Reinforcement Learning-Driven Cluster Head Selection for Reliable Data Transmission in Dense Wireless Sensor Networks

Longyang Du*, Qingxuan Wang, Zhigang ZHANG
School of Artificial Intelligence, Jiaozuo University, Jiaozuo 454000, Henan, China

***Abstract*—Wireless Sensor Networks (WSNs) have made significant advances towards practical applications. Data gathering in WSNs has been carried out using various techniques, such as multi-path routing, tree topologies, and clustering. Conventional systems lack a reliable and effective mechanism for dealing with end-to-end connection, traffic, and mobility problems. These deficiencies often lead to poor network performance. We propose an Internet of Things (IoT)-integrated densely distributed WSN system. The system utilizes a tree-based clustering approach dependent on the installed sensors' density. The cluster head nodes are structured in a tree-based cluster to optimize the process of gathering data. Each cluster's most efficient aggregation node is selected using a fuzzy inference-based reinforcement learning technique. The decision is based on three crucial factors: algebraic connectedness, bipartivity index, and neighborhood overlap. The proposed method significantly enhances energy efficiency and outperforms existing methods in bit error rate, throughput, packet delivery ratio, and delay.**

Keywords—Energy efficiency; wireless sensor networks; clustering; reinforcement learning; fuzzy inference system

I. INTRODUCTION

A. *Overview*

Wireless Sensor Networks (WSNs) represent a paradigm shift in global technological scenarios and consist of many autonomous sensor nodes capable of carrying out extensive sensing, computation, and communication [1]. Through strategic deployment, WSN nodes reside in a wide range of environments [2]. WSNs constitute a fundamental infrastructure that enables computing systems to gather data, process, and transmit it in real-time from the physical world [3]. This pervasive connectivity opens up applications for numerous fields, such as environmental monitoring, medical care, manufacturing, and sustainable communities [4], [5].

The collaborative nature of sensor nodes within WSNs enables the creation of distributed systems capable of collecting and relaying useful information on environmental parameters, object movement, health indicators, and other relevant data [6]. The inherent characteristic of WSNs, which can be modified to accommodate dynamic and adversarial environments, allows the generation of actionable intelligence, enhancing decision-making processes and improving situational awareness [7], [8].

Like constitutive models, which model the behavior of weak rock masses under different states of stress, taking into account pore pressure and temperature [9], WSNs must combine several environmental parameters to achieve optimal data collection and network operation. Yet, the deployment and operation of WSNs also entail inherent constraints, such as energy limitations [10], scalability issues [11], data security concerns [12], and network reliability [13], which demand innovative solutions and algorithms that ensure optimal performance and overcome these constraints. Despite these concerns, WSNs are a fundamental technology that drives innovation, revolutionizing businesses and enhancing our understanding of the world [14].

B. *Motivation and Contribution*

Several methodologies have been proposed, encompassing diverse techniques such as multi-path routing, tree structures, clustering, and cluster trees, yet they often struggle to ensure a robust and reliable system addressing mobility, traffic dynamics, and end-to-end connectivity individually [15-17]. Consequently, these shortcomings frequently lead to suboptimal network performance, hindering their full potential in practical applications. To solve these challenges, this study introduces a novel scheme tailored to a densely distributed WSN system model.

With a tree-based cluster formation strategy, a flexible deployment density for sensor nodes is accommodated under this innovative framework. Each cluster in this meticulously structured architecture is meticulously organized around a singular cluster head node, a design crafted to streamline and optimize energy-efficient data-gathering processes. A distinguishing element in this scheme is incorporating a fuzzy logic engine and reinforcement learning. This sophisticated system dynamically determines optimal data-gathering nodes within clusters embedded within a densely distributed WSN. This decision-making process requires the evaluation of three key metrics: algebraic connectivity, bipartivity index, and neighborhood overlap.

The proposed approach intelligently assigns data collection tasks, promoting efficiency and energy savings in the network. Like machine learning techniques, such as regression and clustering algorithms, assist in analyzing the effects of various factors on business economics [18], this approach enables effective data-driven decision-making for optimizing resource usage and improving network performance. This study has made the following primary contributions:

- Advanced multi-cluster data collection: We introduce a novel multi-cluster data collection strategy tailored for densely distributed WSNs, addressing the complexities of large-scale monitoring applications.

- Cluster formation based on energy and delay factors: We propose a robust approach to select cluster heads within each cluster, leveraging energy and delay factors to optimize the network's performance for lifespan, throughput, packet delivery rate, and reliable links for mobile sensors.
- Reinforcement Learning-based Fuzzy Logic Engine (RL-FLE): We use incorporated RL-FLE for intelligent decision-making, empowering cluster heads to dynamically determine optimal data-gathering nodes based on link efficiency among neighboring nodes.
- Improved performance indicators: The proposed approach demonstrates superior performance over traditional protocols (LEACH, HEED, MBC) by maximizing link stability and enhancing critical performance indicators, including packet delivery rate, bit error rate, end-to-end delay, and throughput.
- Reduced buffer occupancy and network traffic: Our proposed scheme effectively reduces buffer occupancy and minimizes network traffic, verifying its efficiency in managing data flow and ensuring resource optimization compared to existing protocols.
- Potential for energy savings: Experiments reveal the potential for substantial energy savings, emphasizing the energy-efficient nature of the proposed approach for sustainable and long-term network operation.

The study is structured as follows: Section II reviews related research. Section III presents the methodology, detailing the approach, algorithms, and framework used in the study. Section IV analyzes the findings, comparing them with existing studies and discussing their implications. Finally, Section V summarizes the key insights, highlights the contributions, and suggests potential directions for future research.

II. Related Works

Table I summarizes methodologies, key contributions, evaluation metrics, and results from related works concerning data collection, energy efficiency, and network optimization.

TABLE I. OVERVIEW OF RECENT ROUTING PROTOCOLS

Reference	Methodology	Key contributions	Results
[19]	Secure mobile sensor network with cloud integration	Optimizing performance through efficient routing, energy consumption, and security enhancement. Lightweight and congestion equilibrium-focused data collection scheme. Transmission facilitated via AND-OR graph mechanism. Secure access to collected data for cloud computing.	Significant energy savings and enhanced network stability
[20]	Rechargeable WSNs	Far-relay approach for proportional energy consumption. Optimal scheduling with Opt-JoDGE. Buffer-battery-aware adaptive scheduling with NO-BBA.	NO-BBA closely approaches Opt-JoDGE performance, especially in scenarios with acceptable delay levels.
[21]	Data gathering in WSNs with obstacles	Cluster construction with ant colony optimization and hierarchical aggregation. MS tour formation with multi-agent reinforcement learning and Cluster construction with ant colony optimization and hierarchical aggregation.	DGOB addresses energy consumption and data gathering delay challenges in WSNs with obstacles.
[22]	Trust-aware and energy-efficient data gathering	Clustering, tree construction, and watchdog selection with particle swarm optimization. Variable-length particles for the unknown number of watchdogs.	TEDG algorithm significantly improves energy efficiency and extends network longevity.
[23]	RLSSA-CDG for energy efficiency in WSNs	RLSSA-CDG combines CDG with sleep scheduling in a distributed algorithm. Q-learning algorithm for active node selection.	RLSSA-CDG outperforms other algorithms, demonstrating its energy efficiency and superiority in network lifespan extension.
[24]	Clustering with mobile data collector	Mobile data collector traverses the network for effective data collection.	An optimized approach to mobile data collection, demonstrating effectiveness in both balanced and unbalanced network topologies.
[25]	Energy-aware and cluster-based data aggregation	Fuzzy logic and CapSA for clustering and routing.	CEDAR performs better than prior research in delay, packet delivery rate, and network lifespan.
[26]	Multi-channel design for high throughput in WSNs	Utilizes a subset of cluster heads with multiple radios. Genetic algorithm for clustering, routing, and channel assignment.	It achieves a significant increase in throughput, reduced energy consumption, and improved energy utilization compared to previous schemes.

The increasing utilization of lightweight sensors has driven the advancement of emerging technologies in various domains. One notable trend is integrating cloud services in applications that handle large volumes of observational data. However, the dynamic and time-sensitive nature of these environments requires enhanced performance. Eco-friendly systems require stable and reliable data transmission. Additionally, many green network solutions are vulnerable to unforeseen situations resulting from broadcasting on unreserved mediums.

To meet the mentioned demands, Haseeb, et al. [19], have introduced a secure mobile sensor network that integrates cloud technology to optimize performance through efficient routing, energy consumption, and security enhancement. Their work makes significant contributions in several aspects. Firstly, it focuses on establishing a stable and error-free system for collecting data from mobile sensors to reduce unnecessary energy usage. The proposed data-gathering method is lightweight and maintains congestion equilibrium. An AND-OR graph mechanism facilitates green data transmission from mobile data sources to the cloud, which reduces routing gaps and retransmissions. Cloud computing can provide secure access to the collected data from constraint-oriented green environments.

Liu, et al. [20], studied energy harvesting and joint data gathering challenges in battery-powered WSNs by employing mobile sinks. As a mobile sink travels along a predetermined route, sensor nodes harvest energy from its RF circuits. Meanwhile, these nodes relay sensor data to the sink. A far-relay strategy is proposed to address the proportional relationship between energy consumption and energy harvesting at a sensor node, determined by the distance squared among sensor nodes. This strategy aims to choose sensor nodes near the path to facilitate data transmission to nodes located at a greater distance. The far-relay technique involves formulating a network utility maximization issue and introducing an optimum scheduling strategy considering time slot scheduling, relay selection, and power allocation regulations.

To tackle the issue of effectively managing sensor power and minimizing the delay in capturing data, mainly when obstacles are present, Najjar-Ghabel, et al. [21], have introduced DGOB, which gathers data in WSNs with obstacles. DGOB employs node clustering and a mobile sink to collect cluster heads' data, minimizing network energy usage. The algorithm follows a two-phase process: cluster construction and mobile sink tour formation. DGOB employs two methods to create superior clusters in the cluster construction phase. The first phase involves the combination of hierarchical aggregation with ant colony optimization to produce resilient clusters under adverse conditions. In the subsequent iterations, the Genetic Algorithm (GA) updates the current clusters, thus improving the cluster development process. In the second stage, DGOB presents a proficient approach to constructing tours by combining multi-agent reinforcement learning and GA. The success of DGOB is confirmed by comprehensive simulation findings, demonstrating a 34 per cent reduction in energy usage and an 80 per cent increase in network longevity compared to existing techniques.

Soltani, et al. [22], have presented a trustworthy and energy-aware data aggregation algorithm to enhance data collection efficiency in WSNs. It consists of several vital stages, namely clustering, tree formation, and watchdog determination, each framed as an optimization problem and optimized by the Particle Swarm Optimization (PSO) algorithm. The watchdog selection stage can be particularly noteworthy as it involves particles of variable length due to the unknown number of watchdogs. To address this challenge, a new particle representation and initialization scheme is developed. The proposed algorithm has demonstrated significant improvements in performance metrics through extensive simulations. It significantly improves energy efficiency in delivering data to the sink node, decreases nodes' residual energy standard deviation by 81per cent, and extends the network's lifespan to 129per cent.

Wang, et al. [23], have introduced the Reinforcement Learning-based Sleep Scheduling Approach for Compressed Data Collection (RLSSA-CDC) to enhance energy efficiency in WSNs. This algorithm combines Compressive Data Collection (CDC) with sleep scheduling to reduce data transmission and minimize energy consumption in WSNs. Unlike previous approaches that faced challenges with centralized optimization problems and increased control message exchanges, RLSSA-CDC is formulated as a distributed algorithm. This framework minimizes control message exchanges and adapts to the variance in residual node energy, preventing nodes from premature energy depletion.

Meddah, et al. [24], suggest a novel strategy to mitigate energy waste in WSNs by utilizing Mobile Data Collector (MDC) devices. An MDC device collects data efficiently from sensor nodes by traversing the network. Their proposed method, called the Tree Clustering algorithm with MDC, aims to establish an optimized traveling path through a subset of cluster heads while minimizing the travel distance. The cluster heads are chosen by a competitive selection system that considers several factors, including packet transmission rate, closeness to the root of the tree, node energy level, and proximity to the next cluster head. The efficacy of the suggested approach was evaluated using simulation tests done on both balanced and unbalanced network topologies.

Mohseni, et al. [25], developed a Clustered Energy-conscious Data Aggregation Routing protocol called CEDAR, incorporating a fuzzy logic model and Capuchin Search Algorithm (CapSA). It comprises two steps: cluster creation and extra/intra-cluster routing. Initially, sensor nodes are clustered using fuzzy logic. Then, CapSA determines optimal paths between cluster heads, the base station, and cluster nodes. As demonstrated by the simulations conducted in the MATLAB simulator, CEDAR is superior to existing research concerning packet delivery rate, latency, and network lifetime.

Shahryari, et al. [26], have addressed high-throughput WSN requirements by implementing the multi-channel framework designed explicitly for heterogeneous WSNs. This approach aimed to overcome the limitations present in existing multi-channel methodologies, which often suffer from low throughput and significant overhead. Their innovative solution introduced a paradigm that utilized a subset of high-level nodes, known as cluster heads, with multiple radios within the network. These cluster heads efficiently transfer captured data from standard sensors to the base station. To achieve this, an energy-saving and high-throughput algorithm was developed to manage routing, clustering, and channel assignment processes in this diverse WSN setup.

The first stage focused on forming a spanning tree among the super nodes while intelligently determining appropriate channels for their radios. They introduced a novel multi-objective cost function extending the network lifetime over conventional tree construction methods. Additionally, this function effectively manages interference, improving overall throughput across the network. In the subsequent stage, the algorithm determined the optimal selection of cluster heads and channels for standard nodes. Their algorithm demonstrated a substantial increase in throughput through extensive simulations due to multiple channel utilization. Moreover, it achieved notable reductions in energy consumption per transmitted bit to the base station, achieving an impressive improvement of 21.6 per cent and 48.3 per cent, respectively, compared to prior schemes.

III. Proposed Method

The mobile sink-centric information collection approach prevalent in densely distributed WSNs often consumes more energy by sensor nodes closer to the sink node, leading to energy holes. Existing Expectation Maximization (EM)-based clustering approaches have attempted to mitigate this problem by optimizing the number of clusters to minimize energy consumption. However, these approaches struggle to determine cluster heads effectively, especially as the scale and node density of the network increase, resulting in increased energy consumption and shorter network lifetime. To meet these challenges, this study presents a novel tree-based clustering scheme supported by robust cluster heads to prolong network lifetime and improve energy efficiency.

A single cluster head node controls each cluster to ensure efficient information collection. First, RL-FLE determines strong cluster heads within clusters of densely distributed WSNs. This determination relies on three key variables: neighborhood coverage, mathematical connectivity, and bipartivity index. Then, dynamic network reconfiguration is performed by moving the sink to a different position and consolidating the cluster head node when node failures occur in a cluster. This adaptive framework is expected to prolong network lifespan and minimize energy usage. The effectiveness of the suggested strategy is determined through a comparative analysis of existing methods.

A static and energy-efficient routing scheme is introduced for the complex and diverse IoT ecosystem. The evaluation of this approach involved implementing a transmission algorithm in a network with over a thousand nodes deployed in areas of 200 to 300 square meters, with varying amounts of nodes. The evaluation results underline the suitability and effectiveness of static routing methods for mobile IoT applications. The architecture of the proposed approach follows a layered structure, similar to the traditional layered architecture used in network systems. However, the relay layer is excluded because it is not included in the system. The model represents a hierarchical network structure in which all sensor nodes remain static and stick to static routing-based transmission. Fig. 1 visually depicts a wireless sensor architecture based on a mobile sink. This model uses mobile sinks in the network to acquire information gathered by fixed sensor nodes.

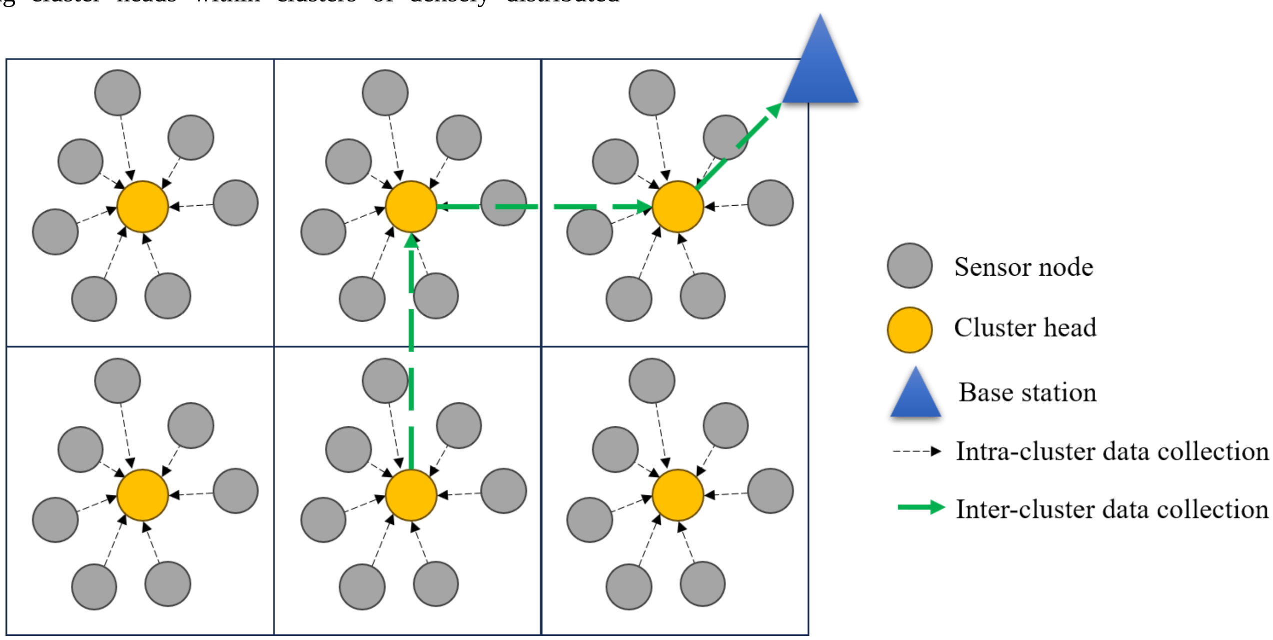

Fig. 1. Network model.

In WSN, the traditional cluster head selection approach considers energy, delay, and distance parameters. However, in the context of IoT networks, it becomes crucial to analyze the specific parameters of IoT devices. Since WSNs are closely linked to IoT devices, it is essential to consider parameters such as temperature and load characteristics of these devices. Therefore, the cluster head selection strategy should consider energy, delay, distance, temperature, and load factors. Ideally, lower temperature, load, delay, and distance values are preferred.

The delay value commonly falls within the range of 0 to 1. Eq. (1) computes the delay sensor nodes encounter when transmitting data to the mobile sink. To decrease this delay, decreasing the number of participants in each cluster is necessary. N denotes the total quantity of sensor nodes, $S(N_v)$ indicates cluster node signal strength, and $S(N_{v'})$ signifies mobile sink node signal strength.

$$D(N_v, N_{v'}) = \frac{\overset{n}{\underset{q=1}{S(N_v) - S(N_{v'})}}}{N} \quad (1)$$

Eq. (2) calculates the distance between the mobile sink and cluster heads. $dist(N_v, N_{v'})$ calculates Euclidean distances between a typical node (N_v) and the mobile sink node ($N_{v'}$) in dense sensor networks.

$$dist(N_v, N_{v'}) = \sqrt{\left(x_{n_v} - x_{n_{v'}}\right) + \left(y_{n_v} - y_{n_{v'}}\right)} \quad (2)$$

To increase the network's longevity, each cluster node's battery level is considered when calculating the remaining power. As packets are forwarded, each node expends energy according to its type, length, frequency, and distance. The power provided by node x_i, denoted as $RP(x_i)$, is determined by the total node count within i^{th} cluster. A higher value of $RP(x_i)$ indicates that a node has more stable and energy-rich power reserves, potentially extending its lifetime and improving network reliability. The residual power of node x_i is calculated as shown in Eq. (3), which helps in identifying stable nodes for long-term cluster membership.

$$RP(x_i) = \frac{\sum_{x_j \in cluster_i} EP_{x_j}}{n_i} \quad (3)$$

Information collection hubs are selected by cluster heads according to three factors: neighborhood overlap (NOVER), Algebraic Connectivity (AC), and Bipartivity Index (BI). NOVER is a quantitative measure to evaluate shared adjacency between the terminal hubs. It is used effectively for group detection, with a lower NOVER value indicating that the connection is likely to connect two different groups, while a higher NOVER value suggests a connection between nodes within the same group. The BI refers to the capacity to partition the vertices of a tree structure into two separate sets so that all edges connect vertices from one set to the other. There are no edges between vertices within the same set. This bipartite property of the graph is an essential consideration in the selection process.

AC is a metric that quantifies a network's resilience regarding link distances. A higher algebraic connectivity value indicates the network is more likely to remain connected even after one or more connection distances, demonstrating its resilience. On the other hand, a lower value suggests that the network could be fragmented by removing links. Fuzzy logic is used to combine these three variables and evaluate the immediate reward of the choice. Fuzzy logic enables the evaluation of connection efficiency from the cluster head node to each neighboring hub.

In addition to the instant reward, the long-term reward of the selection is also considered. The behavior of nearby hubs influences the fairness of selecting data-gathering stations. This is considered by assessing distances between data gathering points and cluster head nodes. The closer the information-gathering hub is to the center, the more favorable it is regarding long-term reward.

By considering these metrics and incorporating fuzzy logic, the cluster head hub can make informed decisions about selecting the information-gathering hub. The evaluation considers both the instant and long-term rewards, ensuring efficient and fair information gathering in the network. The evaluation value for each neighbor is calculated by the cluster head node using the following steps:

- Fuzzification: NOVER, BI, and AC values are converted into fuzzy values by applying predefined membership functions and linguistic terms. These membership functions define membership degrees for fuzzy sets determined by input values. This step allows the crisp values of the metrics to be represented as fuzzy values.

- Defining and applying IF/THEN statements: The fuzzy outcomes derived from the fuzzification process are matched with predetermined IF/THEN criteria. These rules define the relationship between the fuzzy inputs (NOVER, BI, and AC) and the desired output (evaluation value for the neighbor). The rules are designed based on expert knowledge or derived from data analysis. The fuzzy values are aggregated through logical operations (such as AND, OR) embedded inside the IF/THEN rules to determine the ranking of the neighbor.

- Defuzzification: The imprecise numerical value acquired from the preceding phases is transformed into a precise numerical value using a predetermined output membership function and defuzzification process. The output membership function assigns a degree of membership to various numerical values based on the fuzzy value. The defuzzification method calculates a crisp value from the fuzzy output value, typically by taking the centroid or weighted average of the membership function.

The AC value of a system quantifies the network's ability to withstand connection failures. The term refers to the secondary lowest eigenvalue of the Laplacian matrix associated with the system. *Log(A)* is used to measure a system's resilience to failures of connections. A degree vector D_i and adjacency matrix $A(i,j)$ are used to calculate this value. The algebraic connectivity, computed using Eq. (4), measures the resilience of the cluster's internal topology.

$$L(i,j) = \begin{cases} -A(i,j) & for\ i \neq j \\ D_i & for\ i = j \end{cases} \quad (4)$$

The BI is employed to quantify the level of bipartivity in a graph. The range of the value is from 0 to 1. A value of 1 signifies that the graph is bipartite and no frustrated edges connect vertices within the same segment. It will fall below 1 if there are no true bipartite graphs. Eq. (5) provides the bipartivity index, enabling an assessment of structural separation within the communication graph.

$$BPI(G) = \frac{\sum_{j=1}^{n} \cosh(\lambda_j)}{\sum_{j=1}^{n} \sinh(\lambda_j) + \cosh(\lambda_j)} \quad (5)$$

We assess the degree of overlap in neighborhood overlap among the cluster head and neighboring nodes. Due to the difficulty in obtaining an accurate assessment of this overlap in densely deployed WSN settings, we choose immediate neighboring hubs from the cluster head's gauge neighborhood overlap. We define this metric using Eq. (6). NCH (u) refers to the cluster head node and NCN(v) represents its neighbor nodes.

$$NOVER(u-v) = \frac{2 \times |N_{CH}(u) \cap N_{CH}(v)|}{|N_{CH}(u)| + |N_{CH}(v)| - 2} \quad (6)$$

Fig. 2 depicts the fuzzy inference system, with NOVER, BI, and AC as inputs. After performing the fuzzification process, the defuzzification procedure generates an output to determine optimal cluster heads. Input and output membership functions are formulated using a triangular function. Fig. 3 illustrates the fuzzy participation functions for NOVER, BI, and AC. These functions define the degree of belonging to specific linguistic variables (e.g., Bad, Medium, Good) for NOVER, (e.g., Light, Medium, Heavy) for BI, and (e.g., Low, Medium, High) for AC.

Fig. 4 shows the IF/THEN rules the cluster head uses to calculate the rank of participating nodes. These rules map the fuzzy input values (obtained from the participation functions) to the desired output, representing the participating nodes' evaluation rank. Different rules may apply simultaneously, and these rules are combined using the Min-Max strategy. Since multiple rules can apply simultaneously, the evaluation results from different rules are combined using the Min-Max strategy. This strategy selects the minimum (worst) value among the evaluations as the overall evaluation result.

Defuzzification is carried out to produce a precise numerical number that represents the competence value of the node. The output membership function, demonstrated in Fig. 5, assigns degrees of membership to various numerical values according to the fuzzy output value. The Center of Gravity (COG) approach, commonly called the centroid method, is employed for defuzzification. Defuzzified values are represented by x-coordinates, which correspond to a node's competence value.

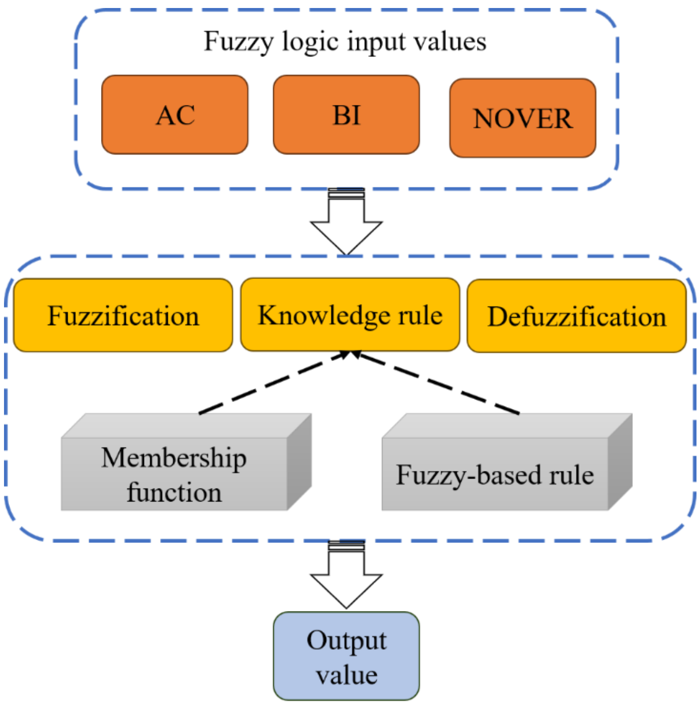

Fig. 2. Fuzzy inference system.

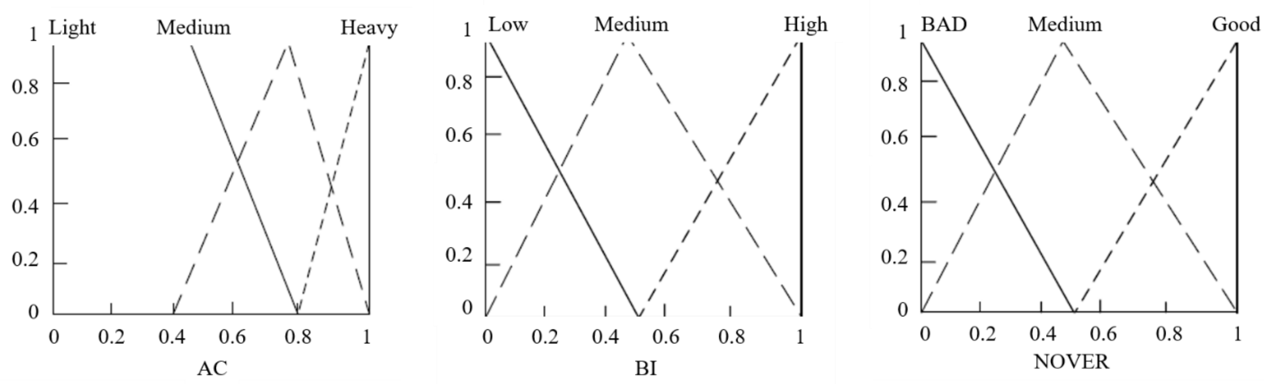

Fig. 3. Fuzzy membership functions for inputs.

1. If (AC is light), (BI is high), and (NOVER is good) then (output is acceptable)
2. If (AC is light), (BI is high), and (NOVER is medium) then (output is unfavourable)
3. If (AC is light), (BI is high), and (NOVER is bad) then (output is bad)
4. If (AC is light), (BI is medium), and (NOVER is good) then (output is unfavourable)
5. If (AC is light), (BI is medium), and (NOVER is medium) then (output is bad)
6. If (AC is light), (BI is medium), and (NOVER is bad) then (output is bad)
7. If (AC is light), (BI is low), and (NOVER is good) then (output is bad)
8. If (AC is light), (BI is low), and (NOVER is medium) then (output is bad)
9. If (AC is light), (BI is low), and (NOVER is bad) then (output is very bad)
10. If (AC is medium), (BI is high), and (NOVER is good) then (output is good)
11. If (AC is medium), (BI is high), and (NOVER is medium) then (output is acceptable)
12. If (AC is medium), (BI is high), and (NOVER is bad) then (output is acceptable)
13. If (AC is medium), (BI is medium), and (NOVER is good) then (output is acceptable)
14. If (AC is medium), (BI is medium), and (NOVER is medium) then (output is bad)
15. If (AC is medium), (BI is medium), and (NOVER is bad) then (output is unfavourable)
16. If (AC is medium), (BI is low), and (NOVER is good) then (output is unfavourable)
17. If (AC is medium), (BI is low), and (NOVER is medium) then (output is unfavourable)
18. If (AC is medium), (BI is low), and (NOVER is bad) then (output is bad)
19. If (AC is heavy), (BI is high), and (NOVER is good) then (output is perfect)
20. If (AC is heavy), (BI is high), and (NOVER is medium) then (output is good)
21. If (AC is heavy), (BI is high), and (NOVER is bad) then (output is acceptable)
22. If (AC is heavy), (BI is medium), and (NOVER is good) then (output is good)
23. If (AC is heavy), (BI is medium), and (NOVER is medium) then (output is acceptable)
24. If (AC is heavy), (BI is medium), and (NOVER is bad) then (output is unfavourable)
25. If (AC is heavy), (BI is low), and (NOVER is good) then (output is good)
26. If (AC is heavy), (BI is low), and (NOVER is medium) then (output is unfavourable)
27. If (AC is heavy), (BI is low), and (NOVER is bad) then (output is bad)

Fig. 4. Fuzzy logic rules.

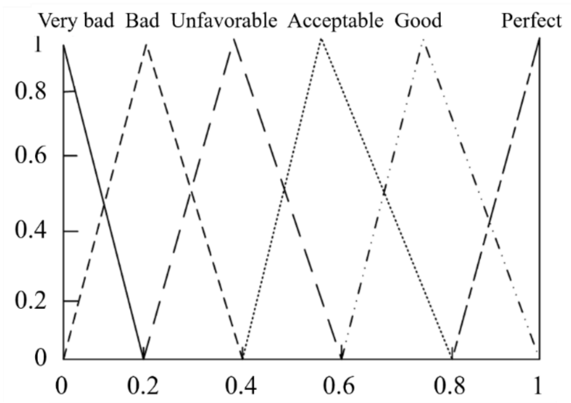

Fig. 5. Fuzzy membership function for the output.

The study presents a fuzzy-based reinforcement learning algorithm to determine the value of action and state relations (*Q(i, a)*) to acquire optimum fuzzy combination rules. It begins with initialization, setting *Q(i, a)* and *V(i, a)* to zero for all states *i* and actions *a*, and introducing control parameters such as *kmax* and *A*. The algorithm progresses through states and actions, selecting actions based on fuzzy combination rules and updating values accordingly.

After each action execution, the Q-value is updated using a reinforcement learning rule considering the observed reward, discount factor, and the maximum Q-value for subsequent actions. The process iterates until it reaches the desired number of iterations (*kmax*). Finally, the algorithm calculates optimal decisions at each state by selecting actions that maximize the learned Q-values. The termination condition marks the conclusion of the algorithm, providing a comprehensive framework for learning optimal fuzzy combination rules through fuzzy-based reinforcement learning.

Fig. 6 depicts the flowchart for the proposed algorithm and provides a visual representation of the steps involved. Fig. 7 provides a pseudocode for the proposed algorithm. To summarize, the algorithm begins by initializing the Q-values and visit counts. It then takes actions based on fuzzy combination rules and updates the Q-values by incorporating rewards and the highest Q-value of the next state. This process is iterated until a specified exit condition is satisfied.

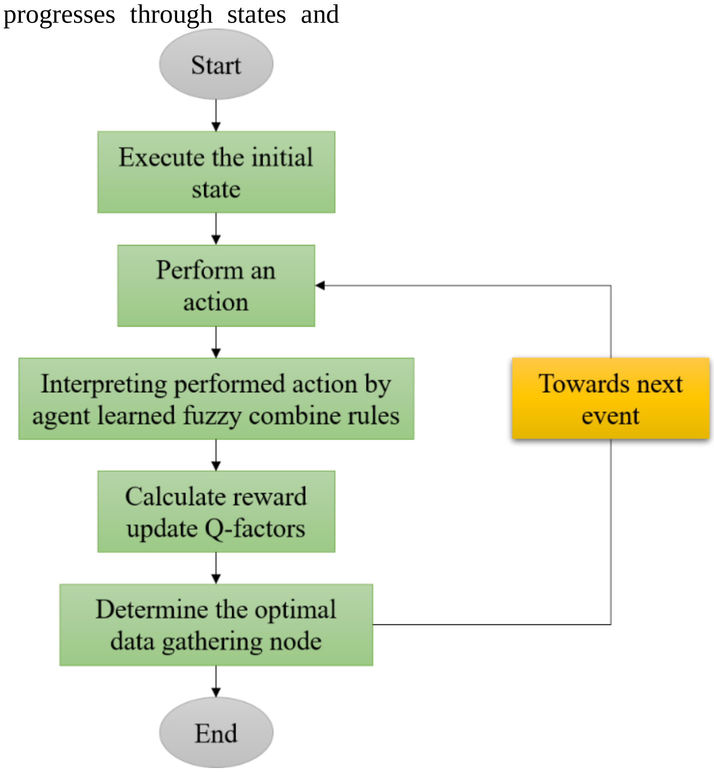

Fig. 6. Flowchart of the proposed algorithm.

Initialization

Set $Q(i, a) = 0$ and $V(i, a) = 0$ for all states i in S and actions a in $A(i)$;

Set $k = 0$, $kmax$, and A as a constant;

Initial state

Calculate the initial state i;

Loop until $k > kmax$

Action selection:

Choose an action a based on the fuzzy combination rules at state i;

Action execution and reward:

Perform action a and observe the resulting reward $r(i, a, j)$;

Update the value $V(i, a)$ by incrementing it by 1: $V(i, a) \leftarrow V(i, a) + 1$;

Calculate the adaptive factor a as A divided by $V(i, a)$: $a = A / V(i, a)$;

Q-value update:

Update the Q-factor associated with the state-action pair (i, a) using the following update rule:

$Q(i, a) \leftarrow (1 - a)Q(i, a) + a[r(i, a, j) + \gamma * max\ b \in A(j)\ Q(j, b)]$,

where γ is the discount factor for future rewards.

Iteration:

Increment the iteration counter: $k = k + 1$;

Set the current state i to the new state j obtained after performing action a;

Optimal decision calculation

Calculate the feasible decisions at each state by selecting the action $a*(i)$ that maximizes the Q-value for state j:

$a*(i) \in arg\ max\ b \in A(j)\ Q(j, b)$

Algorithm termination

Fig. 7. The pseudocode of cluster formation.

IV. Results and Discussion

The proposed method (RL-FLE) was evaluated through simulations performed under different parameter configurations and provided valuable results. To make a comparative analysis with LEACH, CEDAR, PSO, HEED, and MBC, an execution study was carried out using the system test. The simulation scenario encounters 100 identical sensor nodes and 9 cluster heads, all possessing limitless battery capacity, spread across a $1000 \times 1000\ m^2$ region. Furthermore, a fixed sink node with inexhaustible energy stores was strategically placed beyond the surveillance area. The simulations considered specific performance parameters: the MicaZ platform, the ZigBee application with a packet size of 127 bytes, and compliance with IEEE 802.15.4 standards.

The sensor nodes were modeled using a linear battery model with a capacity of 1200 mAh, while the two-ray signal propagation model was employed to capture wireless signal behavior. In the simulation setup, the information envelope within a cluster had a fixed size of 512 bytes. The transmission range within a cluster was limited to 40 m, ranging from 80 m to 120 m between clusters. Notably, the detection range for clustering was set at 20 m. The base station, the central data collection point, was positioned at coordinates (x = 500, y = 1050). Lastly, the energy parameters assigned to each sensing node amounted to 300 mJ.

The performance evaluation of the suggested data-gathering strategy entailed modeling diverse network parameters, including latency, total energy consumption, bit error rate, throughput, and packet delivery rate. Fig. 8 to Fig. 11 illustrate the correlation between network efficiency and node count. It is crucial to emphasize that both MBC and LEACH have restrictions on maximizing throughput, reducing latency, minimizing total energy usage, and maintaining a high packet delivery percentage as the network grows.

Fig. 8 shows the average end-to-end delay across seven clustering protocols under varying node densities. RL-FLE performs better than conventional schemes by minimizing packet delivery latency using intelligent cluster head election through reinforcement learning and link efficiency. In comparison, LEACH and MBC experience higher delays due to their lesser adaptiveness towards dense traffic environments.

Fig. 9 highlights the total energy consumed by sensor nodes during data transmission. RL-FLE indicates the least energy consumption due to its effective data routing via stable and dense cluster heads. The conventional methods, such as HEED and LEACH, cause higher energy consumption through ineffective cluster head rotation or lack of context learning.

Fig. 10 illustrates the network throughput of various protocols. RL-FLE performs better by reducing packet loss and optimizing data flow paths. CEDAR performs similarly, whereas LEACH and HEED perform poorly due to excessive retransmissions and the absence of a dynamic load-balancing mechanism.

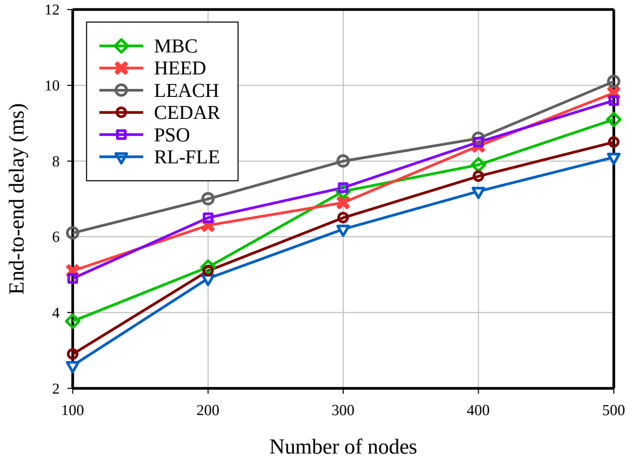

Fig. 8. Network delay comparison.

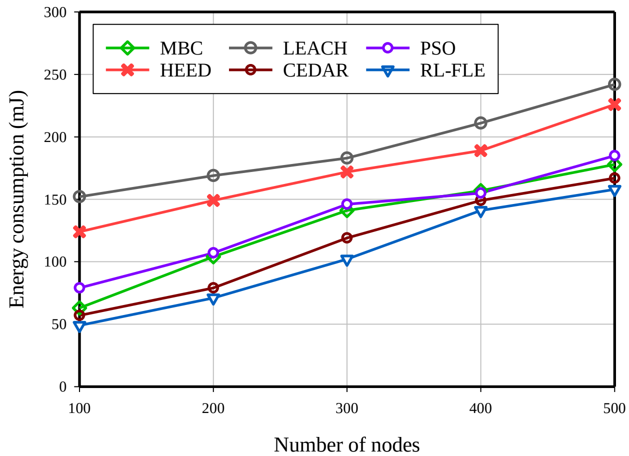

Fig. 9. Energy usage comparison.

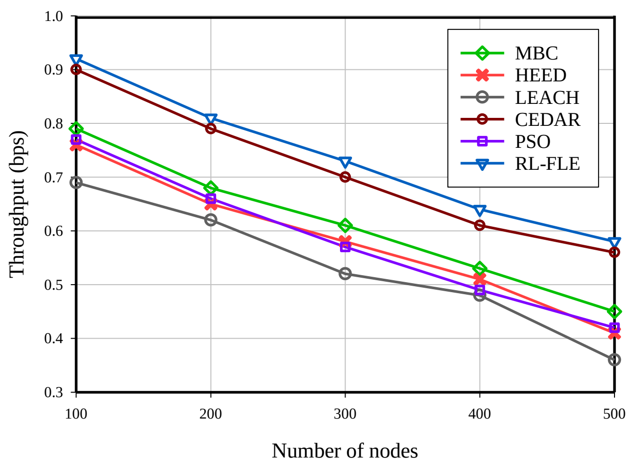

Fig. 10. Throughput comparison.

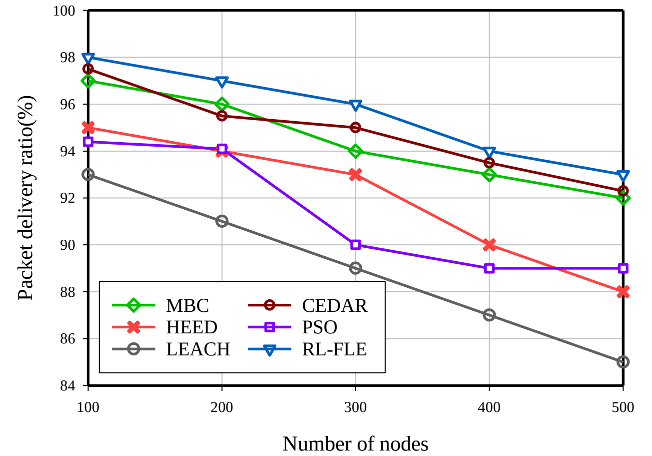

Fig. 11. Performance comparison.

Fig. 11 shows the performance of various protocols under increased node mobility. RL-FLE exhibits low and steady delay despite high-speed movement due to learning-based adaptation. Other protocols experience poor performance due to static cluster head strategies.

The proposed scheme demonstrates superior performance in a mobile sensor environment compared to CEDAR, PSO, LEACH, MBC, and HEED, as indicated in Fig. 12 and Fig. 13. The findings from simulations reveal that the suggested system successfully builds solid links and adjusts to situations with high levels of mobility. Especially in these situations, the recommended strategy results in a better packet delivery ratio and less latency. Furthermore, the strategy improves execution efficiency without regard to the quantity of sensor nodes in the overall setup.

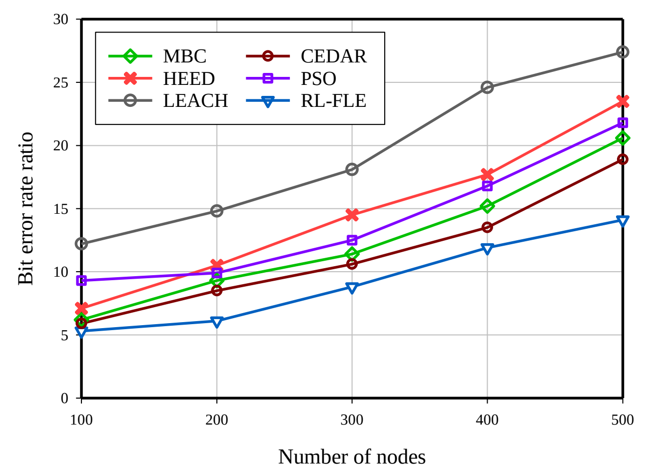

Fig. 12. Bit error rate comparison.

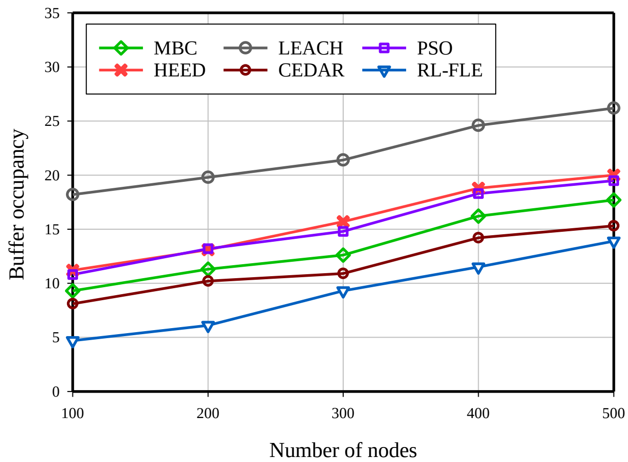

Fig. 13. Throughput comparison.

In WSNs with a significant number of nodes and frequent movement, unstable communication may lead to packet losses, requiring retransmission. In contrast, the proposed scheme ensures stable connections and promotes balanced energy consumption throughout the network. Consequently, it can be concluded that the suggested scheme suits high-mobility environments, enabling the preservation of sensor hub energy, prolonging network lifetime, and enhancing system reliability while maintaining superior communication quality.

V. Conclusion

Effectively monitoring large-scale areas using multiple sensor nodes has become crucial in time-critical military and industrial applications. To address this challenge, the cluster tree network management architecture has emerged as a proficient method. The primary objective is to optimize network performance concerning network lifespan, throughput, packet delivery rate, reliable links for mobile sensors, and energy efficiency. We proposed an effective multi-cluster data collection strategy for WSNs deployed in dense clusters. The energy and delay factors were used to select robust cluster heads for each cluster. Subsequently, the cluster head chooses the data-gathering node based on the link efficiency of neighboring nodes, employing the RL-FLE approach. The proposed scheme was characterized by several advantages, including maximizing link stability and enhancing key performance indicators like packet delivery ratio, delay time, bit error rate, and throughput.

Experimental outcomes suggest our approach effectively reduces buffer occupancy and network traffic while minimizing energy utilization compared to CEDAR, PSO, LEACH, HEED, and MBC protocols. Potential areas for future study might focus on improving energy efficiency, scalability, and flexibility in dynamic contexts. Exploring the incorporation of new technologies like edge computing and machine learning algorithms can enhance the effectiveness of multi-cluster network management architectures for large-scale and time-sensitive applications. This will help to advance the development of WSNs in a rapidly changing environment.

References

[1] B. Pourghebleh, N. Hekmati, Z. Davoudnia, and M. Sadeghi, "A roadmap towards energy - efficient data fusion methods in the Internet of Things," Concurrency and Computation: Practice and Experience, vol. 34, no. 15, p. e6959, 2022.

[2] H. Sallay, "Designing an Adaptive Effective Intrusion Detection System for Smart Home IoT," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 15, no. 1, 2024.

[3] H. Xue, Z. Zhang, and Y. Zhang, "A novel cluster-based routing protocol for WSN-enabled IoT using water-cycle algorithm," Proceedings of the Indian National Science Academy, vol. 89, no. 3, pp. 724-730, 2023.

[4] B. Pourghebleh and V. Hayyolalam, "A comprehensive and systematic review of the load balancing mechanisms in the Internet of Things," Cluster Computing, pp. 1-21, 2019.

[5] A. Janarthanan and V. Vidhusha, "Cycle - Consistent Generative Adversarial Network and Crypto Hash Signature Token - based Block chain Technology for Data Aggregation with Secured Routing in Wireless Sensor Networks," International Journal of Communication Systems, p. e5675, 2023.

[6] O. S. Egwuche, A. Singh, A. E. Ezugwu, J. Greeff, M. O. Olusanya, and L. Abualigah, "Machine learning for coverage optimization in wireless sensor networks: a comprehensive review," Annals of Operations Research, pp. 1-67, 2023.

[7] H. Chaitra et al., "Delay optimization and energy balancing algorithm for improving network lifetime in fixed wireless sensor networks," Physical Communication, vol. 58, p. 102038, 2023.

[8] H. I. Obakhena, A. L. Imoize, F. I. Anyasi, and K. Kavitha, "Application of cell-free massive MIMO in 5G and beyond 5G wireless networks: A survey," Journal of Engineering and Applied Science, vol. 68, no. 1, pp. 1-41, 2021.

[9] A. Azadi and M. Momayez, "Simulating a Weak Rock Mass by a Constitutive Model," Mining, vol. 5, no. 2, p. 23, 2025.

[10] D. Lin, Z. Chen, X. Liu, L. Kong, and Y. L. Guan, "ESWCM: A Novel Energy-sustainable Approach for SWIPT-enabled WSN with Constraint MEAP Configurations," IEEE Transactions on Mobile Computing, 2024.

[11] S. K. Gupta and S. Singh, "Energy Efficient Dynamic Sink Multi Level Heterogeneous Extended Distributed Clustering Routing for Scalable WSN: ML-HEDEEC," Wireless Personal Communications, vol. 128, no. 1, pp. 559-585, 2023.

[12] V. Hayyolalam, B. Pourghebleh, and A. A. Pourhaji Kazem, "Trust management of services (TMoS): investigating the current mechanisms," Transactions on Emerging Telecommunications Technologies, vol. 31, no. 10, p. e4063, 2020.

[13] Y. Y. Ghadi et al., "Machine learning solutions for the security of wireless sensor networks: A review," IEEE Access, vol. 12, pp. 12699-12719, 2024.

[14] S. Kaveripakam and R. Chinthaginjala, "Energy balanced reliable and effective clustering for underwater wireless sensor networks," Alexandria Engineering Journal, vol. 77, pp. 41-62, 2023.

[15] T. M. Ghazal, M. K. Hasan, H. M. Alzoubi, M. Alshurideh, M. Ahmad, and S. S. Akbar, "Internet of Things Connected Wireless Sensor Networks for Smart Cities," in The Effect of Information Technology on Business and Marketing Intelligence Systems: Springer, 2023, pp. 1953-1968.

[16] R. Priyadarshi, "Energy-Efficient Routing in Wireless Sensor Networks: A Meta-heuristic and Artificial Intelligence-based Approach: A Comprehensive Review," Archives of Computational Methods in Engineering, pp. 1-29, 2024.

[17] P. Thippun, Y. Sasiwat, D. Buranapanichkit, A. Booranawong, N. Jindapetch, and H. Saito, "Implementation and experimental evaluation of dynamic capabilities in wireless body area networks: different setting parameters and environments," Journal of Engineering and Applied Science, vol. 70, no. 1, p. 1, 2023.

[18] M. B. Bagherabad, E. Rivandi, and M. J. Mehr, "Machine Learning for Analyzing Effects of Various Factors on Business Economic," Authorea Preprints, 2025, doi: 10.36227/techrxiv.174429010.09842200/v1.

[19] K. Haseeb, Z. Jan, F. A. Alzahrani, and G. Jeon, "A secure mobile wireless sensor networks based protocol for smart data gathering with cloud," Computers & Electrical Engineering, vol. 97, p. 107584, 2022.

[20] Y. Liu, K.-Y. Lam, S. Han, and Q. Chen, "Mobile data gathering and energy harvesting in rechargeable wireless sensor networks," Information Sciences, vol. 482, pp. 189-209, 2019.

[21] S. Najjar-Ghabel, L. Farzinvash, and S. N. Razavi, "Mobile sink-based data gathering in wireless sensor networks with obstacles using artificial intelligence algorithms," Ad Hoc Networks, vol. 106, p. 102243, 2020.

[22] K. Soltani, L. Farzinvash, and M. A. Balafar, "Trust-aware and energy-efficient data gathering in wireless sensor networks using PSO," Soft Computing, pp. 1-24, 2023.

[23] X. Wang, H. Chen, and S. Li, "A reinforcement learning-based sleep scheduling algorithm for compressive data gathering in wireless sensor networks," EURASIP Journal on Wireless Communications and Networking, vol. 2023, no. 1, p. 28, 2023.

[24] M. Meddah, R. Haddad, and T. Ezzedine, "An efficient mobile data gathering method with tree clustering algorithm in wireless sensor networks balanced and unbalanced topologies," Wireless Personal Communications, pp. 1-19, 2022.

[25] M. Mohseni, F. Amirghafouri, and B. Pourghebleh, "CEDAR: A cluster-based energy-aware data aggregation routing protocol in the internet of things using capuchin search algorithm and fuzzy logic," Peer-to-Peer Networking and Applications, pp. 1-21, 2022.

[26] M.-S. Shahryari, L. Farzinvash, M.-R. Feizi-Derakhshi, and A. Taherkordi, "High-throughput and energy-efficient data gathering in heterogeneous multi-channel wireless sensor networks using genetic algorithm," Ad Hoc Networks, vol. 139, p. 103041, 2023.