Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Trust-Aware Hybrid Q-AHA Energy-Efficient secure Clustering for Wireless Sensor Networks with a Mobile Sink Node

Keerthi K S

keerthi.k.s@gmail.com

SJB Institute of Technology (SJBIT) , affiliated to Visvesvaraya Technological University (VTU) Belagavi

Jagadisha N

SJB Institute of Technology (SJBIT) , affiliated to Visvesvaraya Technological University (VTU) Belagavi

Research Article

Keywords: Wireless Sensor Networks, Mobile Sink, Secure Clustering, Q-Learning, Artificial Hummingbird Algorithm, Trust Management

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11385271/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Trust-Aware Hybrid Q-AHA Energy-Efficient secure Clustering for Wireless Sensor Networks with a Mobile Sink Node

Keerthi K.S[1*] and Jagadisha N[2]

[1,2]SJB Institute of Technology (SJBIT), Bengaluru, affiliated to

Visvesvaraya Technological University (VTU) Belagavi, Karnataka,India

*Corresponding author. E-mail: keerthi.k.s@gmail.com;

Contributing authors: jagadisha.n83@gmail.com

Abstract

Mobile Wireless Sensor Networks (MWSNs) have important function in dynamic data collecting, but suffer from severe limitations, such as limited energy, topology instability, and susceptibility to malicious routing attacks. Most traditional clustering approaches regard energy efficiency and security as independent objectives, which may cause high energy consumption and low data dependability. In this study, to overcome these challenges, we propose a Trust-Aware Unified Hybrid Q-Learning and Artificial Hummingbird Algorithm (Q-AHA) clustering algorithm for MWSNs with mobile sink. At the local level, distributed Q-learning agents evaluate the quality of sensor nodes based on their residual energy, mobility, distance from the sink, and a dynamic Beta-reputation-based trust metric. At the global level, the Artificial Hummingbird Algorithm (AHA) is used in the mobile sink to optimize the CH selection based on the learnt node quality scores. Furthermore, an event-triggered control technique is proposed to avoid superfluous control packet broadcasts. Under the same simulation environment, the performance of Q-AHA is tested and compared with RL-PSO, LEACH, LEACH-RLC and M-TAAHA with node numbers 50, 100 and 200 with 10% malicious-node condition. It achieved LND values of 3720, 3723, and 3727 rounds, PDR values of 93.55%, 94.86%, and 95.44%, throughputs of 13.2045, 23.9824, and 45.7374 kbps, and delays of 9.55, 10.1543, and 11.4244 ms, respectively. Q-AHA has good trade-offs between network lifetime, reliability, throughput and delay, showing that it is suitable for application to scalable, energy-efficient MWSNs.

Keywords: *Wireless Sensor Networks, Mobile Sink, Secure Clustering, Q-Learning, Artificial Hummingbird Algorithm, Trust Management.*

1. Introduction

WSNs are the back bone of a wide range of current day applications including environmental monitoring, industrial automation, smart city infrastructure and military surveillance. These networks are built of densely deployed resource constrained sensor nodes that continuously sense, process and transmit the environmental data to a central Base Station (BS) [1]. However, limited energy remains one of the main challenges in WSNs. Sensor nodes have restricted battery power, which can affect the network's performance and lifetime. Sensor nodes are generally deployed in harsh or remote environments where battery replacement is physically or economically infeasible. A key priority in WSN research has been energy-efficient routing and clustering topologies to increase the network lifetime [2].

To address the “energy hole” or “hotspot” problem [3] [4], i.e. the rapid depletion of energy of nodes close to a static base station, Mobile Wireless Sensor Networks (MWSNs) Mobile sink-based methods have been proposed to improve energy efficiency. A mobile sink can help distribute energy consumption more evenly across the network[5], but it causes major challenges of dynamic topology [6] . Moreover, the expanding use of WSNs in unsupervised and decentralized systems renders them more prone to security assaults. Malicious or compromised nodes can perform routing attacks (e.g., blackhole or grayhole (selective forwarding) assaults) [7] by silently dropping packets and ruining the Packet Delivery Ratio (PDR) and the overall Quality of Service (QoS) of the network.

In many conventional clustering protocols such as LEACH [8] and its variations, energy-efficiency and network security are frequently considered as incompatible objectives. Most of the protocols that are heavily optimized for energy balancing assume a fully cooperative network and so are naive to insider attacks. However, using complex cryptographic methods can add extra processing and communication overhead, which may increase energy use in resource-constrained sensor nodes [9] . More recently Reinforcement Learning (RL) [10] and in particular Q-Learning [11] has become a strong way to allow sensor nodes to adapt dynamically to their environment. However, pure multi-agent Q-learning directly applied in a mobile large-scale WSN usually suffers from “curse of dimensionality”. The long convergence times and hefty memory requirements are beyond the capability of the sensor nodes. Swarm intelligence algorithms [12] such as metaheuristic [13] algorithms like the Artificial Hummingbird Algorithm (AHA) are good at finding globally optimal cluster formations but do not have the ability of constant “learning” based on localized prior interactions.

To overcome this gap, this work proposes a Trust-Aware Unified Hybrid Q-Learning and Artificial Hummingbird Algorithm [14] Clustering Protocol to handle energy efficiency and secure routing problem at the same time. The proposed design is built on a two-tier framework. At local level, distributed Q-learning agents continually examine their local state such as residual energy, mobility, distance to sink and a dynamic Beta-reputation trust metric to offer a node quality score. The mobile sink use Q-AHA Operational Framework to process these learning scores globally, thereby overcoming the RL centralization bottleneck to select a highly secure globally optimal Cluster Head (CH) set.

The primary contributions of this work are:

1. We propose a novel dual-tier clustering protocol based on local Q-learning. It employs Artificial Hummingbird Algorithm (AHA) for global multi-objective optimization.

2. The Q-learning reward system adopts a Beta-reputation trust model to evaluate the reliability of nodes in the selection of cluster heads. This strategy is beneficial for detection and elimination of rogue or unstable nodes thereby improving network security and data integrity.

3. An event-triggered state update mechanism is proposed to cut down the communication cost usually brought by the centralized metaheuristics, significantly reducing the control packet transmissions while keeping the clustering accuracy [15] .

The remainder of this paper is organized as follows. Related work on secure clustering and RL-based WSN routing is presented in Section 2. 3. Models of Systems This section describes the system models such the energy, mobility and threat models. Section 4 covers the proposed Trust-Aware Q-AHA protocol simulation results and the conclusion in Section 5.

2. Related Works

In this section, present research works related to clustering and routing in WSNs are described in depth along with methodology and key restrictions. Fofana and Lei [16] A reinforcement learning-based clustering method has been proposed that combines K-Means clustering with the SAC-D algorithm for cluster head selection in WSNs. It uses residual energy, distance, and RSSI information to improve energy efficiency and extend the network lifetime. However, it does not consider trust management, node mobility and security against rogue nodes, so there is need for further improvement. Kumar et al. introduced an ANN-based clustering protocol for WSNs, EECA [17] , which improves energy efficiency through intelligent cluster head selection and eliminates redundant data transmission. The protocol enhances the network lifetime with the help of cluster size control and verification times. However, it does not take into account the trust-aware security, mobility support and meta-heuristic optimization for cluster head selection. Recently, Bourebia and Li [18] proposed a system called GEECS, which enhances energy efficiency in CH selection by using scalable K-means++ and ϵ-greedy multi-armed bandit. The protocol is very efficient to increase the lifetime of the network but it presupposes that the sensor nodes remain completely static. It has no security foundation, hence it is highly vulnerable to malicious routing attacks in mobile scenarios. Singh and Vir [19] proposed a hierarchical Routing Rings technique to address the mobile sinks which enhances the speed of data transmission and decreases the static hotspot problem. The ring architecture is efficient in reducing the latency for multicast but its scalability is restricted by the continuous energy consumption for frequent node-to-node ring communication. Moreover, the framework is quite susceptible to localization difficulties and provides no security notion at all. Jukuntla and Dondeti [20] proposed a mobile-sink data collection method based on LEACH and rendezvous nodes to save energy and improve network lifetime. However, their approach does not combine reinforcement learning, dynamic trust assessment, and AHA-based optimization for secure cluster-head selection. Sangeetha et al. [21] proposed a mobile-sink path-planning method that uses clustering and multi-objective optimization to improve energy-efficient data collection. However, the approach does not include trust-based security or reinforcement learning for cluster-head selection. Choudhary et al.

Sulakshana and Kamatam [22] proposed a mobile-sink data collection method that plans the order in which sensor nodes are visited to make data gathering more efficient. However, the approach does not consider trust-based security or reinforcement learning for cluster-head selection. [23] proposed a hybrid RL-PSO system based on residual energy metrics for the optimization of cluster head selection and physical failure detection. The approach is very successful to maximize stationary network lifetimes but it does not incorporate node mobility issues or dynamic trust management that restrict its utility against hostile nodes in mobile WSNs. Chiraira et al. [24] presented a spatiotemporal cluster-based MARL framework, where the Q-

learning tasks are offloaded to the local cluster agents, which successfully reduces the state-space complexity and the total energy consumption of large-scale WSNs. Although computationally efficient, the system is built on static infrastructure and does not evaluate node trust, thereby constraining its application in networks with mobile sink routing and hostile insider attacks. Mouli et al. [25] proposed a dual reinforcement learning approach which uses LO-SARSA for CH selection and a neural-network based PO-PGM to improve next-hop routing. This methodology leads to large gains in energy efficiency for cooperative nodes, but does not account node trustworthiness or dynamic routing concerns of a mobile sink. Jurado-Lasso et al. [26] suggested a centralized protocol, named LEACH-RLC, which applies Mixed Integer Linear Programming (MILP) for the CH selection and a Reinforcement Learning (RL) agent to regulate the cluster regeneration time. This approach is good to reduce the control overhead and enhance the energy distribution of the network by dynamically learning the ideal moment to trigger new clustering solutions. However, its scalability to large deployments is constrained as it relies on a central controller with a full global view of the network and assumes all sensor nodes to be fully static. ReLeC protocol is introduced by Sharma et al. [27] to solve the energy optimization problem by intelligent routing. The ReLeC protocol is composed of two phases: a cluster creation phase and a Q-learning data transfer phase. This approach can decrease the delivery latency and enhance the network lifetime by rewarding the paths with high residual energy and short hop counts, but it does not take into account the security and dynamic characteristics of the working environment. The absence of trust evaluation method and mobile sink support restricts its scalability and usability in dynamic hostile deployments where malicious routing interference is a problem. Choudhary and Barwar [28] , A clustering approach combining Reinforcement Learning (RL) and Particle Swarm Optimization (PSO) has been proposed to address the energy limitations of WSNs. The method selects cluster heads by considering both energy efficiency and network connectivity, which helps reduce transmission cost and extend network lifetime. However, it mainly focuses on energy optimization in static networks and does not include a strong trust mechanism to protect the clustering process from malicious nodes in dynamic environments. Jayabalan E and Pugazendi [29] suggested a Q-learning based clustering technique to determine optimal routing paths and extend the lifetime of the network in Wireless Body Sensor Networks. The protocol efficiently routes vital healthcare data, but it does not have a trust management architecture at all, which severely restricts its resilience to internal routing assaults or compromised nodes.

Table 1. Contributions and limitations of the literatures

Reference No.	Contribution	Limitation	Used Trust?	Considered Mobile Sink?
Fofana & Lei [16]	Adaptive and energy-efficient CH selection for enhancing network lifetime using proposed K-Means with SAC-D	The work focuses on a simpler two-hop communication model and does not address complex multi-hop routing	No	No
Debasis et al. [17]	ANN-based CH selection using energy, events, distance, and	Single-hop communication only; nodes failing to join	No	No

	neighbors, with verification-based idle-listening reduction	clusters communicate directly with BS (energy exhaustive);		
Bourebia & Li [18]	Proposed GEECS using ε-greedy algorithm for CH selection with scalable K-means++, achieved better network lifetime for 100 nodes	Packet delivery ratio decreases with network growth, clustering performed only once, lacks adaptability to network dynamics	No	No
Singh & Vir [19]	The study proposes an improved Routing Rings approach that uses a mobile sink and intelligent routing to increase data delivery, reduce delay and control overhead, and use network energy more efficiently.	The method depends on accurate sink location information, and maintaining the routing ring can consume extra energy. Its performance may also vary with sink movement, network structure	No	Yes
Sangeetha et al.[21]	Uses clustering and multi-objective optimization to plan the mobile sink's path while reducing energy consumption.	Does not consider trust-based security or reinforcement learning for selecting cluster heads	No	yes
Chiraira et al. [24]	It uses spatiotemporal clustering with MARL/Q-learning to group sensor nodes and make decisions locally. This helps the network scale better, save energy, and reduce communication and computation costs	The study also leaves issues like different sensor types, node failures, network changes, and very large networks for future work.	No	No
Mouli et al. [25]	Uses LO-SARSA, Q-learning, and PO-PGM to choose cluster heads, organize sensor nodes, and find better routing paths.This helps save energy, deliver data more reliably, and keep the network running longer	The method mainly tries to save energy, reduce delays, and improve data delivery, but it does not address security issues like detecting untrusted or malicious sensor nodes.	No	No
Jurado-Lasso et al. [26]	Uses MILP and DQN to choose energy-efficient cluster heads and decide when clusters should be updated.This helps reduce extra control messages, save energy, and keep the network running longer	The clustering process needs a lot of computation, which can make real-time use difficult. It also assumes that the sensor nodes and the base station stay in fixed locations.	No	No
Sharma et al. [27]	ReLeC uses clustering and reinforcement learning to select suitable cluster heads and efficient routes based on remaining energy and hop count.This helps the sensor network use less energy and operate for a longer time.	Real-world testing needed; computational viability requires further analysis; limited scalability testing	No	No

Choudhary et al.[28]	Uses RL and PSO together to select efficient cluster heads, detect faults, and help the network last longer	The method mainly focuses on energy, connectivity, and distance, without considering trust-based security or mobile-sink operation.	No	No
Jayabalan & Pugazendi [29]	Uses clustering and Q-learning to choose suitable cluster heads and find efficient routes for sending health data.This helps save energy, reduce delay, and improve packet delivery and network performance	Choosing the best route with Q-learning add computation time, which can make the routing process slower.	No	No
Anil Kumar Jemla Naik et al.[30]	Proposes trust aware multiobjective AHA for secure cluster-head selection and multihop routing using residual energy, trust, distance and cluster balancing with a mobile sink.	Assumes static sensor nodes, no Q-learning based node evaluation, event triggered state updates	Yes	yes

2.1. Research Gap

Although much progress has been made in energy-efficient clustering for Wireless Sensor Networks (WSNs) based on machine learning and metaheuristic algorithms, most of the available protocols presume static and completely cooperative environments. However, the available hybrid models are successful in optimizing energy consumption but lack the vital dynamic trust-aware security features and are highly vulnerable to internal routing assaults as well as compromised nodes. Moreover, these frameworks always not well address the complicated topological problems and ongoing routing overhead generated by movable sink nodes in large scale deployments. Thus, there is a significant research gap in a holistic framework which integrates Reinforcement Learning (Q-learning) with advanced swarm intelligence (e.g. Artificial Hummingbird Algorithm) to optimize energy efficiency, routing that adjusts automatically as the mobile sink moves through the network. and continuously checks node trust to keep the network more secure.

3 Methodology

In this section, we describe the architecture and operational model of the proposed Trust-Aware Unified Hybrid Q-Learning and Artificial Hummingbird Algorithm (Q-AHA) Clustering Protocol. The protocol is designed for routing security and energy consumption optimization in Wireless Sensor Networks (WSNs) with a mobile sink. This architecture combines the local decision-making ability of the Reinforcement Learning (Q-learning) in the optimal evaluation of the Cluster Head (CH) with the global optimization ability of the Artificial Hummingbird Algorithm (AHA) that significantly alleviates the high energy consumption cost of the dynamic topologies. The protocol importantly includes a robust β-reputation trust model in the clustering logic to secure against rogue nodes.

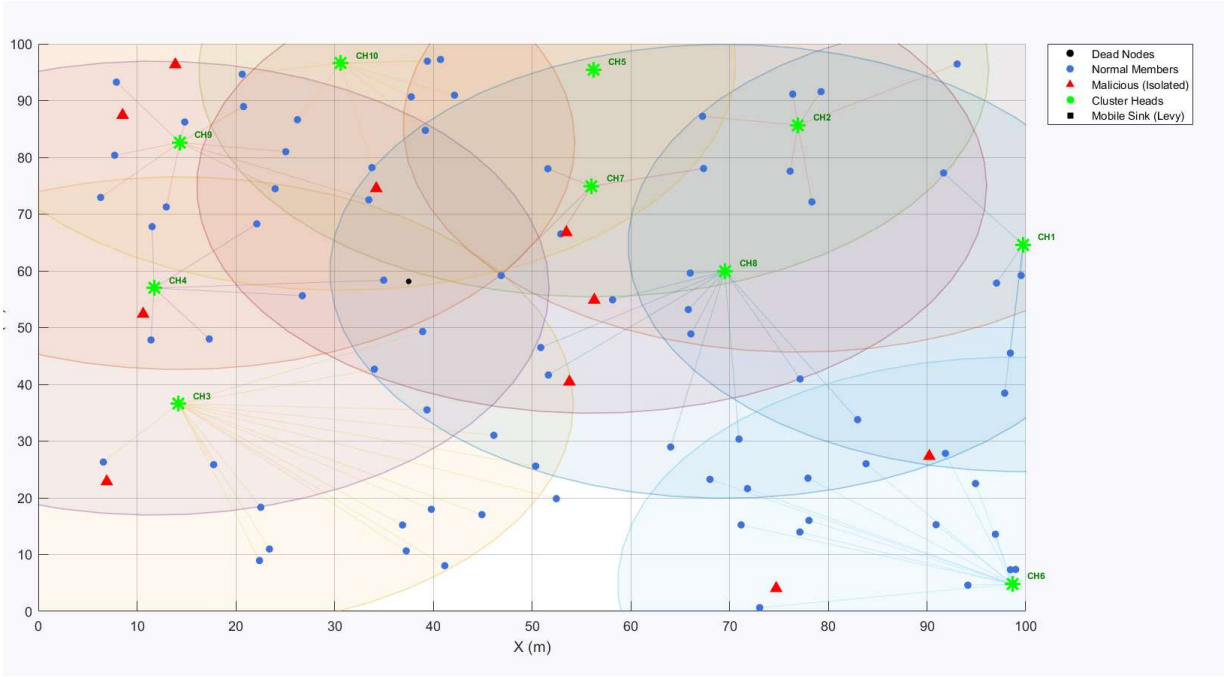


Figure 1. Adaptive Cluster Formation and Node Allocation

3.1 Network and Mobility Model

The network comprises sensor nodes scattered across a target geographical area, alongside a mobile sink that starts at an initial position, $P_{sink}(0)$. The sensor nodes and the mobile sink follow the Random Waypoint Mobility Model (RWMM) to describe the dynamic aspect of the network. In each simulation round, the position of each sensor node and the mobile sink is updated based on their selected random way point and mobility parameters. The position of sensor node i at time t is expressed as $P_i(t)=(x_i(t),y_i(t))$ while the position of mobile sink is expressed as $P_{sink}(t)=(x_{sink}(t),y_{sink}(t))$.
The Euclidean distance is used to determine the instantaneous distance between sensor node and mobile sink. It is denoted as $D_{BS}(i,t)$ as:

$$D_{BS}(i,t) = \sqrt{\left(x_i(t) - x_{sink}(t)\right)^2 + \left(y_i(t) - y_{sink}(t)\right)^2} \quad (1)$$

where $D_{BS}(i,t)$ is the current distance of node to sink at round . Due to the mobility of the sensor nodes and the more evenly across the network, the coordinates of them and therefore $D_{BS}(i,t)$ are updated dynamically in each round of simulation. Then, this distance is utilized in the energy consumption, communication range, link-quality, routing and cluster-head selection procedures of the proposed network model.
The proposed framework employs Random Waypoint Mobility (RWP) [31] model to represent the mobility of sensor nodes. Each sensor node randomly chooses a location in the sensing field and moves with a speed uniformly distributed in the range of 0.5-5 m/s. When it reaches the destination, the node waits for 5 s and then chooses the next destination. The mobile sink moves across the planned sojourn areas while the mobile sensor nodes move independently according to the RWP model. The mobility model is a realistic topology change and is extensively used to evaluate mobile wireless sensor networks.

3.2 Mobility Energy Model

In the proposed topology-aware Q-AHA protocol, each node is considered mobile. Consequently, the energy dissipation is dynamic, and the Residual Energy (E_r) is updated continuously at every time step t by integrating communication costs, overheads, and physical mobility costs. The energy consumption for transmitting a message of length m bits over distance $d(t)$ at time t is modeled as,

$$E_{TX}\big(l, d(t)\big) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d(t)^2, & d(t) < d_0 \\ lE_{elec} + l\varepsilon_{mp}d(t)^4, & d(t) \geq d_0 \end{cases} \quad (2)$$

Here, E_{TX}represents the energy needed to transmit a message, while $E_{elec}$is the energy used per bit by the transceiver circuitry. The parameters $\varepsilon_{fs}$and ε_{mp}represent the amplifier energy for free-space and multipath communication, respectively. The threshold distance $d_0 = \sqrt{\varepsilon_{fs}/\varepsilon_{mp}}$ determines which propagation model is used. The energy required to receive a message of l bits is given by:

$$E_{RX}(l) = lE_{elec} \quad (3)$$

Where E_{RX} denotes the energy consumption of receiving the same message, and distance $d(t)$ is not static but dynamically updated based on node mobility. For node i, its position at time t is represented by $\mathrm{p}_i(t) = (x_i(t), y_i(t))$, where $x_i(t)$and $y_i(t)$show its current coordinates. Moreover, a Random Waypoint Model (RWM) is employed to select a destination for a node that randomly p_i^{dest} within the deployment area and moves forward with a velocity of $v_i(t)$ using magnitude $v_i \in [v_{min}, v_{max}]$. The position of the mobility node after reaching the random position is updated as,

$$\mathrm{p}_i(t + \Delta t) = \mathrm{p}_i(t) + \mathrm{v}_i(t)\Delta t \quad (4)$$

Here, $p_i(t)$represents the current position of node i, $v_i$indicates its velocity, and $\Delta t$is the time step. When the node reaches a randomly selected destination, it pauses for a short random period before choosing a new destination for data transmission. The distance travelled in each round is specified by $\Delta s_i(t) = \|\mathrm{p}_i(t + \Delta t) - \mathrm{p}_i(t)\|$ which helps to obtain the mobility energy cost, which is modelled as,

$$E_{\text{move},i}(t) = e_m \cdot \Delta s_i(t) \quad (5)$$

Here, $e_m$represents the energy used for movement per meter, measured in joules per meter. The total energy of node $i$at time $t + \Delta t$is then calculated as follows

$$E_i(t + \Delta t) = E_i(t) - \big(E_{TX}(l, d(t)) + E_{RX}(l) + E_{\text{move},i}(t) + E_{ctrl}(t) + E_{loc}(t)\big) \quad (6)$$

Here, E_i denotes the total energy of node i, and $E_{ctrl}(t)$ refers to the energy spent in transmitting or receiving control beacons owing to mobility-based neighbor discovery. $E_{loc}(t)$ specifies that the cost of localization is the position estimation performed. Thus, the mobility model influences the communication distances and adds an explicit mobility cost to ensure that the RE reflects both the data transfer and the physical

movement of the nodes. After selecting the next destination, the process is performed iteratively.

3.3 Trust and Security Model

The proposed framework supports dynamic cluster formation and routing while protecting the network from malicious or compromised sensor nodes. Each node's trust level is measured using a Beta-reputation model based on its successful and failed communication interactions [32]. For a node i evaluating node j, the direct trust value is calculated from the packet-forwarding behavior as:

$$T_{ij}^{D}(t) = \frac{S_{ij}(t)+1}{S_{ij}(t)+F_{ij}(t)+2} \quad (7)$$

where $S_{ij}(t)$ and $F_{ij}(t)$ represent the number of successful and unsuccessful interactions, respectively, observed between nodes i and j until time t. Laplace smoothing is achieved by adding one to the numerator and two to a denominator, so that the trust value never becomes exactly zero or one when there are only a limited number of observations. Therefore, a node that always forwards packets will have a high trust value, and a node that drops packets or fails to forward them repeatedly will have a lower trust value.

To enhance trust estimation in cases where direct observations are lacking, an indirect trust component is derived from neighboring nodes. The indirect trust can be expressed as

$$T_{ij}^{I}(t) = \frac{\sum_{k\in N_i} w_{ik} T_{kj}^{D}(t)}{\sum_{k\in N_i} w_{ik}} \quad (8)$$

where $T_{kj}^{D}(t)$ is the direct trust reported by neighbor k for node j, and w_{ik} represents the reliability weight assigned to the recommendation from node k. The overall trust value is then determined by aggregating direct and indirect trust as:

$$T_{ij}(t) = \lambda T_{ij}^{D}(t) + (1-\lambda) T_{ij}^{I}(t) \quad (9)$$

where $0 \leq \lambda \leq 1$ controls the relative contribution of direct observations and indirect recommendations. This combination enables the framework to distinguish reliable nodes from potentially compromised nodes while reducing the dependence on a single source of trust information. Recent trust-based WSN methods also use information from neighboring nodes along with direct observations to identify unreliable or malicious nodes more effectively [33] .

Besides communication behavior, the proposed model also includes Received Signal Strength Indicator (RSSI) information for estimating the mobility suitability of sensor nodes. The RSSI measurements are utilized to recognize unstable communication links and to evaluate the suitability of a node for cluster-head (CH) selection, as node mobility can lead to fast changes in link quality and frequent changes in topology. A node with stable RSSI and reliable trust value is more suitable to become CH because it can keep communicating with its cluster members for a longer period. In contrast, nodes with low trust or highly fluctuating RSSI are assigned lower suitability for CH selection and packet forwarding.

The proposed framework combines trust and mobility information with reinforcement learning and the Artificial Hummingbird Algorithm (AHA). Each node is evaluated using four main factors: residual energy, trust level, mobility, and distance, represented as $S = \{E, T, M, D\}$. Therefore, CH selection and routing decisions are not only based on geographical proximity and residual energy, but also jointly consider node reliability, communication stability, mobility and energy status. This integrated security mechanism reduces the probability of choosing compromised or unstable nodes as CHs, enhances the packet-forwarding reliability, and improves the overall security and robustness of the mobile wireless sensor network.

3.4 The Q-AHA Operational Framework

The execution of the Q-AHA protocol proceeds in distinct, sequential phases iteratively over discrete time steps (t) throughout the network lifecycle.

Phase 1: Initialization and Dynamic Mobility Update

At the initial deployment stage $(t = 0)$, the network is set up with a group of homogeneous sensor nodes with uniform maximum energy capacity, E_"max". The parameters of the local reinforcement learner are instantiated at the same time. The state-action value matrix (Q-table) is formally denoted as Q(s,a)=0 and is initialized to zero for every independent sensor node which acts as a Q-learning agent. This is a naive learning phase before any interaction with the environment or data transfer.

In every subsequent operational round or time step t, the network experiences a dynamic topological reconfiguration according to the Random Waypoint Mobility Model (RWMM). In this mobility architecture an active node i randomly chooses a target destination in the sensing perimeter and moves toward it with a velocity vector $V_i(t)$, uniformly chosen from a predefined velocity range $[v_{\text{min}}, v_{\text{max}}]$. As Q(s,a)=0, standard kinematic trajectory updates are employed to compute. the new spatial coordinates of the node:

$$P_i(t + \Delta t) = P_i(t) + V_i(t)\Delta t \tag{10}$$

where Δt is the duration of the current time step. During this transition, the protocol accurately records the physical distance traveled by the node, $\Delta d_i(t)$. This tracking is required to correctly calculate the mechanical mobility energy expenditure $E_{\text{mob}}(i, t)$,. This makes the residual energy assessment of the node consider both typical communication expenses and real physical movement.

To be able to retain a global spatial awareness in the highly dynamic environment, each node has to recompute its relative Euclidean distance to the mobile sink, $D_{BS}(i, t)$.Then nodes emit lightweight control beacon with their last location vectors. This broadcast strategy is localized. The control packet overhead is purposely minimized for saving battery power. It allows the mobile sink and the surrounding nodes to build a very accurate topological map in real time, which is very critical for the right state evaluations in the later Q-learning phase.

Phase 2: Q-Learning Based Localized Evaluation

To support smart and decentralized decision-making in a fast-changing mobile environment, each sensor node acts as an independent Reinforcement Learning (RL)

agent. Because the network topology is continuously changing, the state space seen by any node i is strictly time dependent. The state is a 4-tuple of critical operational parameters at each discrete time step t:

$$S_i(t) = \{E_i(t), T_v(i,t), M_f(i,t), D_{BS}(i,t)\} \quad (11)$$

where $E_i(t)$ represents the normalized residual energy, $T_v(i,t)$ denotes the established β-reputation trust value, $M_f(i,t)$ signifies the mobility factor derived from spatial variance, and $D_{BS}(i,t)$ is the calculated Euclidean distance to the mobile sink.

The agent chooses one of the two actions in the action space A={0,1} according to the observed state $S_i(t)$, Here 0 means staying as Normal Node and 1 means volunteering as Candidate CH. The agent adopts a ϵ-greedy action selection policy in an effort to balance the exploitation of the optimal actions known so far and the exploration of topological configurations not yet explored.

The effectiveness of the chosen action is instantaneously quantified by a comprehensive multi-objective reward function:

$$R_i(t) = w_1 E_i(t) + w_2 T_v(i,t) - w_3 M_f(i,t) - w_4 D_{BS}(i,t) \quad (12)$$

Here, w_1, w_2, w_3,and $w_4$are predefined weights whose total is equal to 1. The reward function gives higher rewards to nodes with more energy and higher trust levels and a penalty to high mobility (which implies link instability) and long distances to the sink.

After the reward acquisition, the state-action value of the agent is updated iteratively with the standard Bellman equation:

$$Q(S_i, a) \leftarrow Q(S_i, a) + \alpha \left[R_i(t) + \gamma \max_{a'} Q(S_i', a') - Q(S_i, a) \right] \quad (13)$$

where α∈(0,1] is the learning rate that determines how quickly the agent learns and γ∈[0,1) is the discount factor that gives priority to immediate rewards. Q-learning has also been applied to energy-efficient routing in WSN to support adaptive routing decisions based on network conditions [34] If it converges in the current topology, the resulting Q-value is the learned node quality score $Q_v(i,t)$. This score is a very compact measure of the local viability of the node and is sent back to the mobile sink to drive the global Artificial Hummingbird Algorithm (AHA) optimization phase.

In the proposed Q-AHA framework, Q-learning is used to evaluate the suitability of sensor nodes for cluster-head selection rather than selecting the cluster heads directly. The reward function considers important factors such as residual energy, trust, mobility, and distance to the mobile sink to assess node performance as network conditions change. Although the reward does not directly depend on the selected action, the Q-values are updated over time based on previous observations, allowing the system to maintain a history of node quality. These learned Q-values are then used by the Artificial Hummingbird Algorithm (AHA), along with other network parameters,

to identify suitable cluster heads. In this way, Q-learning supports the evaluation of individual nodes, while AHA performs the final cluster-head selection by considering the overall network conditions.

Phase 3: Global Optimization via Artificial Hummingbird Algorithm (AHA)

The heavy combinatorial computation of clustering the whole network is offloaded to the mobile sink, which is resource rich, to avoid the computational limitations of each sensor node in the proposed protocol. The mobile sink gathers the decentralized learned quality scores (Q_v) from the local Q-learning agents, and then initiates the optimization phase of Artificial Hummingbird Algorithm (AHA). AHA dynamically finds the globally optimal number and spatial distribution of Cluster Heads (CHs) adapted to the current physical topology of the network. The algorithm begins with an initial population of candidate cluster head (CH) configurations—representing potential routing topologies—and evaluates each set X using a multi-objective fitness function governed by dynamic weights tailored to the current time step

$$F(X,t) = \beta_1 E_i(t) + \beta_2 T_v(t) + \beta_3 Q_v(t) + \beta_4 N_d(t) - \beta_5 D_{BS}(t) - \beta_6 C_l(t) \quad (14)$$

Here, $E_i(t)$shows the total remaining energy, $T_v(t)$represents the network trust level, and $Q_v(t)$gives the local RL score. $N_d(t)$indicates the nearby node density, $D_{BS}(t)$represents the distance to the sink, and $C_l(t)$shows the expected cluster load. The weights $\beta_1$to $\beta_6$are used to balance these different factors during optimization. AHA has also been extended to multi-objective optimization problems [35]

The proposed AHA optimization framework does not rely on static weighting coefficients to guaranty the long term network survivability and avoid early depletion of overused routing nodes. Instead, it employs a time-varying dynamic weight adaptation mechanism. As the network matures in its lifecycle, the priorities of the routing protocol need to change fundamentally; early stage operations require a focus on minimal latency and shorter routing paths, while late stage operations require to aggressively focus on energy conservation and trust to avoid catastrophic data loss.

This switch of dynamics is governed by a time-dependent transition parameter δ, which is given by the fraction of the current time step t and the maximum number of simulated rounds $T_{\max}$:

$$\delta = \frac{t}{T_{\max}} \quad (15)$$

The optimization weights, i.e., the residual energy coefficient $\beta_1(t)$ and the distance penalty coefficient $\beta_5(t)$, are scaled dynamically with a linear progression model:

$$\beta_1(t) = \beta_{1,\min} + \delta \times \left(\beta_{1,\max} - \beta_{1,\min}\right) \quad (16)$$

$$\beta_5(t) = \beta_{5,\max} - \delta \times \left(\beta_{5,\max} - \beta_{5,\min}\right) \quad (17)$$

As the network ages (i.e., δ→1), the framework mathematically increases the importance of residual energy (β_1) and relaxes the penalty on distance (β_5) allowing

the selection of geographically longer, but highly energetic and secure, routing paths. To ensure stability of AHA foraging behavior algorithmically, each of dynamically adjusted coefficients are strictly normalized at each time step t such that the sum of them is unity:

$$\beta_k^{\text{norm}} = \frac{\beta_k(t)}{\sum_{j=1}^{6} \beta_j(t)} \quad (18)$$

This continuous normalization ensures the multi-objective fitness landscape is mathematically bounded, which allows the Artificial Hummingbird Algorithm to consistently converge to optimal, energy-aware, and highly secure Cluster Head configurations, irrespective of the network's operational age.

The AHA applies three mathematically modeled, bio-inspired foraging behaviors to iteratively improve these candidate solutions:

Guided Foraging (Exploitation): The algorithm guides the candidate solutions toward the best known CH configuration. It mathematically models the goal-directed flight behavior of hummingbirds to conduct a localized search, exploiting the high-quality routing regions to strictly minimize the intra-cluster distances and reduce the immediate transmission costs.Territorial Foraging (Exploration): This phase forces candidate solutions to explore adjacent, unvisited search spaces to prevent the algorithm from converging too early on suboptimal network topologies. It perturbs current CH selections by evaluating the neighbor nodes to find potentially better and highly secure configurations that Guided Foraging might miss.Migration Foraging (Maintenance of Diversity): If a CH candidate set displays persistently low fitness (like an exhausted nectar source), the algorithm systematically replaces such low-fitness solutions with completely new candidate solutions, randomly distributed over the sensing area. This mechanism allows for high population diversity and prevents the optimization matrix from stagnating in case of rapid changes of node mobility.The global optimal clustering topology is chosen as the candidate set with the highest multi-objective fitness value after the algorithm has iterated the prescribed number of times, which is mathematically expressed as

$$CH^* = \text{argmax}(F(X,t)))$$

Finally, this authoritative CH selection matrix is broadcasted by the mobile sink back to the decentralized sensor array. This initiates the physical cluster formation and subsequent data transmission phase.

Phase 4: Cluster Formation, Data Transmission, and Energy Depletion

Following the AHA broadcast, normal member nodes associate with the nearest designated CH to form discrete clusters based on their current coordinates $P_i(t)$. Intra-cluster communication follows a single-hop pattern (Node →CH). The cluster heads (CHs) aggregate the incoming sensor data to remove redundancies before performing inter-cluster routing to transmit the final payload to the mobile sink.

Importantly, at the conclusion of this data transmission phase, the Residual Energy $E_i(t)$ for every node is updated. The system deducts the dynamic transmission costs $E_{Tx}(t)$, reception costs $E_{Rx}(t)$, the physical mobility costs calculated in Phase 1 $E_{mob}(t)$,

and operational overheads, strictly adhering to the Dynamic Energy Consumption Model.

Algorithm: Q-AHA

Input: N, Tmax, α, γ, ε, w1–w4, β1–β6, λ, θE, θP, Δmax, θRSSI, κ, Npop, Imax

Output: Optimal CH set for each round, performance metrics

1. Initialize Ei ← Emax, Qi(s,a) ← 0, Ti ← 1, Sij ← 0, Fij ← 0; initialize node and sink positions.
2. for t = 1 to Tmax do
3. if all nodes are dead then terminate.
4. δ ← t / Tmax.

Phase 1: Mobility and Event-Triggered Beaconing

5. for each active node i do
6. Update Pi(t); compute travelled distance Δdi, mobility energy Emob(i), and sink distance DBS(i).
7. if energy change > θE OR position change > θP OR role changes OR beacon interval ≥ Δmax then
8. Broadcast {position, residual energy, role}; update last-broadcast state.
9. end if
10. end for

Phase 2: Local Q-Learning Evaluation

11. for each active node i do
12. Construct Si = {normalized energy, trust, mobility, normalized sink distance}.
13. if t > 1 then
14. Update Q-value using the Bellman equation.
15. end if
16. Select ai ∈ {0,1} using ε-greedy policy.
17. Ri = w1·Energy + w2·Trust − w3·Mobility − w4·SinkDistance.
18. Qv(i) ← Q(Si, CH-action); send Qv(i) to the mobile sink.
19. end for

Phase 3: Global CH Optimization using AHA

20. Update dynamic AHA weights using δ and normalize β1…β6.
21. Generate an AHA population of candidate CH sets from eligible nodes.
22. for each candidate solution Xp do
23. F(Xp) = β1·ResidualEnergy + β2·Trust + β3·QScore + β4·NodeDensity − β5·SinkDistance − β6·ClusterLoad.
24. end for
25. Xbest ← candidate with highest fitness.

for iter = 1 to Imax do
  for each candidate Xp do
    Apply Guided, Territorial, or Migration Foraging.
    Repair Xp if required and recompute fitness.
    if F(Xp) > F(Xbest) then Xbest ← Xp.
  end for
end for
CH* ← Xbest; broadcast CH* to all active nodes.

Phase 4: Clustering, Transmission, Trust, and Energy Update

Assign each normal node to the nearest selected CH.
Perform Sensor Node → CH → Mobile Sink transmission; CHs aggregate before forwarding.
for each communication link (i,j) do
  On success Sij ← Sij + 1; otherwise Fij ← Fij + 1.
  TDij = (Sij + 1)/(Sij + Fij + 2).
  if informed neighboring nodes exist then
    TIij = Σ(wik·TDkj)/Σwik.
  else TIij ← TDij.
  Tij = λ·TDij + (1−λ)·TIij.
  if RSSI variation > θ RSSI then Tij ← Tij·κ.
end for
for each node i do
  Ei(t) ← Ei(t−1) − (ETX + ERX + Emob + Ectrl + Eloc).
  if Ei(t) ≤ 0 then mark node i as dead.
end for
Record alive nodes, FND, HND, LND, PDR, throughput, delay, JFI, and residual energy.
end for
Return CH history, trust state, residual energy, final topology, and performance metrics.

4 Simulation results and discussion

The proposed Trust-Aware Hybrid Q-AHA Energy-Efficient Secure Clustering Protocol is evaluated through simulations using 50, 100 and 200 sensor nodes randomly placed within a defined area of 100 x 100 m. Each node was initiated with 1.00 J of energy . The Q-learning part was with 81 state space and two actions. The population size of the AHA algorithm was 20. A mobile sink node was introduced to increase the efficiency of data collecting and to decrease the connection overhead.

The proposed Q-AHA protocol along with benchmark protocols namely RL-PSO, LEACH, LEACH-RLC and M-TAAHA were deployed and tested in the same environment to get a fair and consistent comparison between them. The network settings of all the protocols were the same, including the deployment area, the number of sensor nodes, the initial sensor-node energy, the packet size, the ratio of malicious nodes, the radio energy parameters, the duration of rounds, and the length of the simulation. Only the clustering, routing, learning, trust and optimization mechanisms of the protocols were kept based on their algorithms. Hence, the observed performance difference is mostly due to the way each protocol works and not because of the simulation setup.

Table 2

Metric	**Value**
Number of nodes	50,100,200
Simulation area	$100\ \mathrm{m} \times 100\ \mathrm{m}$
Initial energy/node	1.00 J
E_{elec}	50×10^{-9}J/bit
Data packet size	4000 bits
Round Duration	20s
γ	0.9
α	0.1
Node mobility	0.5-5 m/s

4.1 Performance Analysis

The proposed Q-AHA was compared with RL-PSO, LEACH, LEACH-RLC and M-TAAHA in three different network sizes of 50, 100 and 200 sensor nodes. To perform the comparison, the performance metrics used were: Network Lifetime metrics such as First Node Death (FND), Half Node Death (HND), and Last Node Death (LND); Packet Delivery Ratio (PDR); Throughput; End-to-End Delay. The results obtained demonstrate that Q-AHA yields good overall performance, including high network lifetime, packet delivery and throughput.

4.1.1 Network Lifetime Analysis

Q-AHA had 1788, 2244 and 3720 rounds for FND, HND and LND respectively for 50 nodes. It has LND around 65.85% higher than the second highest LEACH-RLC. In the case of 100 nodes, LEACH-RLC had the maximum FND of 1942 rounds and Q-AHA had 1746 rounds. Q-AHA, however, had the highest HND of 2058 rounds and LND of 3723 rounds which was about 65.76% better than LEACH-RLC. LEACH-RLC recorded the best FND and HND, while Q-AHA had the best LND with 3727 rounds, which is approximately 56.40% more than the next best LEACH. Generally, Q-AHA does not guarantee the best early node survival for larger networks, but maintains very good LND for long-term energy balancing and for network lifetime.

Network Size	**Metric**	**RL-PSO**[23]	**LEACH**[8]	**LEACH-RLC** [26]	**M-TAAHA** [30]	**Q-AHA (Proposed)**

50 nodes	FND	970	863	1749	1182	1788
	HND	1183	1251	1812	1747	2244
	LND	1295	1509	2243	1816	3720
100 nodes	FND	854	1046	1942	1062	1746
	HND	1212	1581	1997	1836	2058
	LND	1355	1892	2246	1919	3723
200 nodes	FND	1105	1277	1970	1005	1477
	HND	1203	1900	1995	1931	1861
	LND	1332	2383	2107	1981	3727

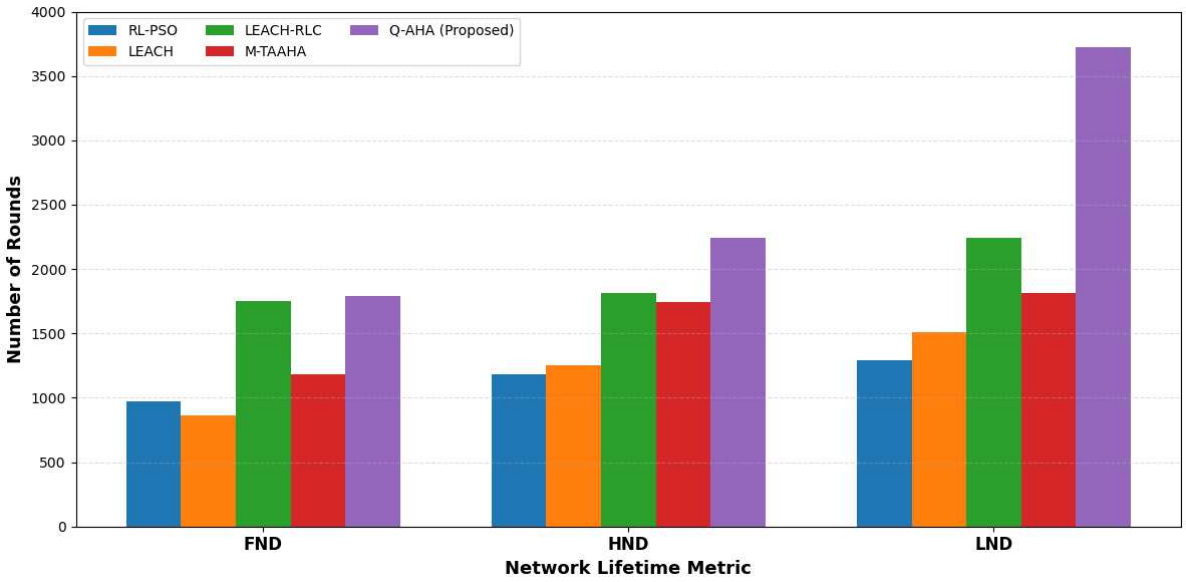


Fig 2 Comparison of FND, HND and LND for 50 Nodes

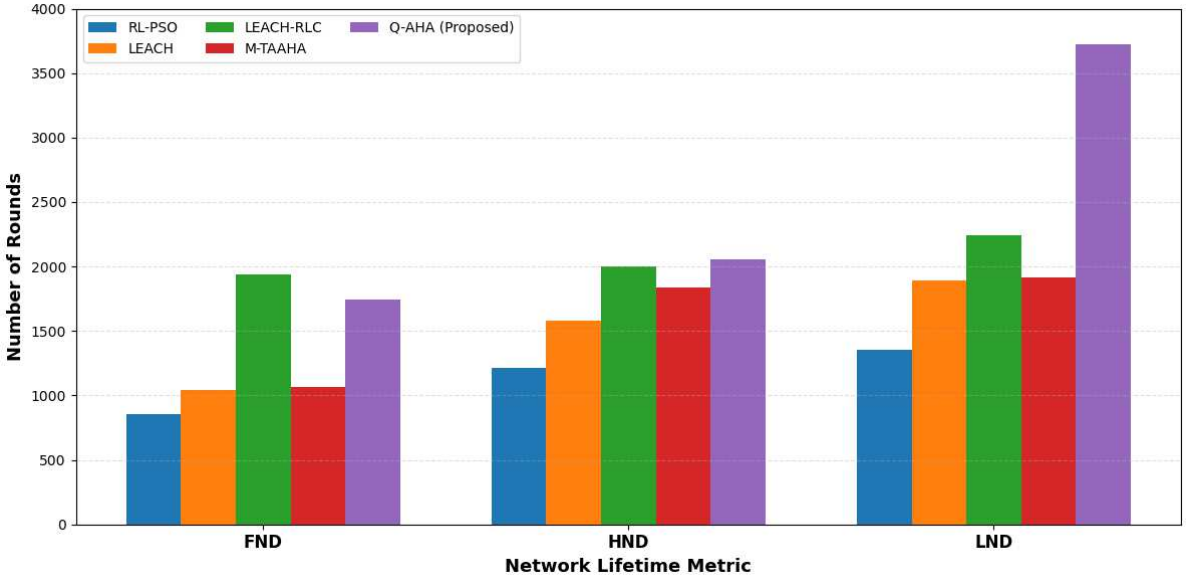


Fig 3 Comparison of FND, HND and LND for 100 Nodes

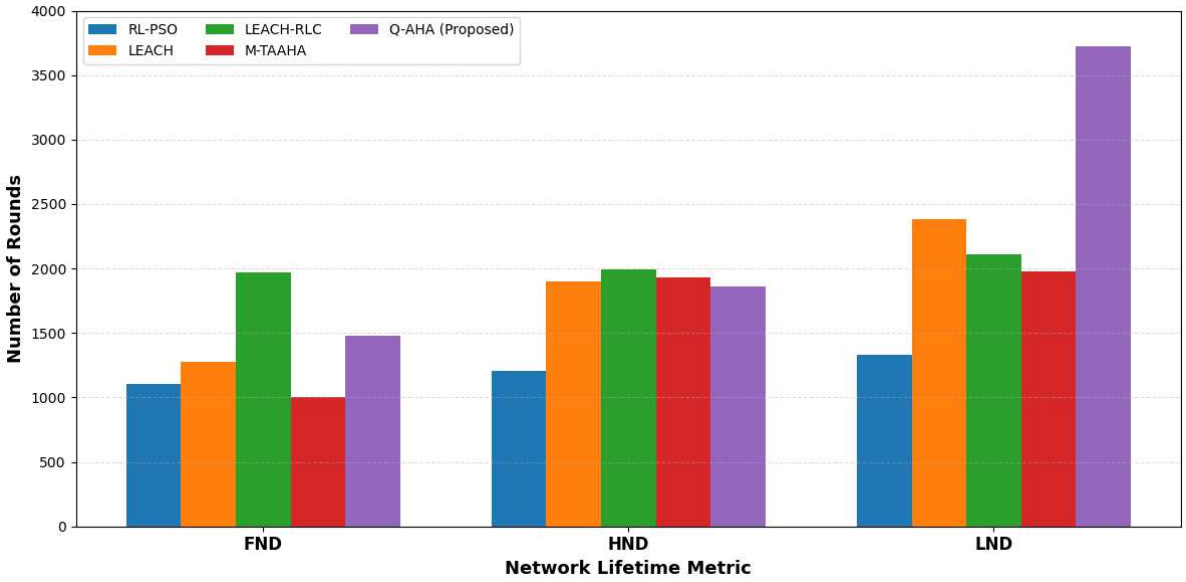


Fig 4 Comparison of FND, HND and LND for 200 Nodes

4.1.2 Packet Delivery Ratio Analysis

Q-AHA achieved average PDR values of 93.55%, 94.86%, and 95.44% for 50, 100, and 200 nodes, respectively. For 50 nodes LEACH-RLC slightly perform better than the Q-AHA with 94.15%. But Q-AHA has the maximum PDR at 100 nodes and 200 nodes with an enhancement of around 1.76% and 2.20% with respect to LEACH-RLC respectively. These results demonstrate that Q-AHA can deliver reliable packet delivery service and is scalable in the sense of increasing network size.

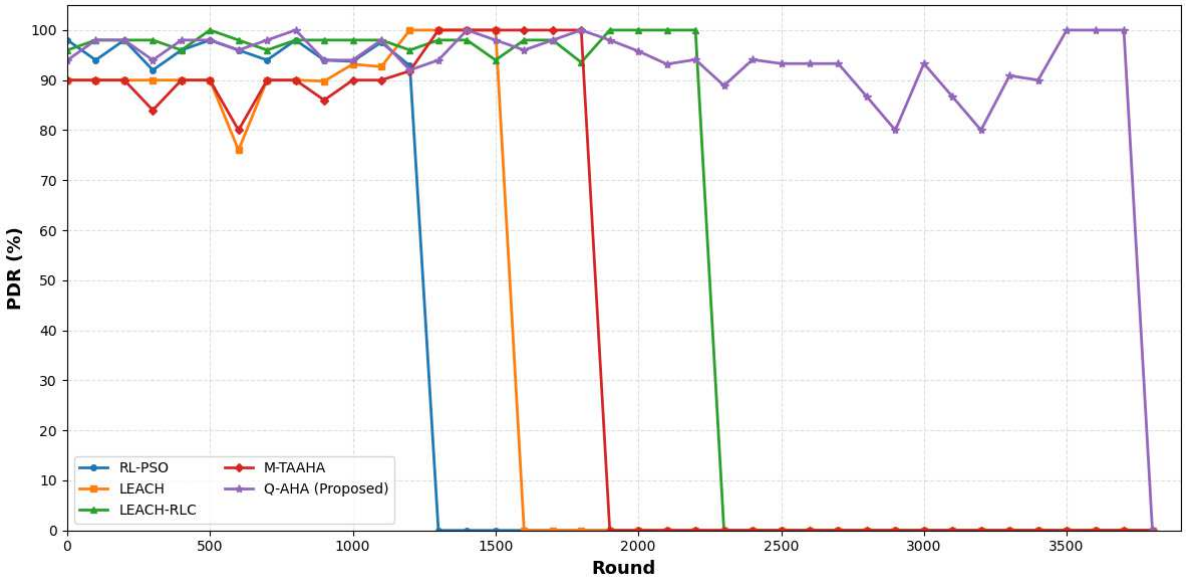


Fig 5 Comparison of PDR for 50 Nodes

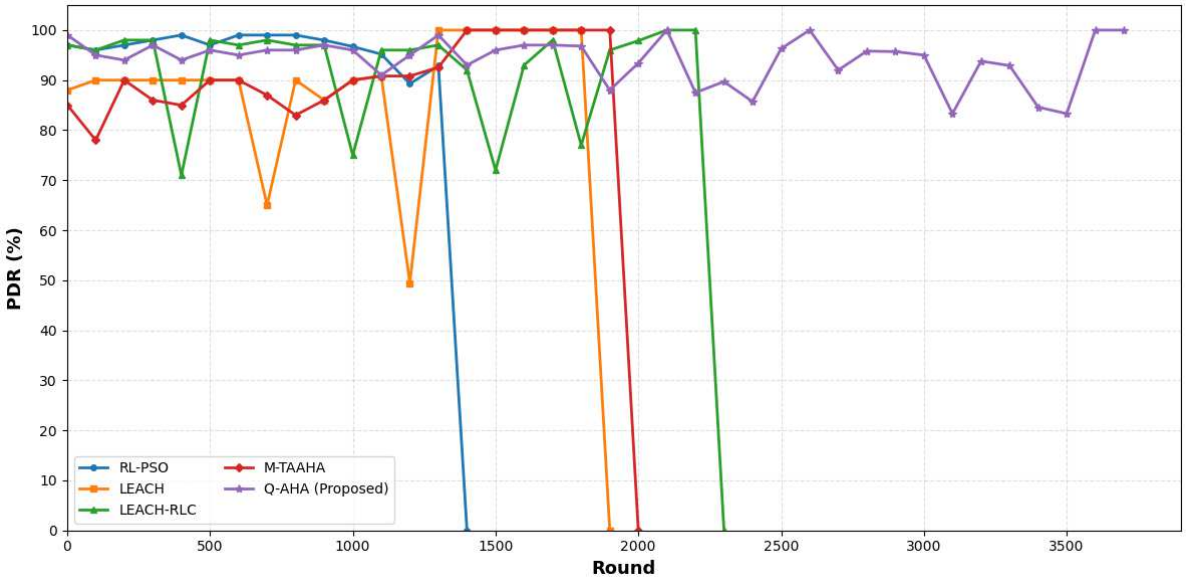


Fig 6 Comparison of PDR for 100 Nodes

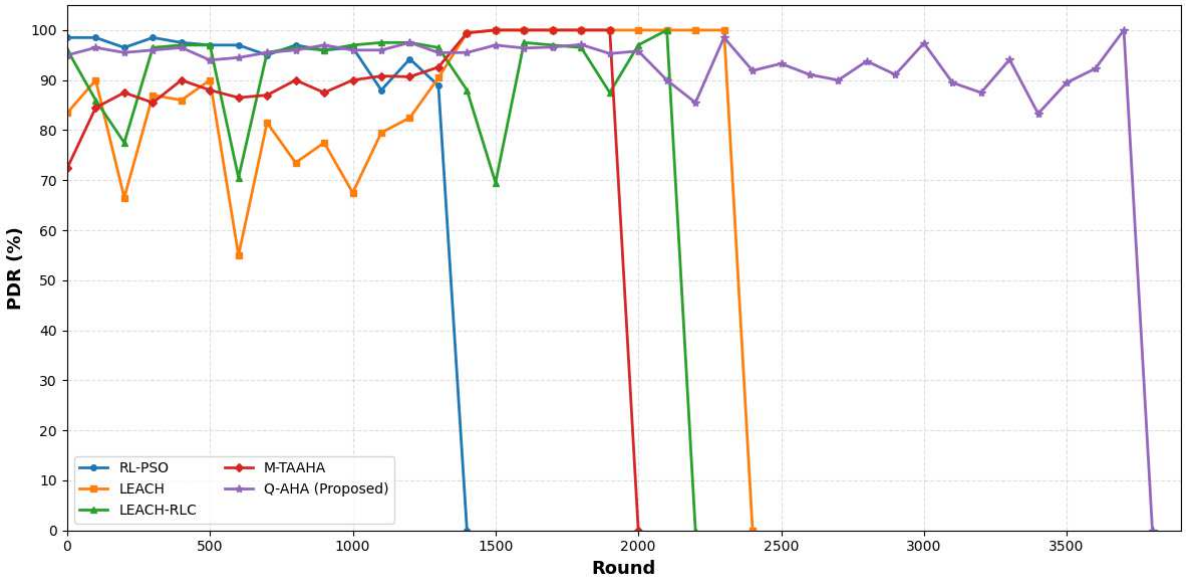


Fig 7 Comparison of PDR for 200 Nodes

4.1.3 Throughput Analysis

Q-AHA had the best performance on all sizes of networks. It recorded 13.2045 kbps, 23.9824 kbps, and 45.7374 kbps for 50, 100, and 200 nodes, respectively. For 50 nodes, 100 nodes, and 200 nodes, the throughput of Q-AHA was 56.93% higher, 43.16% higher, and 29.40% higher respectively than the best competing protocol. Q-AHA is a scale-up product and can support higher traffic while ensuring good data delivery performance.

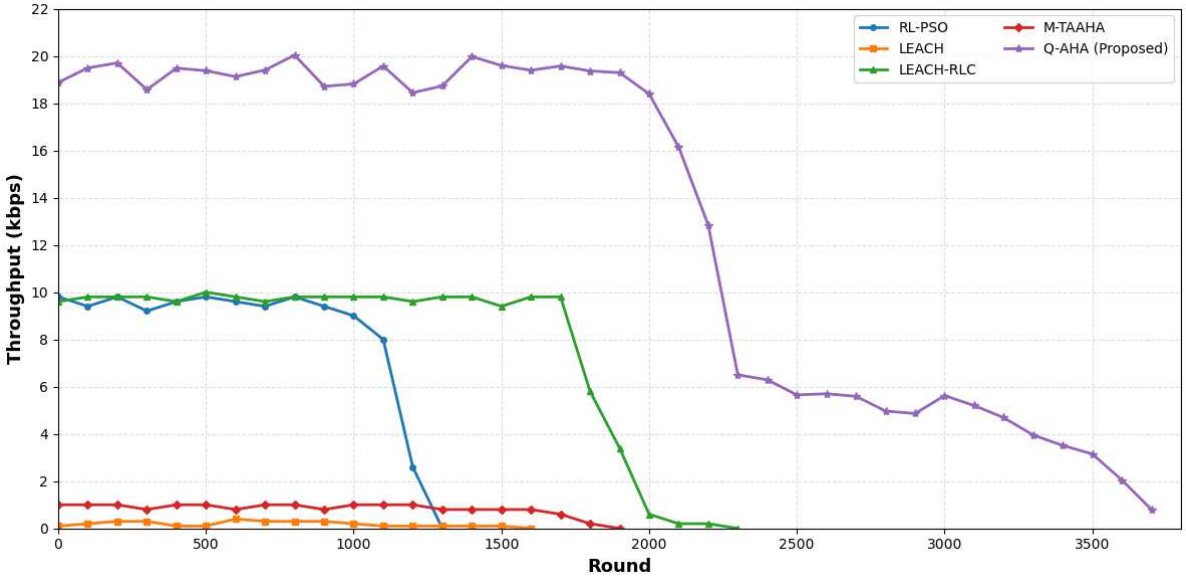


Fig 8 Comparison of Throughput (kbps) for 50 Nodes

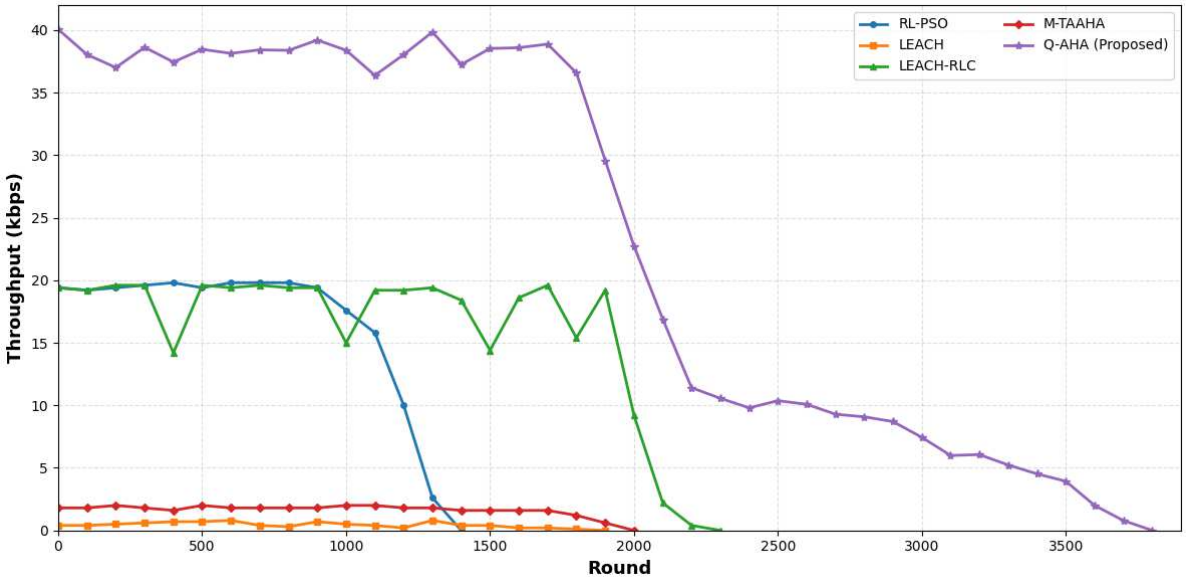


Fig 9 Comparison of Throughput (kbps) for 100 Nodes

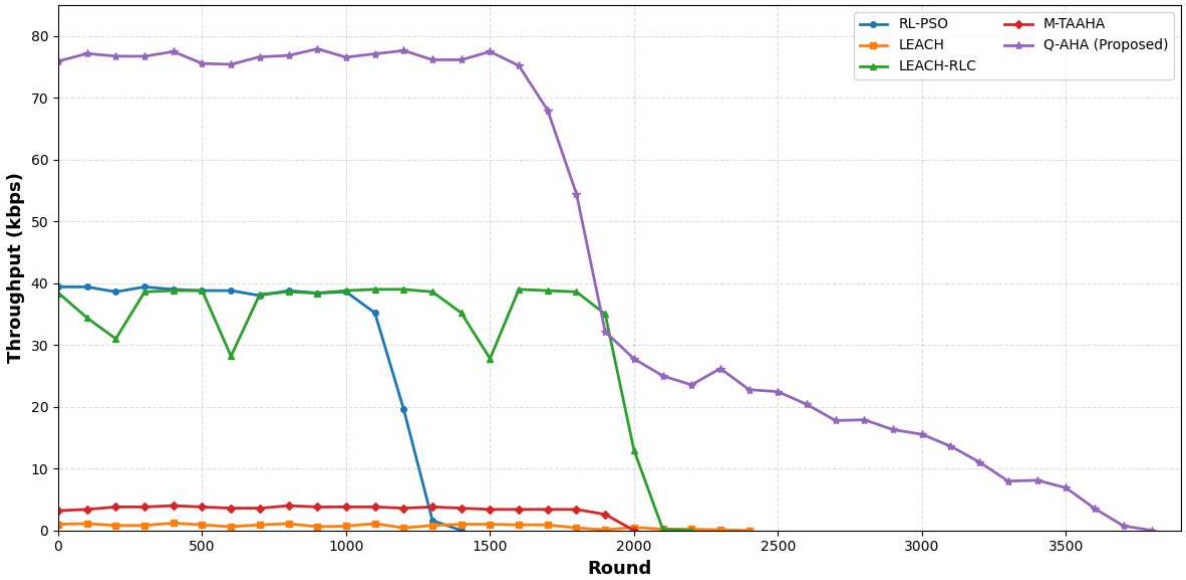


Fig 10 Comparison of Throughput (kbps) for 200 Nodes

4.1.4 End-to-End Delay Analysis

This gives best result for 50 and 100 nodes with delay 9.550ms and 10.1543ms respectively which is around 28.25% and 5.31% less than LEACH-RLC. For 200 nodes, the delay of LEACH-RLC was the lowest (9.369 ms) and the delay of Q-AHA was 11.4244 ms, which is also performed better than RL-PSO, LEACH, M-TAAHA. On average, Q-AHA offers low delay in various network sizes and the PDR and throughput are high.

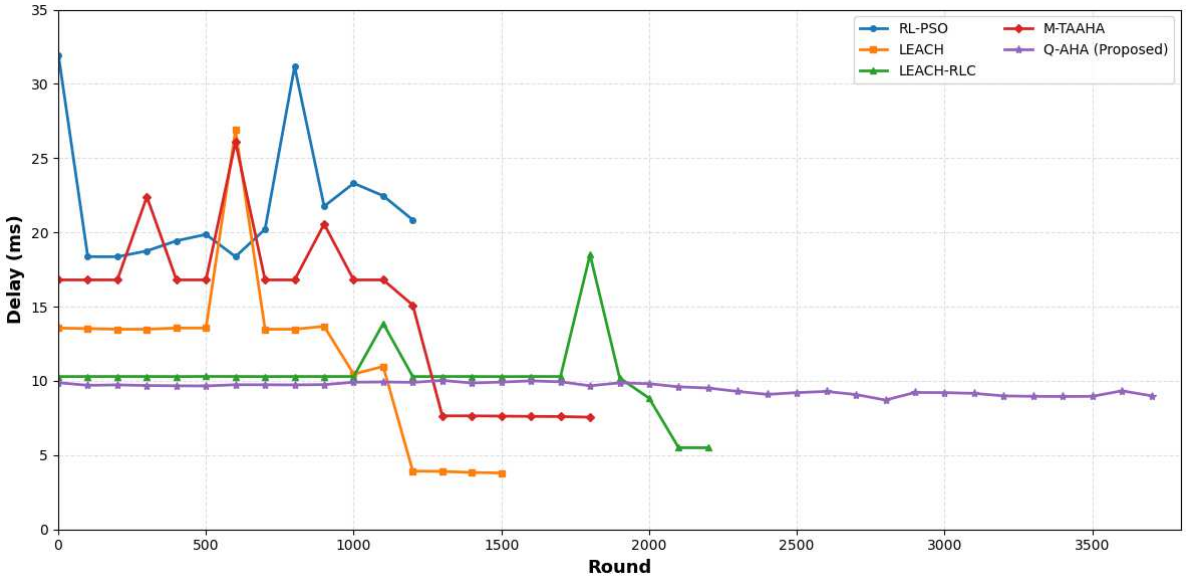


Fig 11 Comparison of Delay (ms) for 50 Nodes

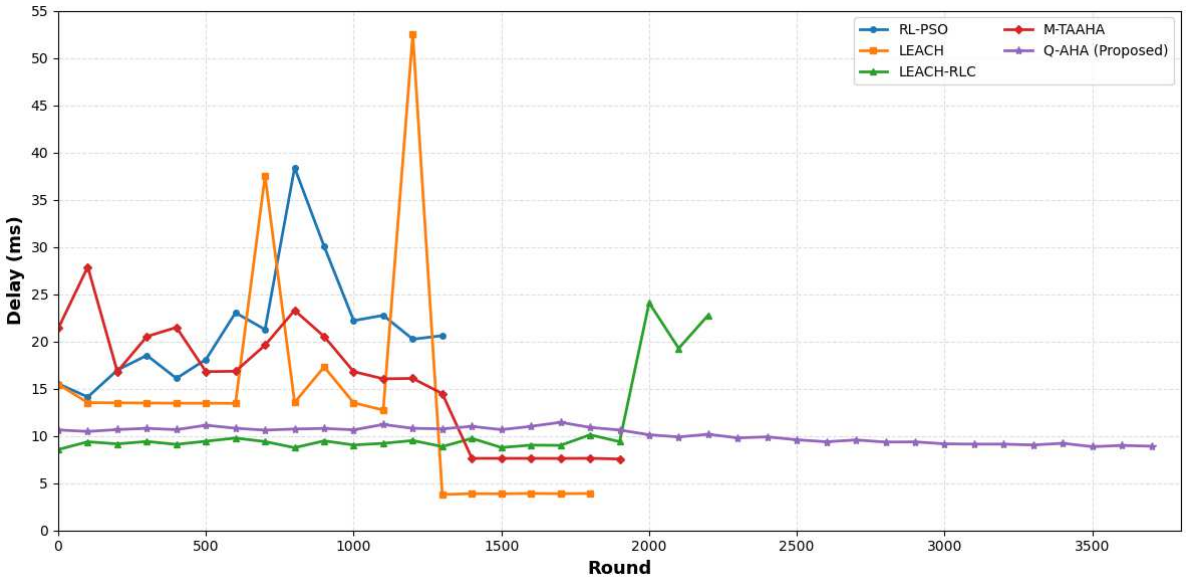


Fig 12 Comparison of Delay (ms) for 100 Nodes

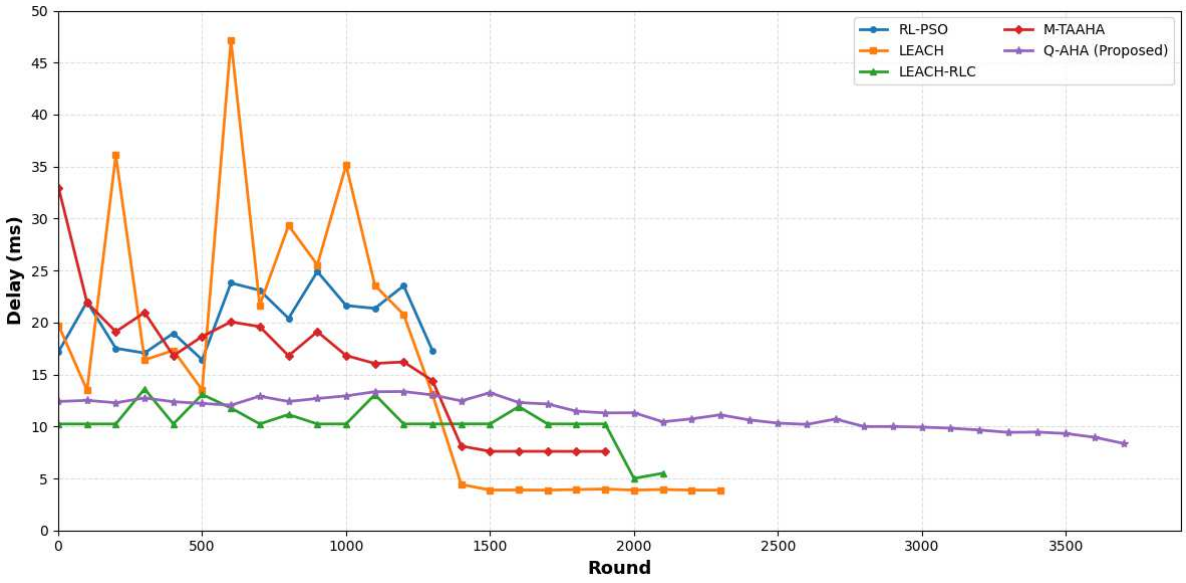


Fig 13 Comparison of Delay (ms) for 200 Nodes

Q-AHA offers a good tradeoff between network lifetime, PDR, throughput and delay. It performed best in terms of LND for 50,100 and 200 nodes and got the highest throughput all the time. The best PDR achieved by Q-AHA was 100 nodes and 200 nodes while the lowest delay was achieved for 50 nodes and 100 nodes. While LEACH-RLC was more efficient in some FND and delay scenarios, Q-AHA provides the optimum tradeoff between scalability, reliability in delivering data, and network longevity for scalable WSN applications.

5 Conclusion

In this paper, we presented Trust Aware Unified Hybrid Q-Learning and Artificial Hummingbird Algorithm (Q-AHA) for secure and energy efficient clustering and routing in Mobile Wireless Sensor Networks. The proposed technique combines Q-learning for localized trust aware node evaluation and AHA for optimum Cluster Head selection.

An event-triggered mechanism is used to decrease needless communication overhead.

To make a fair comparison, these algorithms (Q-AHA, RL-PSO, LEACH, LEACH-RLC and M-TAAHA) were tested in the same simulation environment of 50, 100 and 200 nodes. Q-AHA recorded the highest LND value for rounds 3720, 3723 and 3727, as well as the highest throughput for all network sizes, and the highest PDR for 100 and 200 nodes. It also had the smallest delay of 50 and 100 nodes. Overall, Q-AHA offers a good trade-off between network lifetime, reliability, throughput, and delay, which are appropriate for scalable and secure MWSNs. In future, larger heterogeneous networks, more types of attacks, better mobile-sink path planning, and real-world validation will be taken into account.

Acknowledgements

Not applicable.

Author Contributions

Keerthi K.S.: Conceptualization, methodology, software development, algorithm implementation, simulation experiments, formal analysis, data visualization, and writing – original draft preparation.

Jagadisha N.: Supervision, methodology review, validation, interpretation of results, and writing – review and editing.

Both authors reviewed and approved the final version of the manuscript.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Clinical Trial Registration

Not applicable. This study does not involve clinical trials or human participants.

Ethics Declaration

Not applicable. This study is based on computational research and simulation of wireless sensor networks and does not involve human participants or animals.

Consent to Publish declaration: not applicable

Consent to Participate declaration: not applicable

References

[1] S. Mody, S. Mirkar, R. Ghag, and P. Kotecha, “Cluster Head Selection Algorithm For Wireless Sensor Networks Using Machine Learning,” in *2021 International Conference on Computational Performance Evaluation (ComPE)*, IEEE, Dec. 2021, pp. 445–450. doi: 10.1109/ComPE53109.2021.9752264.

[2] M. F. Tewelgne, S. A. Demilew, and D. W. Girmaw, “Energy efficient inter-cluster multi-hop communication routing protocol for wireless sensor network based on centralized energy efficient clustering routing protocol,” *Discover Applied Sciences*, vol. 7, no. 7, Jul. 2025, doi: 10.1007/s42452-025-07348-9.

[3] H. Li, Y. Dai, Q. Chen, D. Liao, and H. Jin, “Energy efficient mobile sink driven data collection in wireless sensor network with nonuniform data,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-79825-x.

[4] M. Krishnan and Y. Lim, “Reinforcement learning-based dynamic routing using mobile sink for data collection in WSNs and IoT applications,” *Journal of Network and Computer Applications*, vol. 194, Nov. 2021, doi: 10.1016/j.jnca.2021.103223.

[5] G. Kaur and M. Bhattacharya, “Green Fault Tolerant AIoT-Enabled Mobile Sink Data Collection Scheme in Sensor Networks,” *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15385–15394, 2024, doi: 10.1109/TVT.2024.3400880.

[6] M. J. Kavitha, M. R. Geetha, and R. I. Sajan, “Knowledge-based adaptive routing for energy efficiency and attack detection in ad hoc wireless sensor networks,” *Computer Networks*, vol. 259, Mar. 2025, doi: 10.1016/j.comnet.2025.111086.

[7] S. Suma and B. Harsoor, “Detection of malicious activity for mobile nodes to avoid congestion in Wireless Sensor Network,” in *Proceedings of IEEE 2022 4th International Conference on Advances in Electronics, Computers and Communications, ICAECC 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICAECC54045.2022.9716673.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-Efficient Communication Protocol for Wireless Microsensor Networks,” 2000.

[9] G. Thahniyath and M. Jayaprasad, “Secure and load balanced routing model for wireless sensor networks,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 7, pp. 4209–4218, Jul. 2022, doi: 10.1016/j.jksuci.2020.10.012.

[10] X. Wang and H. Chen, “A Reinforcement Learning-Based Dynamic Clustering Algorithm for Compressive Data Gathering in Wireless Sensor Networks,” *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/2736734.

[11] K. C. Chilukuri, “Throughput and Energy Efficient Routing using Reinforcement Learning for Wireless Sensor Networks,” in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 1629–1639. doi: 10.1016/j.procs.2025.04.118.

[12] Q. Tang and F. Nie, “Clustering routing algorithm of wireless sensor network based on swarm intelligence,” *Wireless Networks*, vol. 30, no. 9, pp. 7227–7238, Dec. 2024, doi: 10.1007/s11276-023-03584-2.

[13] R. K. Yadav and R. P. Mahapatra, “Hybrid metaheuristic algorithm for optimal cluster head selection in wireless sensor network,” *Pervasive Mob. Comput.*, vol. 79, Jan. 2022, doi: 10.1016/j.pmcj.2021.101504.

[14] W. Zhao, L. Wang, and S. Mirjalili, “Artificial hummingbird algorithm: A new bio-inspired optimizer with its engineering applications,” *Comput. Methods Appl. Mech. Eng.*, vol. 388, Jan. 2022, doi: 10.1016/j.cma.2021.114194.

[15] L. Liu, R. Song, and Q. Wei, “Event-Triggered Energy Optimization of Wireless Sensor Networks,” in *ICNSC 2022 - Proceedings of 2022 IEEE International Conference on Networking, Sensing and Control: Autonomous Intelligent Systems*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICNSC55942.2022.10004164.

[16] A. H. R. Fofana and J. Lei, “Reinforcement Learning Based Cluster Formation Algorithm in Wireless Sensor Networks,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Dec. 2023, pp. 367–371. doi: 10.1145/3654446.3654512.

[17] K. Debasis, L. D. Sharma, V. Bohat, and R. S. Bhadoria, “An Energy-Efficient Clustering Algorithm for Maximizing Lifetime of Wireless Sensor Networks using Machine Learning,” *Mobile Networks and Applications*, vol. 28, no. 2, pp. 853–867, Apr. 2023, doi: 10.1007/s11036-023-02109-7.

[18] N. E. H. Bourebia and C. Li, “A greedy energy efficient clustering scheme based reinforcement learning for WSNs,” *Peer. Peer. Netw. Appl.*, vol. 15, no. 6, pp. 2572–2588, Nov. 2022, doi: 10.1007/s12083-022-01368-7.

[19] P. Singh and R. Vir, “Enhanced energy-aware routing protocol with mobile sink optimization for wireless sensor networks,” *Computer Networks*, vol. 261, Apr. 2025, doi: 10.1016/j.comnet.2025.111100.

[20] A. Jukuntla and V. Dondeti, “Energy-efficient mobile sink data collection in wireless sensor networks using LEACH with rendezvous nodes,” *Measurement (Lond).*, vol. 221, Nov. 2023, doi: 10.1016/j.measurement.2023.113541.

[21] S. Sangeetha, T. A. A. Victoire, M. Premkumar, and R. Sowmya, “ExAq-MSPP: An Energy-Efficient Mobile Sink Path Planning Using Extended Aquila Optimization Algorithm,” *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, Dec. 2024, doi: 10.1007/s44196-024-00670-x.

[22] G. Sulakshana and G. R. Kamatam, “Data accumulation in WSNs using a mobile sink: A linear programming approach,” *Measurement: Sensors*, vol. 27, Jun. 2023, doi: 10.1016/j.measen.2023.100743.

[23] A. Choudhary, N. C. Barwar, and V. Chouhan, “Energy Efficient Cluster Head Selection Using Hybrid RL-PSO Approach,” *Journal of Intelligent Systems and Internet of Things*, vol. 11, no. 2, pp. 8–21, 2024, doi: 10.54216/JISIoT.110201.

[24] I. Chiraira *et al.*, “A Scalable Reinforcement Learning Framework for Wireless Sensor Networks: Spatiotemporal Cluster-Based Multi-Agent Reinforcement Learning,” in *International Conference on Artificial Intelligence, Computer, Data Sciences, and Applications, ACDSA 2025*, Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/ACDSA65407.2025.11166147.

[25] S. S. Mouli, T. V. Talagondapati, and M. P. Singh, “Multi-Faceted Reinforcement Learning Frameworks for Dynamic Cluster Head Selection and Energy-Efficient Routing in Wireless Sensor Networks,” *SN Comput. Sci.*, vol. 6, no. 5, Jun. 2025, doi: 10.1007/s42979-025-04031-z.

[26] F. F. Jurado-Lasso, J. F. Jurado, and X. Fafoutis, “LEACH-RLC: Enhancing IoT Data Transmission with Optimized Clustering and Reinforcement Learning,” Mar. 2025, doi: 10.1109/JIOT.2025.3552126.

[27] T. Sharma, A. Balyan, R. Nair, P. Jain, S. Arora, and F. Ahmadi, “ReLeC: A Reinforcement Learning-Based Clustering-Enhanced Protocol for Efficient Energy Optimization in Wireless Sensor Networks,” *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/3337831.

[28] A. Choudhary and N. C. Barwar, “Optimizing Clustering in Wireless Sensor Networks: A Synergistic Approach Using Reinforcement Learning (RL) and Particle Swarm Optimization (PSO),” *SN Comput. Sci.*, vol. 5, no. 6, Aug. 2024, doi: 10.1007/s42979-024-03080-0.

[29] E. Jayabalan and R. Pugazendi, “An efficient routing protocol for wireless body sensor networks using reinforced learning algorithm in clusters,” *Measurement: Sensors*, vol. 27, Jun. 2023, doi: 10.1016/j.measen.2023.100730.

[30] A. K. Jemla Naik, M. Parameswarappa, and M. N. Ramachandra, “Multiobjective, trust-aware, artificial hummingbird algorithm-based secure clustering and routing with mobile sink for wireless sensor networks,” *ETRI Journal*, vol. 46, no. 6, pp. 950–964, Dec. 2024, doi: 10.4218/ETRIJ.2023-0330.

[31] N. Temene, C. Sergiou, C. Georgiou, and V. Vassiliou, “A Survey on Mobility in Wireless Sensor Networks,” Feb. 01, 2022, *Elsevier B.V.* doi: 10.1016/j.adhoc.2021.102726.

[32] W. Fang, C. Zhang, Z. Shi, Q. Zhao, and L. Shan, “BTRES: Beta-based Trust and Reputation Evaluation System for wireless sensor networks,” *Journal of Network and Computer Applications*, vol. 59, pp. 88–94, Jan. 2016, doi: 10.1016/j.jnca.2015.06.013.

[33] Y. Han, H. Wang, Y. Li, and L. Zhang, “Trust-aware and improved density peaks clustering algorithm for fast and secure models in wireless sensor networks,” *Pervasive Mob. Comput.*, vol. 105, Dec. 2024, doi: 10.1016/j.pmcj.2024.101993.

[34] X. Su, Y. Ren, Z. Cai, Y. Liang, and L. Guo, “A Q-Learning-Based Routing Approach for Energy Efficient Information Transmission in Wireless Sensor Network,” *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 1949–1961, Jun. 2023, doi: 10.1109/TNSM.2022.3218017.

[35] W. Zhao, Z. Zhang, S. Mirjalili, L. Wang, N. Khodadadi, and S. M. Mirjalili, “An effective multi-objective artificial hummingbird algorithm with dynamic elimination-based crowding distance for solving engineering design problems,” *Comput. Methods Appl. Mech. Eng.*, vol. 398, Aug. 2022, doi: 10.1016/j.cma.2022.115223.