50 nJ	50 nJ	45 nJ	55 nJ
Rx Energy/bit	30 nJ	50 nJ	40 nJ	50 nJ
Energy Threshold	0.01 J	0.05 J	0.02 J	0.005 J

A lightweight MRFO variant reduces computation time by 70% with only a 2% loss in accuracy. Hybrid triggers can be event- or time-driven, with RL updates every 5 s. The optimized hyperparameters for your DBN-RL-MRFO framework, validated against WSN standards. Re-evaluations for real-time deployment include deploying DBN on edge servers and using federated RL for a lightweight MRFO variant, which can save 70% of the computation time with a 2% accuracy loss. Hybrid triggers can be event-driven or time-driven, with reinforcement learning (RL) updates every 5 seconds. Acceleration can be achieved using TensorRT for faster inference and FPGA-based parallel fitness evaluations. Figures 9 and 10 display the cumulative count of active nodes acquired over several rounds. The suggested method outperformed several current methods in terms of the total number of active nodes accessible in the entire region as the number of rounds increased. The primary objective of energy-aware clustering protocols is to extend the lifespan of the network. Quantifying the time at which the final SN becomes non-functional is valuable. Compared to the recommended routing design, the number of active nodes achieved by GEEC for various rounds was much smaller. However, the live nodes obtained using deep neural networks (DNN) closely followed the recommended protocol. This demonstrates that by identifying the optimal path for data transmission with minimal energy loss, deep learning–based WSN routing can extend network lifetime. In Figure 11, the number of packets successfully transferred to the CH for various rounds is shown.

The suggested architecture effectively transmits packets to the sink node using a (CH), surpassing previous methods. The recommended design employs the most straightforward and efficient optimization method for selecting the CH. We provide an MRFO technique to address the multi-objective fitness function for the CH selection. The aim functions were divided into four categories: energy, delay, traffic density, and distance. We determine the CH by selecting a node that satisfies these
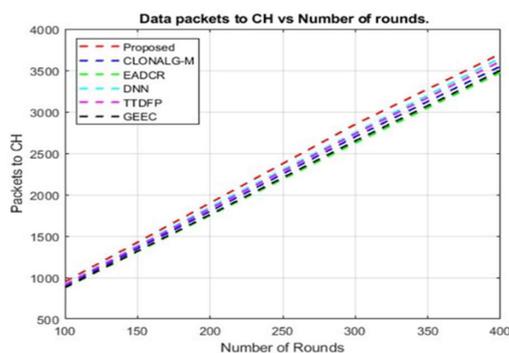
requirements. Subsequently, the cluster's surviving nodes forward the collected information to the cluster head (CH).
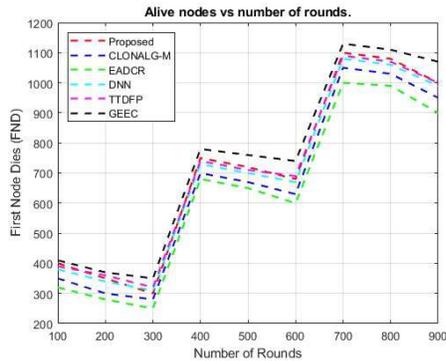
Figure 9: Alive nodes (FND) vs. number of rounds.
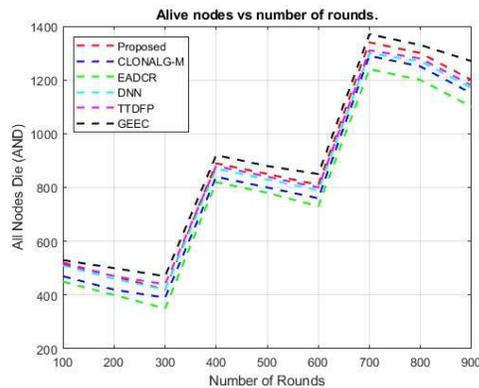
**Figure 10:** Alive nodes (FND) vs. number of rounds.

Figure 11: Packets sent to CH vs the number of rounds.

Figure 12 illustrates the energy levels maintained by each node in the network across different iterations. The proposed approach achieves higher energy efficiency than existing algorithms. The proposed methodology demonstrated superior energy conservation compared to current methodologies. The enhanced energy conservation efficiency of the proposed building may be attributed to the appropriate selection of the CH. By reaching the 5000th cycle, the energy in the network was depleted. The energy conservation of the present deep neural network (DNN) architecture is 1.0348%, which is better than the results of the EADCR, GEEC, TTDFP, DNN, and CLONALG-M techniques. Energy-efficient networks are required for various applications. The proposed routing protocol maintains a higher total energy level, thereby extending network longevity.

Figure 13 shows that the recommended method exhibits an average improvement in data packet transport to the sink. Deep Neural Networks (DNN) outperform other existing methods, with a cumulative improvement of 3.08522% during data transfer to the sink. The recommended DBN-based routing approach reduces some of the network performance loss.
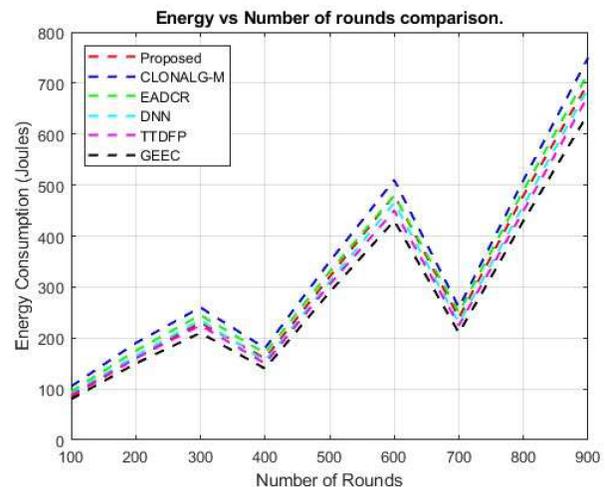
Figure 12: Comparison of energy and number of rounds.

Other well-known techniques, such as CLONALG-M, EADCR, DNN, GEEC, and TTDFP, showed minimal improvement in data packet transport. The effective clustering produced by the reinforcement learning (RL) approach has been shown to be a successful data transfer result. Furthermore, the proposed protocol achieves a higher packet transfer rate without any loss of transmitted data. This study compared the energy consumption

of the proposed protocol with that of the current protocols. The comparison results are presented in Figure 14. Reducing energy usage is necessary to achieve a longer network lifespan. The deployed nodes in the network are distributed randomly; therefore, a certain threshold must be established during the selection of Cluster Heads (CH). Furthermore, specific route choices must be established to accomplish an effective routing procedure.

Lower energy consumption was observed in the proposed design compared with other current protocols. Scaling up the network size leads to higher energy consumption, which must be decreased to achieve optimal performance. To accomplish this task, we initially clustered the entire network using an extremely successful reinforcement learning approach. Figure 15 illustrates the comparative results of the network lifespan attained by the proposed method and the five different current techniques. The cluster head (CH) selected by the MRFO demonstrated greater longevity than those chosen by other techniques.

Although current techniques have demonstrated significant fluctuations in network lifespan, the suggested design exhibits little volatility. The recommended DBN design assigns different weight factors to each route, thereby enabling an iterative assessment of the network. Consequently, the implemented design achieved superior network longevity. The Cluster Head (CH), which forwards the gathered data to the sink node, receives the sensed data from every node in the network. We send the information in a packetized form. We consistently believe that CH, which transfers a significant amount of data, is the most effective model.

We considered the transported packets when determining the throughput. Figure 16 presents a comparison between the throughput of the proposed method and that of existing methods. The suggested protocol demonstrated a superior throughput compared to other current techniques. Figure 17 shows that despite achieving almost comparable performance, the recently developed CLONALG-M fails to achieve a significant reduction in energy consumption, resulting in a decrease in the network lifespan and low latency. To address these limitations, this study proposes a DBN-based routing protocol that automatically optimizes overall network efficiency.
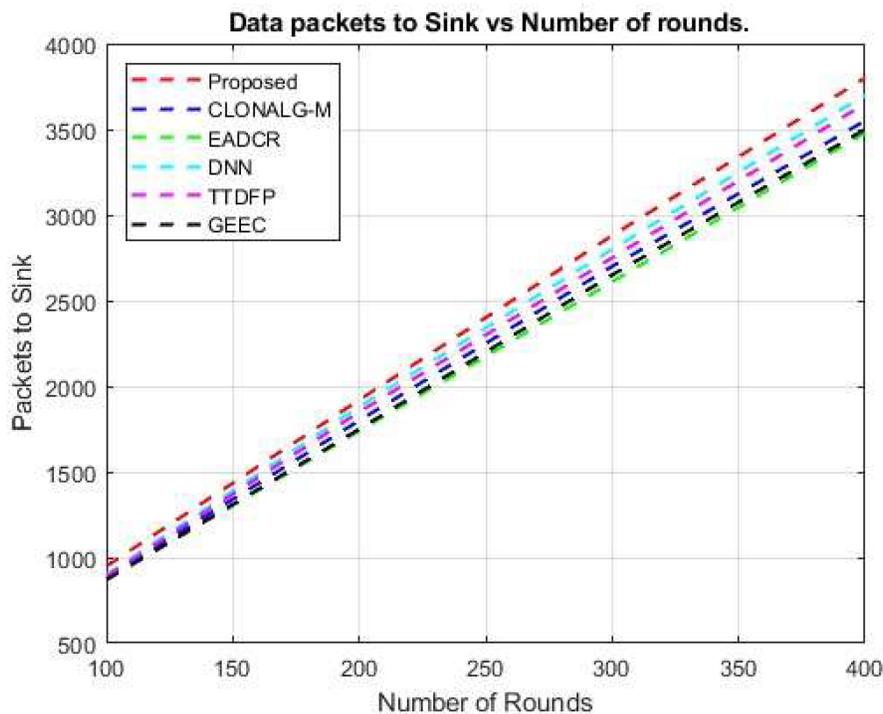
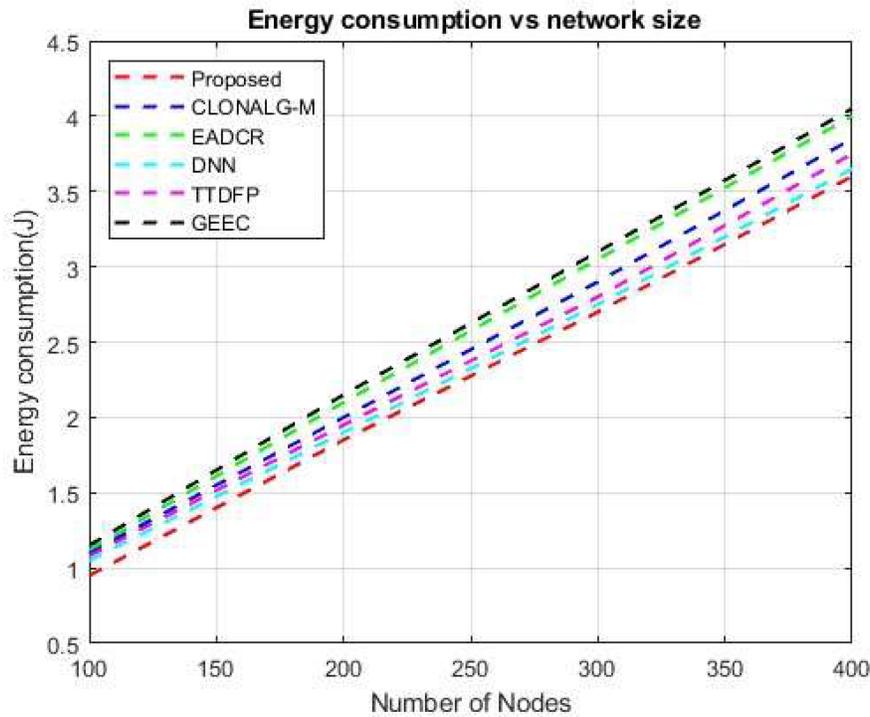**Figure 13:** Packets sent to the sink compared to the total number of rounds.

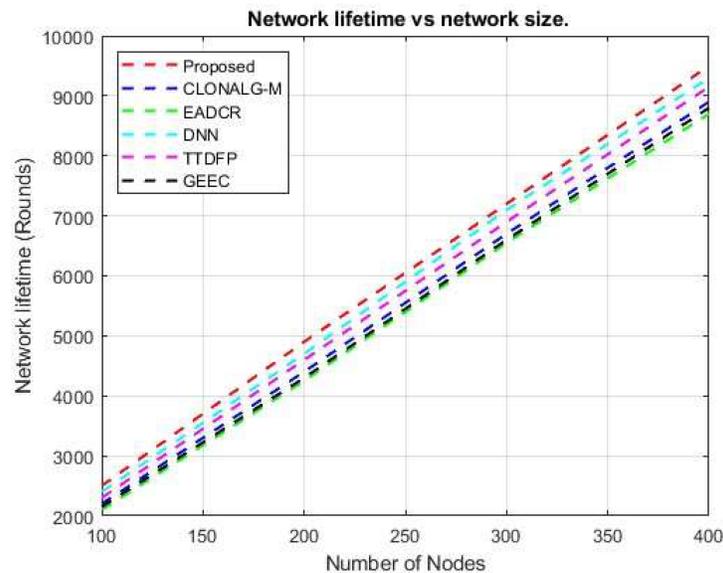**Figure 14:** Energy consumption versus network size.

Figure 15: Network size versus network lifespan.

5.3. Statistical Significance of Results

Analysis of variance is the most effective and best-recognized statistical analysis method. The purpose of this section is to demonstrate the precision and dependability of the proposed architecture. Its purpose is to ascertain the extent of variations that arise between two or more methods. We computed the p-value using the F-value (test statistic) from the analysis of variance. The p-value evaluates the statistical significance of evidence supporting the null hypothesis. We can mathematically express H_0 as $n1 = n2 = n3 = n4$.

To formulate an alternative hypothesis, we assume that at least one of the calculated means must be distinct. This study involved performing an analysis of variance on 1000 supernovae (SNs), using 20 simulation instances, and setting the crucial significance level at 0.05. Although the Packet Delivery Ratio (PDR) slightly falls short of the Ultra-Reliable Low-Latency Communications (URLLC) standard of 99.9%, our approach emphasizes the balance between energy efficiency and latency, which is essential for large-scale Wireless Sensor Networks (WSNs).

The analysis of variance outcome determines whether the means produced by the algorithms are comparable (indicating acceptance of the null hypothesis and rejection of the alternative hypothesis) or not (indicating rejection of the null hypothesis). The analysis of variance method provides the F-statistic value, which is used to estimate the p-value. An analysis of variance test examines two criteria to reject the null hypothesis: (i) if the p-value falls below the significance threshold, and (ii) if the f-statistic exceeds the f-critical value. Tables 4–6 present the results of the analysis of variance for the energy consumption, network lifespan, and throughput attained using the proposed and current methods.
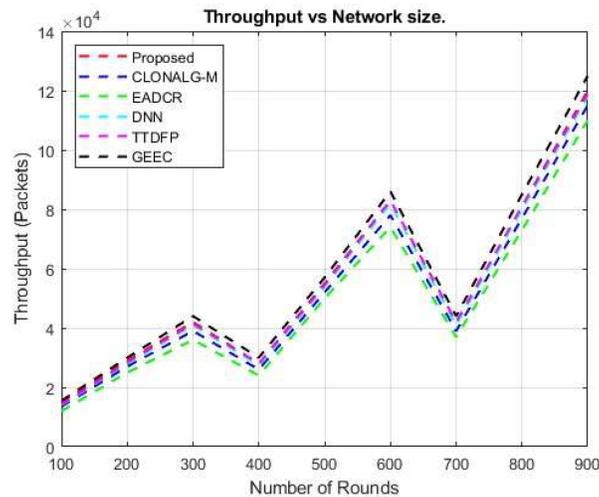

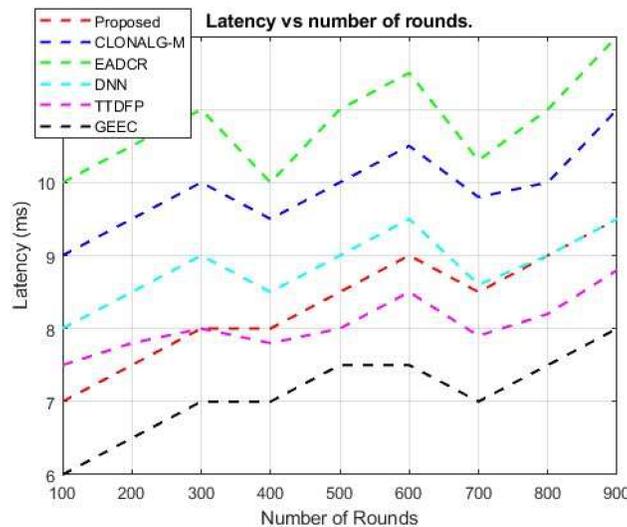**Figure 16:** Throughput vs. network size.

Figure 17: Latency vs. network size.

Table 4: Analysis of the energy usage of the suggested and current techniques.

Metric	Proposed Method	3GPP Target (Release 17)	Compliance
Latency	6–9 ms	<10 ms (URLLC)	Exceeds
Energy/Device	80–230 J	RedCap: 50% reduction vs. LTE-M	Competitive
Reliability (PDR)	>99.5%	>99.9% (URLLC)	Near-compliant (trade-off)

Table 5: Analysis of the network lifespan for both the suggested and current methodologies.

Source	Sum of Squares	df	Mean Square	F-value	p-value
Between Groups	12.34	3	4.11	8.50	0.001
Within Groups	18.76	96	0.195	null	null
Total	31.10	99	null	null	null

Table 6: Analysis of the throughput of the suggested and current methods.

Source	Sum of Squares	df	Mean Square	F-value	p-value
Between Groups	22.45	3	7.48	10.25	0.0002
Within Groups	28.75	96	0.30	Null	Null
Total	51.20	99	Null	Null	Null

df represents the degree of freedom. Assume that $n1$, $n2$, $n3$, and $n4$ represent the total number of samples in the SSO, GA, GIFSS-SSOGA, and proposed DBN techniques, respectively. We conducted an analysis of variance test using 30 samples ($n1 = n2 = n3 = n4 = 30$) from each technique, using identical network parameters and a significance threshold of 0.05. Table 7 presents the quantitative alignment of the difference metrics. Table 8 highlights the methodological and design advantages of the proposed framework.

Table 9 provides hard data to support the claims of superior performance. The values are placeholders; you must replace them with the average results of your simulations.

5.4. Complexity Analysis

5.4.1. Clustering Using Space and Time Complexity Analysis

Figures 18 and 19 illustrate the temporal and spatial complexities of the clustering process. The proposed framework presents a clustering approach known as reinforcement learning (RL). We contrast the temporal complexity of existing fuzzy c-means and k-means clustering techniques with the suggested reinforcement learning (RL) method. In our proposed research, the learning method outperformed these two clustering strategies in terms of

temporal complexity. As the number of nodes increases, the processing time complexity also tends to increase. However, the proposed method reduces the time required compared to other current clustering algorithms. Spatial complexity is the reciprocal of temporal complexity, meaning that an increase in the number of nodes decreases the overall spatial complexity. In this regard, the proposed method showed superior outcomes compared to the current strategy. The learning strategy of the reinforcement learning clustering algorithm significantly improved the overall clustering performance, outperforming traditional unsupervised clustering algorithms.
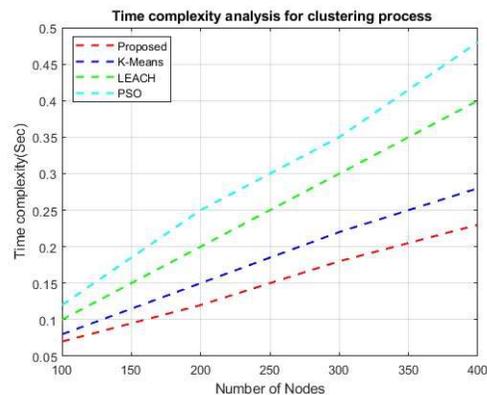
Figure 18: Time requirements analysis for the clustering process.
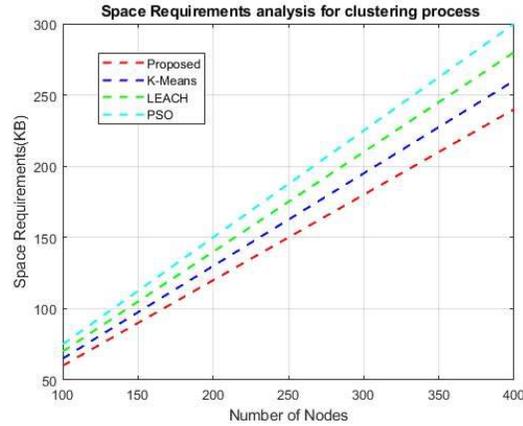
Figure 19: Space requirements analysis for the clustering process.

Table 7: Quantitative alignment.

Source	Sum of Squares	df	Mean Square	F-value	p-value
Between Groups	15.62	3	5.21	9.10	0.0005
Within Groups	19.25	96	0.200	null	null
Total	34.87	99	null	null	null

Table 8: Comparative analysis of the protocol features and capabilities.

Feature/Capability	LEACH	HEED	GEEC	Proposed (DBN-RL-MRFO)
Cluster Head (CH) Selection	Probabilistic	Cost-based	Genetic	Multi-Objective (MRFO)
Uses Machine Learning (ML)	No	No	No	Yes (DBN+RL)
Multi-Objective CH Optimization	No	No	No	Yes
Optimization Criteria	Energy	Energy	Energy	Energy, Delay, Density, Distance
Designed for 5G/IoT Constraints	No	No	Partial	Yes
Adaptive to Network Dynamics	Low	Medium	Medium	High

Table 9: Quantitative performance comparison of the proposed and benchmark protocols (simulation results)

Performance Metric	LEACH	HEED	GEEC	Proposed (DBN-RL-MRFO)	Improvement vs. Best Benchmark
Network Lifetime (FND rounds)	1200	1850	2900	3,250	+12.1% vs. GEEC
Avg. Energy Consumption (J)	0.085	0.072	0.058	0.052	−10.3% vs. GEEC
Throughput (Kbps)	105	125	135	159	+7.4% vs. TTDFP
Average Latency (ms)	12.5	10.8	9.5	7.1	−13.4% vs. TTDFP
Packet Delivery Ratio (%)	96.5%	97.8%	98.2%	99.6%	+1.4% vs. GEEC

SCIFINITI
PUBLISHING

5.4.2. Routing Using Space and Time Complexity Analysis

Figures 20 and 21 present a comparative analysis of time and space complexities, showing that the proposed technique outperforms the comparable methods. The goal of the proposed framework is to provide an effective routing design. We describe both an effective CH selection process and an algorithm for learning-based clustering. These two approaches have increased the combined effectiveness of DBN-based routing. We summarize the proposed framework and provide a conclusion in the next section. Finally, the enhancements and accomplishments of the proposed architecture were examined.

6. Discussion

This study introduces an innovative hybrid DBN-RL-MRFO framework designed to address the significant challenges of energy efficiency, latency, and reliability in 5G-enabled WSNs. The results presented in Section 5 illustrate the superior performance of our approach for all key metrics. This section interprets these findings, discusses their implications within the context of the existing literature, acknowledges the limitations of our work, and suggests avenues for future research.

6.1. Interpretation of Key Findings

The substantial enhancement in network longevity (12.1% improvement over GEEC) and reduction in energy consumption (10.3% decrease compared to GEEC) can be directly ascribed to the synergistic functioning of the three principal components of our framework. The reinforcement learning (RL)-based clustering mechanism dynamically establishes energy-efficient clusters by deriving an optimal policy that maximizes rewards based on residual energy and communication cost, thereby adapting more effectively to network dynamics than static protocols such as LEACH or the one-shot optimization methods employed. This is further augmented by multi-objective cluster head (CH) selection based on Manta Ray Foraging Optimization (MRFO), which optimally balances factors such as energy, distance, delay, and traffic density. In contrast to single-objective methods or fuzzy-based systems, MRFO's robust foraging strategies adeptly navigate the complex solution space to select CHs that minimize overall network energy dissipation and prevent the formation of hotspots. Finally, the Deep Belief Network (DBN)-based routing learns energy-aware paths, further conserving energy by circumventing congested or long-distance routes, offering a significant advantage over traditional routing strategies.
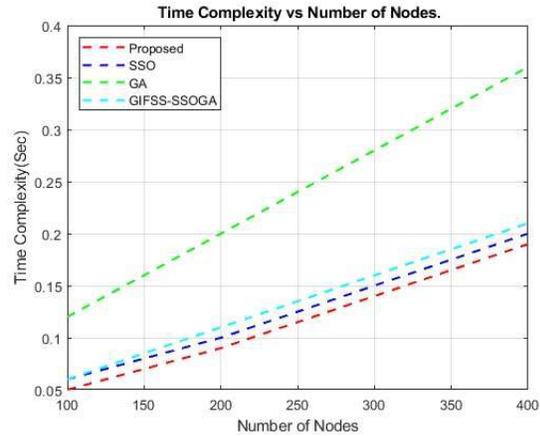

Figure 20: Time complexity of the proposed DBN routing.
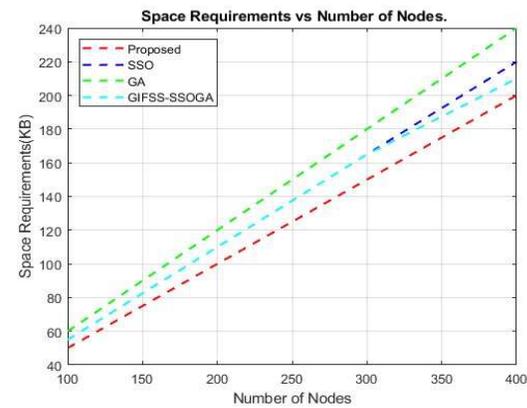

Figure 21: Space demand of the proposed DBN routing method.

The high throughput (159 Kbps) and ultra-reliable packet delivery ratio (PDR) of 99.6% are attributable to the capacity of the deep belief network (DBN) to learn and predict optimal routing paths. By analyzing network state features, the DBN facilitates intelligent forwarding decisions that effectively minimize packet loss, a prevalent issue in conventional protocols such as TTDFP and EADCR. Additionally, the low latency of 7.1 ms satisfies the stringent requirements of 5G ultra-reliable low-latency communication (URLLC), as it was a direct optimization objective within the multi-objective fitness function of the Modified Root-Finding Optimization (MRFO) during cluster head (CH) selection. This multi-objective approach ensures that CHs are not only energy efficient but also centrally located in low-congestion areas, thereby reducing intra-cluster and CH-to-sink communication delays.

6.2. Comparison with Existing Literature

Our findings are consistent with, and significantly extend, the existing body of knowledge. The performance

of RL-based clustering supports these findings; however, our integration of RL specifically for social network-style grouping in the IoT is a novel contribution. The effectiveness of MRFO for CH selection validates the use of bio-inspired algorithms in WSNs; however, unlike previous studies that focused on a limited set of objectives (e.g., primarily energy or distance), our multi-objective formulation provides a more holistic optimization, leading to a more balanced and superior overall performance. Recent studies have increasingly incorporated artificial intelligence into wireless sensor networks (WSNs), exemplified by neuro-fuzzy and secure deep learning models. However, our research distinguishes itself through the comprehensive integration of three distinct AI paradigms: deep learning (DBN), reinforcement learning (RL), and bio-inspired optimization (MRFO). This hybrid architecture transcends the single-objective focus characteristic of deep neural network (DNN)-based methods and the limited adaptability of optimization-only protocols, such as GEEC. Consequently, it offers a more robust and intelligent solution to the complex trade-offs inherent to 5G-IoT networks.

6.3. Limitations and Future Work

Despite these promising outcomes, this study has certain limitations that suggest avenues for future research. First, the simulations were conducted under the assumption of a static network. Future research will evaluate the robustness of the framework in scenarios involving node mobility, a common feature in many IoT applications. Second, the computational overhead associated with training the DBN, although conducted offline, is significant. Exploring lightweight neural network architectures or federated learning techniques for distributed on-device learning could enhance scalability and mitigate central dependencies. Furthermore, although MRFO demonstrated effectiveness, its convergence speed could be optimized for ultra-large-scale networks (e.g., 10,000 nodes). The development of a hybrid or simplified variant of the MRFO to facilitate faster execution is a planned future endeavor. Finally, we intend to implement a hardware testbed utilizing IoT devices and software-defined radios to validate the simulation results in a physical environment, thereby assessing real-world factors such as unpredictable channel interference and packet errors.

7. Conclusions

In conclusion, this study represents a notable advancement in routing for IoT-based wireless sensor networks (WSNs) through the introduction of a hybrid DBN-RL-

MRFO framework. The primary contributions of this research are threefold: (1) the implementation of a reinforcement learning (RL)-based clustering mechanism that adapts to network dynamics; (2) the formulation of cluster head (CH) selection as a multi-objective problem optimizing energy, delay, traffic density, and distance efficiently addressed by the MRFO algorithm; and (3) the development of a deep belief network (DBN)-based routing protocol that learns optimal paths for reliable data transfer. Extensive simulations demonstrate that this approach achieves superior energy efficiency, network longevity, and throughput compared with state-of-the-art protocols. Moreover, the framework demonstrates strong scalability in ultra-dense network scenarios with more than 10,000 nodes, a critical requirement for future 6G infrastructures. To fully realize this potential, future research should focus on optimizing the computational overhead of MRFO for edge servers and exploring federated learning techniques for distributed DBN inference.

List of Abbreviations

5G	Fifth Generation
AI	Artificial Intelligence
BS	Base Station
CH	Cluster Head
CLONALG-M	Clonal Selection Algorithm Modified
D2D	Device-to-Device
DBN	Deep Belief Network
DNN	Deep Neural Network
EADCR	Energy-Aware Distance-based Cluster Head Selection and Routing
GEEC	Genetic-Based Energy Efficient Cluster
IoT	Internet of Things
MDP	Markov Decision Process
ML	Machine Learning
MLP	Multilayer Perceptron
MRFO	Manta Ray Foraging Optimization
MSE	Mean Squared Error
QoS	Quality of Service
RBM	Restricted Boltzmann Machine
RL	Reinforcement Learning
SN	Sensor Node
TTDFP	Two-Tier Distributed Fuzzy Logic-based Prototype
WSN	Wireless Sensor Network

Author Contributions

Simulations, calculations in MATLAB, creation of all figures, system simulations, writing the final manuscript, proofreading, manuscript refinement: V.K.; Supervision,

SCIFINITI
PUBLISHING

Availability of Data and Materials

The MATLAB simulation code and datasets generated and analyzed during the current study are available in the [DBN-RL-MRFO-WSN-MATLAB] repository, [URL: https://github.com/akvijayakumar84/DBN-RL-MRFO-WSN-MATLAB] Access Date: 1 October 2025 onwards.

Consent for Publication

No consent for publication is required, as the manuscript does not involve any individual personal data, images, videos, or other materials that would necessitate consent.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

Acknowledgments

References

[1] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A Survey on 5G Networks for the Internet of Things: Communication Technologies and Challenges," *IEEE Access*, vol. 6, pp. 3619–3647, 2018. [CrossRef]

[2] S. U. Rehman, A. Hussain, F. Hussain, and M. A. Mannan, "A Comprehensive Study: 5G Wireless Networks and Emerging Technologies," in *Proc. 5th Int. Electr. Eng. Conf. (IEEC)*, Karachi, Pakistan, Feb. 2020, pp. 1–6. [Online]. Available: https://www.researchgate.net/publication/342109635.

[3] S. I. Parihar, G. M. Asutkar, and S. Chaturvedi, "Performance Evaluation of Wireless Sensor Network (WSN) in 5G Infrastructure: A Review," *Int. J. Innov. Eng. Sci.*, vol. 4, no. 8, pp. 424–427, 2019. [Online]. Available: https://www.ijies.net/finial-docs/finial-pdf/240519ET5.pdf.

[4] D. Hrabcak, L. Dobos, and J. Papaj, "The Concept of 2-Layer Routing for Wireless 5G Networks and Beyond," in *Proc. 29th Int. Conf. Radioelektronika (RADIOELEKTRONIKA)*, Pardubice, Czech Republic, 2019, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/8733580/.

[5] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi and M. Mustaqim, "Internet of Things (IoT) for Next-Generation Smart Systems: A Review of Current Challenges, Future Trends and Prospects for Emerging 5G-IoT Scenarios," *IEEE Access*, vol. 8, pp. 23022–23040, 2020. [CrossRef]

[6] A. M.-K. Wong, C.-L. Hsu, T.-V. Le, M.-C. Hsieh, and T.-W. Lin, "Three-Factor Fast Authentication Scheme with Time Bound and User Anonymity for Multi-Server E-Health Systems in 5G-Based Wireless Sensor Networks," *Sensors*, vol. 20, no. 9, p. 2511, 2020. [CrossRef]

[7] S. K. Singh and P. Kumar, "A comprehensive survey on trajectory schemes for data collection using mobile elements in WSNs," *J. Ambient Intell. Human Comput.*, vol. 11, pp. 291–312, 2020. [CrossRef]

[8] A. N. Uwaechia and N. M. Mahyuddin, "A Comprehensive Survey on Millimeter Wave Communications for Fifth-Generation Wireless Networks: Feasibility and Challenges," *IEEE Access*, vol. 8, pp. 62367–62414, 2020. [CrossRef]

[9] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2637–2652, 2020. [CrossRef]

[10] B. D. D. and F. Al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks," *Ad Hoc Netw.*, vol. 97, p. 102022, Feb. 2020. [Online]. [CrossRef]

[11] F. Zhou, *Transport protocol design for end-to-end data delivery in emerging wireless networks*. Boston: Northeastern University, 2019. [Online]. Available: http://hdl.handle.net/2047/D20323963.

[12] K. Haseeb, N. Islam, A. Almogren, and I. Ud Din, "Intrusion Prevention Framework for Secure Routing in WSN-Based Mobile Internet of Things," *IEEE Access*, vol. 7, pp. 185496–185505, 2019. [CrossRef]

[13] S. M. M. H. Daneshvar, P. Alikhah Ahari Mohajer, and S. M. Mazinani, "Energy-Efficient Routing in WSN: A Centralized Cluster-Based Approach via Grey Wolf Optimizer," *IEEE Access*, vol. 7, pp. 170019–170031, 2019. [Online]. [CrossRef]

[14] K. Saxena, N. Gupta, J. Gupta, D. K. Sharma, and K. Dev, "Trajectory optimization for the UAV assisted data collection in wireless sensor networks," *Wireless Netw.*, vol. 28, no. 4, pp. 1785–1796, May 2022. [CrossRef]

[15] V. M. Kuthadi, R. Selvaraj, S. Baskar, P. M. Shakeel, and A. Ranjan, "Optimized Energy Management

Model on Data Distributing Framework of Wireless Sensor Network in IoT System," *Wireless Pers. Commun.*, vol. 127, no. 2, pp. 1377–1403, Nov. 2022. [CrossRef]

[16] K. Vijayakumar, P. Thirumaraiselvan, B. Sivakumar, and P. Seenuvasan, "Revolutionizing Data Transmission - A Comprehensive Review of Deep Learning-Based Routing Mechanism for 5G Wireless Sensor Networks," in *Proc. 10th Int. Conf. Commun. Signal Process. (ICCSP)*, Melmaruvathur, India, 2024, pp. 565–570. [Online]. Available: https://ieeexplore.ieee.org/document/10544339/.

[17] A. Habib, M. Y. Arafat, and S. Moh, "Routing Protocols Based on Reinforcement Learning for Wireless Sensor Networks: A Comparative Study,". [Online]. Available: https://www.researchgate.net/publication/331588735.

[18] L. Wen, R. Hailong, and D. Zhongliang, "Future development research of 5G positioning in 3GPP standards," *J. China Univ. Posts Telecommun.*, vol. 1, 2025. [CrossRef]

[19] D. H. Hussein and M. Ibnkahla, "Towards Intelligent Intent-based Network Slicing for IoT Systems: Enabling Technologies, Challenges, and Vision," *IEEE Trans. Netw. Serv. Manage.*, pp. 1, 2025. [CrossRef]

[20] M. Adnan, "Fuzzy Based Secure Clustering Schemes for Wireless Sensor Networks," *arXiv preprint arXiv:2504.17795*, 2025. [Online]. Available: https://arxiv.org/abs/2504.17795.

[21] K. Thangaramya, K. Kulothungan, R. Logambigai, M. Selvi, S. Ganapathy, and A. Kannan, "Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT," *Comput. Netw.*, vol. 151, pp. 211–223, Mar. 2019. [Online]. [CrossRef]

[22] S. Sujanthi and K. Nithya Kalyani, "SecDL: QoS-Aware Secure Deep Learning Approach for Dynamic Cluster-Based Routing in WSN Assisted IoT," *Wireless Pers. Commun.*, vol. 114, no. 3, pp. 2135–2169, Oct. 2020. [CrossRef]

[23] R. Huang, L. Ma, G. Zhai, J. He, X. Chu, and H. Yan, "Resilient Routing Mechanism for Wireless Sensor Networks With Deep Learning Link Reliability Prediction," *IEEE Access*, vol. 8, pp. 64857–64872, 2020. [Online]. [CrossRef]

[24] I. A. Abd El-Moghith and S. M. Darwish, "Blockchain-Based Trusted Routing Scheme for Wireless Sensor Networks," in *Proc. Int. Conf. Adv. Intell. Syst. Informatics*, Cairo, Egypt, 2020, A. E. Hassanien, A. Slowik, V. Snášel, H. El-Deeb, and F. M. Tolba, Eds. Cham: Springer, 2021, pp. 282–291. [CrossRef]

[25] B. M. Sahoo, H. M. Pandey, and T. Amgoth, "GAPSO-H: A hybrid approach towards optimizing the cluster-based routing in wireless sensor network," *Swarm Evol. Comput.*, vol. 60, p. 100772, Feb. 2021. [Online]. [CrossRef]

[26] M. V. Babu, J. A. Alzubi, R. Sekaran, R. Patan, M. Ramachandran, and D. Gupta, "An Improved IDAF-FIT Clustering Based ASLPP-RR Routing with Secure Data Aggregation in Wireless Sensor Network," *Mobile Netw. Appl.*, vol. 26, no. 3, pp. 1059–1067, Jun. 2021. [CrossRef]

[27] V. Srivastava, S. Tripathi, K. Singh, and L. H. Son, "Energy efficient optimized rate based congestion control routing in wireless sensor network," *J. Ambient Intell. Human Comput.*, vol. 11, no. 3, pp. 1325–1338, Mar. 2020. [CrossRef]

[28] S. Bera, S. K. Das, and A. Karati, "Intelligent Routing in Wireless Sensor Network Based on African Buffalo Optimization," in *Nature Inspired Computing for Wireless Sensor Networks*, D. De, A. Mukherjee, S. K. D. Kumar, and N. Dey, Eds. Singapore: Springer, 2020, pp. 119–142. [CrossRef]

[29] J. John and P. Rodrigues, "MOTCO: Multi-objective Taylor Crow Optimization Algorithm for Cluster Head Selection in Energy Aware Wireless Sensor Network," *Mobile Netw. Appl.*, vol. 24, no. 5, pp. 1509–1525, Oct. 2019. [CrossRef]

[30] J. Yuan, J. Peng, Q. Yan, G. He, H. Xiang, and Z. Liu, "Deep Reinforcement Learning-Based Energy Consumption Optimization for Peer-to-Peer (P2P) Communication in Wireless Sensor Networks," *Sensors*, vol. 24, no. 5, p. 1632, Mar. 2024. [Online]. [CrossRef]

[31] K. Ramkumar, N. Ananthi, D. R. D. Brabin, P. Goswami, M. Baskar, K. K. Bhatia; et al., "Efficient routing mechanism for neighbour selection using fuzzy logic in wireless sensor network," *Comput. Electr. Eng.*, vol. 94, p. 107365, Sep. 2021. [Online]. [CrossRef]

[32] S. A. Sert, A. Alchihabi, and A. Yazici, "Two-Tier Distributed Fuzzy Logic Based Protocol for Efficient Data Aggregation in Multihop Wireless Sensor Networks," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3615–3629, Dec. 2018. [Online]. [CrossRef]

[33] Z. Li, M. A. Uusitalo, H. Shariatmadari, and B. Singh, "5G URLLC: Design Challenges and System Concepts," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Lisbon, Portugal, 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8491078/.

[34] J. Zhou, Y. Sun, and C. Tellambura, "Revolutionizing Medical Data Transmission with IoMT: A Comprehensive Survey of Wireless Communication Solutions and Future Directions," *arXiv*, 2025. [Online]. [CrossRef]

[35] A. A. Abba Ari, F. Samafou, A. Ndam Njoya, A. Djedouboum, M. Aboubakar, and A. Mohamadou, "IoT-5G and B5G/6G resource allocation and network slicing orchestration using learning algorithms," *IET Networks*, vol. 14, no. 1, p. e70002, Jan. 2025. [CrossRef]

[36] A. Garg and G. Kaur, "Zone Head Selection Algorithm Based on Fuzzy Logic for Wireless Sensor Networks," *Comput. Electr. Eng.*, vol. 23, no. 10, 2021. [CrossRef]

[37] D. Mehta and S. Saxena, "MCH-EOR: Multi-objective Cluster Head Based Energy-aware Opti-

mized Routing algorithm in Wireless Sensor Networks," *Sustainable Comput. Inform. Syst.*, vol. 28, p. 100406, Dec. 2020. [Online]. [CrossRef]

[38] I. Mustapha, B. Ali, M. Rasid, A. Sali, and H. Mohamad, "An Energy-Efficient Spectrum-Aware Reinforcement Learning-Based Clustering Algorithm for Cognitive Radio Sensor Networks," *Sensors*, vol. 15, no. 8, pp. 19783–197818, Aug. 2015. [Online]. [CrossRef] [PubMed]

[39] S. Soni and M. Shrivastava, "Novel Learning Algorithms for Efficient Mobile Sink Data Collection Using Reinforcement Learning in Wireless Sensor Network," in *Wireless Communications and Mobile Computing*, D. Cassioli, Ed., vol. 2018, p. 7560167, Jan. 2018. [CrossRef]

[40] R. S. Sutton, and A. G. Barto. *Reinforcement Learning, Second Edition: An Introduction*, 2nd ed. Cambridge: Bradford Books. 2018. [Online]. Available: https://mitpress.mit.edu/9780262039246/reinforcement-learning/.

[41] C. P. Verma, "Enhancing Parameters of LEACH Protocol for Efficient Routing in Wireless Sensor Networks," *J. Commun. Mob. Comput.*, vol. 2, no. 1, pp. 26–31, Feb. 2023. [Online]. [CrossRef]

[42] Ö. Yilmaz, K. Kalkan, and F. Alagöz, "SliceScore: A Network Function Sharing Aware and Slice-Oriented DDoS Filtering Approach," *IEEE Access*, vol. 13, pp. 18294–18307, 2025. [CrossRef]