



An intelligent fault detection approach based on reinforcement learning system in wireless sensor network

Tariq Mahmood^{1,2} · Jianqiang Li^{1,3} · Yan Pei⁴ · Faheem Akhtar⁵ · Suhail Ashfaq Butt² · Allah Ditta² · Sirajuddin Qureshi¹

Accepted: 12 July 2021 / Published online: 3 August 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The Internet of Things (IoT) has developed a well-defined infrastructure due to commercializing novel technologies. IoT networks enable smart devices to compile environmental information and transmit it to demanding users through an IoT gateway. The explosive increase of IoT users and sensors causes network bottlenecks, leading to significant energy depletion in IoT devices. The wireless network is a robust, empirically significant, and IoT layer based on progressive characteristics. The development of energy-efficient routing protocols for learning purposes is critical due to environmental volatility, unpredictability, and randomness in the wireless network's weight distribution. To achieve this critical need, learning-based routing systems are emerging as potential candidates due to their high degree of flexibility and accuracy. However, routing becomes more challenging in dynamic IoT networks due to the time-varying characteristics of link connections and access status. Hence, modern learning-based routing systems must be capable of adapting in real-time to network changes. This research presents an intelligent fault detection, energy-efficient, quality-of-service routing technique based on reinforcement learning to find the optimum route with the least amount of end-to-end latency. However, the cluster head selection is dependent on residual energy from the cluster nodes that reduce the entire network's existence. Consequently, it extends the network's lifetime, overcomes the data transmission's energy usage, and improves network robustness. The experimental results indicate that network efficiency has been successfully enhanced by fault-tolerance strategies that include highly trusted computing capabilities, thus decreasing the risk of network failure.

Keywords Energy efficient · Fault tolerant · Wireless sensor networks · Internet of things · Cluster head · Reinforcement learning

✉ Yan Pei
peiyang@u-aizu.ac.jp

Extended author information available on the last page of the article

1 Introduction

Internet of Things (IoT) has become a unique infrastructure for network traffic created on an individual basis by numerous small components and devices [1, 2]. IoT has achieved significant attention recently owing to the range of its implications, which include smart healthcare, smart home, traffic management, mobility, and new technologies. IoT methods have the potential to drastically improve the resilience, stability, and effectiveness of intelligent services. The primary aim of IoT is to collect information so that the presence of the most critical and vulnerable perceptual layer in the system is ensured at all times. Therefore, the perception layer is the most crucial and vulnerable because numerous sources are needed when the node energy is depleted. Additionally, the consistency of the data collected by the perception layer has aided in the adoption and maturity of IoT. Wireless sensor networks (WSNs) are extensively used in IoT-based systems to obtain the information needed by smart surroundings. Typically, a WSN is comprised of sensor nodes integrated with wireless transmission equipment. Sensor nodes are self-contained and widely dispersed. The architecture of a WSN is elaborated in (Fig. 1), which usually includes source sensor nodes, cluster head nodes, a sink node, and a manager node. These sensor nodes in operation can achieve real-time monitoring of the physical surroundings, provide extensive data for back-end services to evaluate, and develop the sensing layer's smart structure. However, the typical characteristics of these nodes usually are inadequate storage capacity, energy, and computing capabilities. Since wireless sensor networks aim to perceive and communicate data efficiently, designing the appropriate routing strategy will directly impact the network's

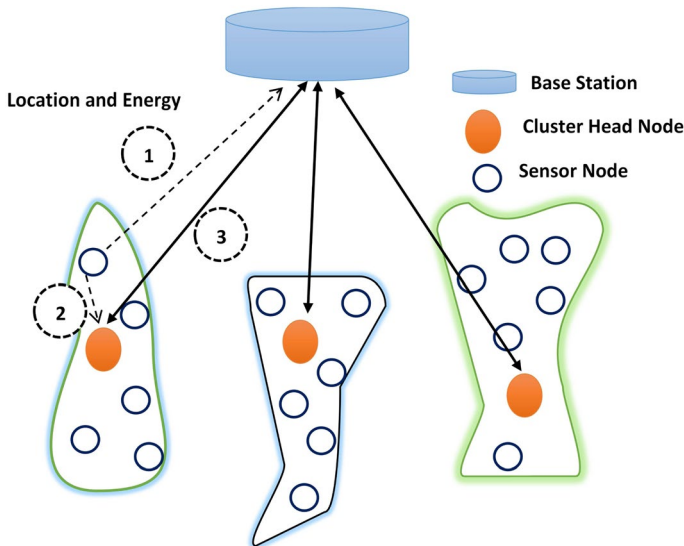


Fig. 1 Cluster-based architecture that allows end-users to monitor the environment remotely. The data collected by the different nodes are sent to a central base station through a cluster head node, which is usually located remotely inside the sensor network and enables access to the information to end-users

performance [3]. However, the primary purpose of the routing strategy for wireless sensor networks is to minimize energy dissipation, improve network damage stability, and extend network lifetime.

The routing strategy of wireless sensor networks has attracted the concern of several academics and institutions. The scale of IoT applications and services is increasing daily, resulting in expanded network architectural complexity. Due to WSN's limited processing and transmission capabilities, the network's architecture will be altered to include a suitable node-energy-saving technique. The main challenge is determining the most energy-efficient multi-hop routing scheme between source and destination nodes [4]. Consequently, technologies such as low-energy and rechargeable WSNs have been developed to ensure the long-term stability of WSNs and the IoT perception layer [5]. Implementing a low-energy routing protocol is significant for research because it enables the sustained functioning of WSNs and the knowledge layer of the IoT. Simultaneously, the amount of data that sensor nodes acquire and transmit varies, resulting in an unequal distribution of residual energy. Due to the limited amount of energy accessible to nodes, the routing protocol must prioritize the energy system to balance power between nodes while sacrificing QoS and other criteria to extend the network's life [6]. To detect information, the sensor node performs preliminary processing and relies on multi-hop transfer to the sink node. Similarly, the proposed algorithm is responsible for determining the order of the cluster head node. In addition to reducing computational complexity, communication overhead is also reduced, because nodes that are not involved in tracking will not send data. The wireless sensor network's transport layer protocol has a comparatively low burden [7].

Recent research indicates that multi-path routing is one of the most effective techniques for adapting to network topology modifications in WSN. Based on the efficient fault tolerance method, an enhanced LEACH (Low Energy Adaptive Clustering Hierarchy) [8, 9] routing method was developed by incorporating cluster member node failure fault and cluster head node failure. LEACH is a clustering routing protocol that periodically selects cluster heads with equal probabilities to avoid cluster head nodes due to excessive energy consumption. In LEACH, an irregular clustering arises from the random selection of the cluster heads. The transmission of data between the cluster head (CH) and base station (BS) is based on frames that limit the use of energy, thus having multiple stable frames reduces the overall energy cost of the network. Data fusion may be performed on the collected data through cluster head nodes, substantially reducing the number of data packets in the entire network. Implementing the LEACH protocol on the TinyOS platform was hindered by several technical problems that had to be overcome [10]. TinyOS has a more significant feature set than other WSN operating systems and is capable of running concurrent applications with low memory needs. TinyOS was used in this study because it is an open-source, customizable, component-based, and application-specific sensor network operating system. We demonstrated that the energy savings and optimum scheduling methods enable the operating system to perform optimally in capacity constraints situations [11, 12]. It can be seen from the analysis that algorithms have their limitations, so it is necessary to flexibly choose a routing protocol suitable for different situations according to various applications.

Therefore, the multipath routing approach often exceeds the processing capability of nodes and increases signaling overhead, prompting more research. To overcome these limitations, the software and hardware of the WSN must be very fault-tolerant to provide high reliability and robustness. The objective of fault-tolerant computing technology is to provide very reliable computer services that enable for the completion of scheduled operations.

Recently, machine learning has gained academic attention in network routing and may be used to develop the routing strategy for wireless sensor networks. Reinforcement learning (RL) may be a promising learning technique for finding the optimum routing path in real-time applications [13]. It is a technique of machine learning in which a learner referred to as an agent selects actions based on the present state of the environment through interaction with it. The agent performs in such a way as to maximize the long-term reward. However, in conventional routing algorithms, the node sends data through a pre-defined routing path. It does not correctly reflect the network's actual status because routing tables are built in advance. The RL-based system assigns a Q-value to each possible action, which represents the action's quality.

According to these limitations, this study aims to present a multi-path routing strategy for wireless sensor networks based on reinforcement learning for network life optimization [14, 15]. There is no need for prior knowledge of network topology. Nodes continuously learn, self-configure over time, and update their information to build intelligent route selections. However, the Q-Learning algorithm, which is based on reinforcement learning, determines the optimum route for the cluster head node. The Q-Learning route planning method initializes the Q-value by supplementing it with search heuristics. The cluster head node utilizes the auxiliary information to actively select the next route point, resolving the blind search problem early on. Additionally, increasing the number of nodes in a cluster increased the probability that more nodes will become cluster heads, leading to energy efficiency. The primary goal of the research is to decrease overall end-to-end (E2E) latency using the node energy consumption scheme before data transmission started. The presented algorithm dynamically adjusts the routing strategy based on the node's power consumption, avoiding data loss due to node failure and easing the load of data collecting under the premise of continuous operation. However, nodes that use multi-path routing protocol may learn quickly to find the optimum parent node-set iteratively. The development of the suggested energy-aware protocol has increased the long-term stability of WSNs, enabling the robustness of collecting data at the IoT perception layer. Consequently, the proposed study increased the network's lifetime by optimizing the cluster head selection technique and changing the selection thresholding and fault tolerance routing technology.

The rest of this paper is organized accordingly. A description of the related work, deficiencies in retrospective studies, and the scope of the study are discussed in Sect. 2. Section 3 describes the paper methodology and architecture of the proposed algorithms. Section 4 describes the performance study of the proposed protocol and numerical results. In Sect. 5, the simulation results are discussed regarding various performance parameters. Finally, Sect. 6 concludes the achievement and future directions of the research work.

2 Related studies

WSN has revolutionized human interaction with nature and enables a more immediate sense of the real world. The research on WSN has great scientific significance, and the application value has aroused great attention from academia, military departments, and industry in the world [16]. WSN routing is distinct from standard fixed network routing in that there is less infrastructure. In a WSN, the wireless connection is insecure, the sensor node may fail, and the routing protocol must also adhere to rigorous energy-saving constraints. Typically, WSN nodes have limited energy capability that cannot be replaced or recharged. However, to increase the life of WSNs, energy conservation considerations are addressed while constructing the network. The quality of the protocol design has a considerable impact on how well a WSN performs. Consequently, the presented study is based on an in-depth analysis of the LEACH protocol. A novel cluster-head voting mechanism was developed to optimize energy consumption and network life.

To overcome these challenges, the software and hardware of the WSN must be exceedingly fault-tolerant to provide high reliability and robustness. The goal of fault-tolerant computing technology is to provide highly reliable computing services that permit the completion of scheduled operations [17, 18]. The system automatically reconfigures the errors and ensures the normal execution of tasks to enhance the network's fault tolerance. The WSN routing protocols are responsible for transmitting data packets from the sensing node to the sink, determining the optimum route between the source and sink nodes [19, 20]. The data that ordinary nodes monitor and send to sink nodes include redundant information.

Many wireless routing approaches have been developed to minimize redundancy by using less bandwidth, storage space, and transmit power. LEACH [3] is a clustering routing protocol that periodically selects cluster heads with equal probabilities to avoid cluster head nodes due to excessive energy consumption. LEACH is divided into initializing phase and stable period phases. During the initiation phase, a cluster head is randomly picked and transmits information to its surroundings. During the stable period, the node gathers monitoring data continually and sends it to the cluster head for processing. The proposed study analyses various protocols used in WSNs, focusing on their data processing methods, route optimization, data transfer, data query caching, and network architecture [16, 21]. Data fusion may be performed on the collected data through cluster head nodes, substantially reducing the number of data packets in the whole network. It can be seen from the analysis that both the planar routing algorithm and the hierarchical routing algorithm have their advantages [22]. However, it also has its limitations, so it is necessary to flexibly choose a routing protocol suitable for different situations according to different applications. Loscri et al, [23] proposed two levels of hierarchal energy-efficient protocol. Two-Level Hierarchy LEACH (TL-LEACH) that uses the random rotation of local-cluster base-stations (primary cluster heads and secondary cluster heads) and localized the coordination to achieve robustness. TL-LEACH ensures the energy load in the network, particularly in a dense network. Flooding [24] is the most classic data transmission routing protocol

in sensor networks. Each node communicates with its neighbors through a data packet broadcast. After receiving the data, the neighbor node stores it and verifies that it was successfully delivered. The Gossiping algorithm [25] is an improvement to the problems in the Flooding algorithm. The node transmitting data only randomly selects one direction for data forwarding, and the forwarding direction is allowed to be reversed. SPIN protocol (Sensor Protocol for Information via Negotiation) [26] is the data-centric communication routing protocol proposed to solve the flooding based on negotiation and resource adaptation. Data transmission is done through a negotiation mechanism. When a node needs to transmit data, meta-data will be sent between the nodes for negotiation.

PEGASIS (Power-Efficient Gathering in Sensor Information Systems) [27] is a LEACH-based routing system with an effective cluster head selection. The sensor nodes are organized in a chain topology to reduce the high communication energy consumption associated with frequent selection. Each node determines the position of the different nodes and uses the greedy method to locate the closest neighbor nodes to transmit and collect data. Each node changes the cluster head in turn (chain head). The chain's information is transmitted from one node to another node, aggregated, and delivered toward the sink node, which initiates a new adoption cycle and communication cycle. The TEEN (Threshold sensitive Energy-Efficient sensor-Network protocol) [28], comparable to LEACH, uses a clustering approach, a routing solution tailored to reactive WSN. The TEEN technique establishes two hard and soft thresholds during the cluster creation phase, thus reducing the amount of data that passes through the filter. Two threshold parameters and data attenuation through TDMA must be broadcast after the selection of the cluster head. The hard threshold defines the smallest quantity of data to send, while the soft threshold provides the lowest possible extent of information changes to detect. Direct diffusion [29] is a planar routing protocol that completes data requests through queries. In a directed diffusion, the data acquisition request is broadcasted to the whole network through the sink node in interest messages. The node, according to the received interest messages, determines data forwarding. Younis et al. [30] proposed a hybrid energy-efficient distributed (HEED). The method selects the cluster head based on the remaining energy and distribution factors of the nodes. The cluster head selection utilizing the primary and secondary parameters. The primary parameter indicates how much energy remains in the node. Qiu et al. [31] proposed an energy-aware fault tolerance WSN algorithm, which decreasing the non-cluster head nodes and selects the target of data transmission. Liu et al. [32] proposed a fault-tolerant centralized algorithm based on structural health monitoring that disambiguates structural damage from sensor faults. Cheraghlou et al. 2014 [33] proposed a protocol that differentiates the live and faulty sensor nodes that enhance the network's fault tolerance capacity and avoids energy loss by eradicating rework. Jiang et al. [34] developed a method for scheduling packet transfers in intelligent IoT networks. The DQN model improved the system's throughput capacity by modifying two variables, including the energy required to transmit packets over several channels and packet discarding. They used the policy to optimize system performance using a stacked auto-encoder as a Q-function optimization algorithm and a utility-based incentive mechanism. Tang et al. [35] introduced a novel approach based on deep reinforcement learning for

dynamically distributing radio resources online in a heterogeneous wireless network with high motility.

In recent years, energy-harvesting technologies have been included in traditional WSNs to overcome the limitations of conventional WSNs that rely on energy resources with a limited capacity. The routing algorithm's primary goal is to reduce the energy consumption of WSN nodes, assure the stability and speed of the data transmission line, and optimize network throughput. By analyzing the advantages and limitations of conventional routing techniques applying our sensor concepts, we developed a novel clustering approach for reducing energy dissipation in sensor networks. This research aims to evaluate the energy conservation and fault-tolerance problems inherent in the classic LEACH-based technique. The random cluster head voting method in the proposed method achieves dynamic resource utilization; minimal consideration is required for the nodes' remaining energy level. Additionally, the presented technique can perform computation locally in each cluster, which reduces the quantity of data sent to the base station. This leads to a substantial reduction in energy usage, considering computation is much more energy-efficient than transmission.

2.1 Comparison analysis of existing approaches

This paper analyzes the characteristics of various protocols used in WSN and provides a comprehensive comparison of their information processing, routing optimization, data transmission, data query caching, data aggregation, and network architecture performance. The performance of several WSN protocols in terms of data processing, data fusion, scalability, route optimization, network topology, and energy efficiency is compared in (Table 1). The route optimization ability refers to the possibility of optimizing the route according to channel characteristics during selection. Multiple routes may pick one or more superior data transmission routes. The routing fault tolerance refers to the resilience and the dependability of the data to the destination using a routing algorithm. For example, if a sensor node fails, the routing algorithm may circumvent this point by repairing its route. When the channel error rate is significant in the system, or if the sensor node environment impacts signal transmission, it is possible to ensure reliable data transfer. The data transmitting technique refers to whether data packets are delivered by point-to-point transmission or through intermediary nodes directly into the sensor network. The period of network survival refers to how long the network can operate efficiently.

The data transmission method indicates whether the data are transmitted or forwarded by the query. The query cache is whether the routing algorithm establishes a buffered packet to hold a local copy of the data to be sent back to cache controls such as the query instructions received. An aggregation of data means that all data packets will be sent before being transported to the destination, to a specified intermediate node for processing, or the aggregation node of a particular node in an area. The data gathered are then fused through the cluster head node, reducing the quantity of data packets in the network to lengthen the life cycle of most nodes. The success of the WSN's MAC layer protocol is evaluated using indicators

Table 1 Performance comparison of different wireless sensor network protocols in terms of information processing, routing optimization, data transmission, data query caching, data aggregation, and network design

Routing algorithms	Network Topology	Fault Tolerance	Energy Efficiency	Data Fusion	Network lifetime	Scalability	Algorithms Complexity	MAC layering	Route optimization	Multipath routing
SPIN	Flat	+	-	+	-	-	+	✓	×	✓
Directed Diffusion	Flat	+	++	++	-	-	-	×	✓	✓
DSR	Flat	-	++	-	++	-	+	×	✓	✓
HEED	Hierarchical	+	+	+	+	+	+	✓	×	✓
LEACH	Hierarchical	++	-	++	+	-	-	✓	×	×
TL-LEACH	Hierarchical	++	+	+	++	+	-	✓	✓	✓
TEEN	Hierarchical	++	++	+	+	+	++	✓	✓	×
PEGASIS	Hierarchical	+	+	+	+	-	++	✓	✓	×

In the table, the symbols indicate X → Not Applicable, ✓ → Implemented, - → low, + → Moderate, and ++ → High

such as bandwidth requirements, energy consumption, bandwidth competition, and network connection. The network's MAC layer protocol's primary objective is to ensure latency and priority while reducing header overhead and energy consumption [36]. The MAC protocol, which is based on competition, conflict avoidance, or a mix of the two, may improve energy savings. This is due to a reduction in header overhead and conflicts, which result in a drop in network energy consumption. The typical network-wide clustering process also leads to substantial energy demands in a comprehensive clustering method such as LEACH. Transfer of data is completed through cluster communication, which decreases the flexibility of routing to some degree. It is not beneficial to boosting the overall energy efficiency of the network. From the previous research, it can be observed that both the planar routing method and the hierarchical routing method have their own benefits and limits. Therefore, routing protocols that are suited for diverse conditions in different applications should be freely selected.

2.2 Deficiencies in retrospective algorithms

The LEACH algorithm's basic aim is to select a cluster head that randomly selects and dispenses the energy load of the whole network to every node. By comparing with planar multi-hop and static cluster protocols, it is found that LEACH protocol can extend network lifetime more than 15%, and clustering can also optimize resource allocation. Still, there are many limitations in LEACH [3, 8].

- The LEACH algorithm randomly selects cluster heads, prone to unreasonable clustering, and distributed in network nodes.
- The LEACH protocol uses long-distance data transmission to communicate directly between the cluster head and base stations. The massive amount of energy is consumed in the cluster head node while considering multi-hop routing within the cluster head to the base station.
- Establish and operate clusters through local collaboration and control.
- The cluster's topology varies dynamically, and cluster head nodes are selected randomly.
- Local data fusion is used to minimize the overall communication burden.

Because wireless communication's energy consumption is proportional to the transmission distance, the energy consumption of nodes at various locations varies, resulting in an unequal energy distribution across nodes. All nodes must operate concurrently in particular applications, and the system's appropriate operating time is determined by the node with the highest energy usage.

2.3 Scope and objective

The initial analysis in this paper is the state of WSN's development and its obstacles leading to practical uses of routing algorithms in WSN. The present study focuses

on the reliability of routing protocols, the characteristics of WSN, and the reinforcement learning technology used to work with the best:

- The greedy method optimizes the selection of the head node by prioritizing adopting the current optimum path. The node explores other nodes to prevent getting a single choice and allocates a probability of exploring a new route, enabling the data transmission route to accumulate progressively. Data forwarding, time delay, energy consumption, and energy harvesting models all are contributing to optimal results. Finally, the algorithm's efficacy was validated by simulation and comparison studies.
- Due to the unpredictable dynamics of WSNs, a central controller is included in the model, and a main real-time algorithm based on reinforcement learning is proposed. This approach incorporates a centralized controller into the network, which monitors the network's topology and creates real-time routing methods. The programme may dynamically alter the transmission channel based on node energy variations, preventing data loss due to node power failure. The central algorithm no longer selects the node independently for each hop. Path selection, on the other hand, occurs during the initial stages of data transmission. The node information is updated until the data transmission is complete, at which point it is reactivated during the subsequent learning cycle.

Although the widely used clustering procedure can balance energy usage, it does not reach optimum usage. This study aims to balance the node's energy usage and proposes a technique for delivering data with the lowest hop count, which ultimately minimizes the data transmission's power usage and removes redundant information through data fusion. The simulation results demonstrate that the suggested technology can significantly lower energy consumption and extend the lifespan of wireless sensor networks.

3 Architecture of proposed routing algorithm

The TL-LEACH and LEACH algorithms are used to mitigate the high energy consumption coupled with prolonged communication between cluster nodes caused by unconstrained clustering. The fixed cluster number determination pertains to the LEACH algorithm's examination of the ideal cluster number. However, the cluster head node is selected randomly, shows instability in the number and distribution of cluster heads state. So when the frequency of cluster heads is inadequate, layering becomes ineffective. The cluster head is linked with the distant base station across many cluster heads, leading to the whole network's excessive energy consumption. Due to the inadequacies of the TL-LEACH and LEACH routing protocols, this study proposed a LEACH-EFT (Energy-Aware and Fault-Tolerant LEACH Routing Protocol). During the clustering phase, the LEACH-EFT algorithm modifies the cluster head election depending on thresholding. To extend the network's lifespan, fault-tolerant routing methods are used. The cluster head and non-cluster head node

failure occurred in each round to enhance network fault tolerance. The proposed protocol is designed on the concept of selecting the cluster head and segmenting each round into a cluster establishment phase and a transmission phase.

The schematic diagram of the proposed protocol related to each round is shown in (Fig. 2). Unlike the TL-LEACH and LEACH routing protocols, the LEACH-EFT protocol identifies the cluster head relying on remaining energy during the cluster establishment stage. The time division multiple access (TDMA) table threshold value is applied to the scheduling of each cluster member node. The main cluster head incoming feedback from cluster member nodes from the many clusters. When a cluster member sends a full packet of data, the cluster head compresses and sends it to the base station using data fusion techniques. However, when cluster nodes fail, the proposed protocol uses a fault-tolerant approach to improve network fault tolerance for both cluster head and non-cluster head node failures. The LEACH-EFT protocol's key technologies are to design the algorithm for the optimal selection of a cluster head, as summarized in (Fig. 3).

3.1 Reinforcement learning

Reinforcement learning (RL), a significant subfield of machine learning, is concerned with performing action learning in response to the environment to maximize the expected outcome [37, 38]. The standard Q-learning algorithm based on the decision process from Markov has no previous knowledge of the environment,

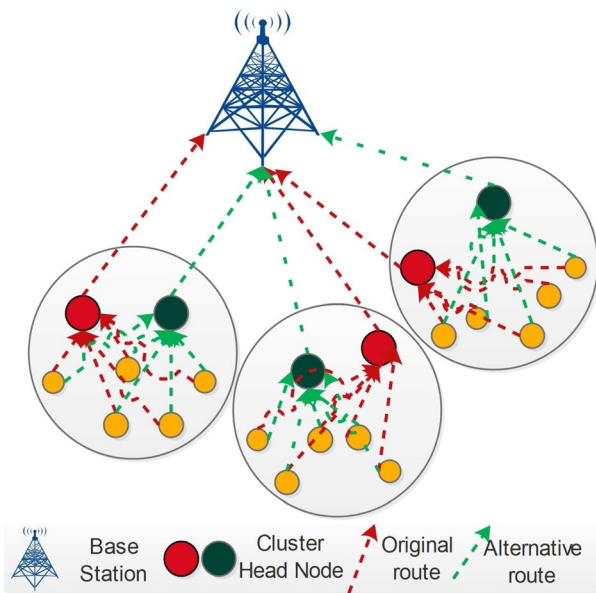


Fig. 2 Schematic diagram of LEACH-EFT protocol is composed of primary cluster head nodes (red), standby cluster head nodes (green), ordinary nodes and base station. The fault-tolerant method is used to improve network fault tolerance for both cluster head and non-cluster head node failures

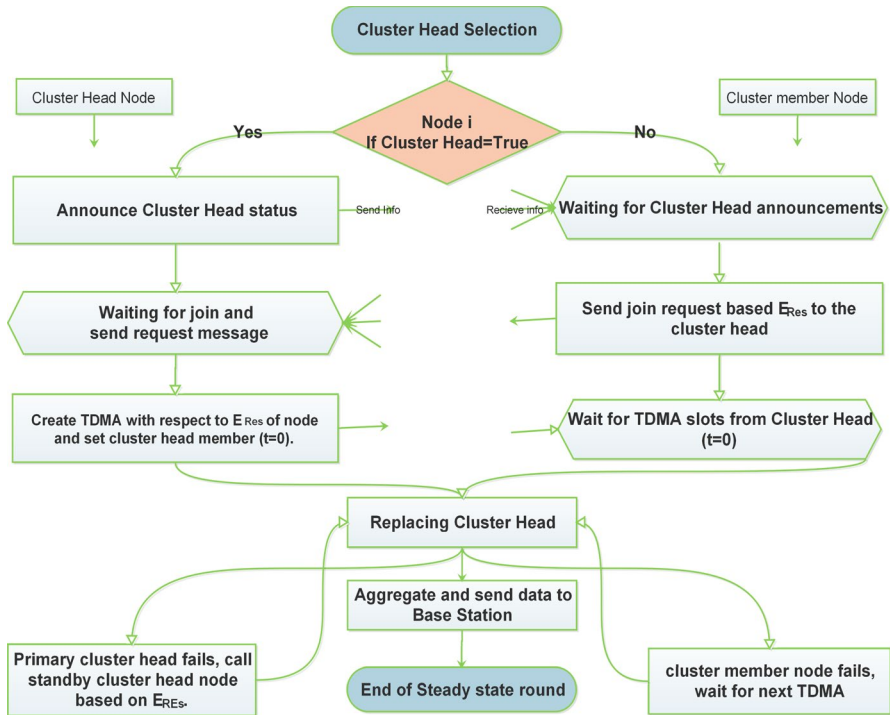


Fig. 3 Schematic diagram of the workflow of each round of the LEACH-EFT protocol. The architecture selects cluster head member nodes and cluster head nodes using a fault-tolerant approach coupled with TDMA Scheduling

leading to slow training speed and low iteration efficiency throughout the learning process as regards path planning optimization. Finally, better route planning advantages may be achieved with shorter training periods suggesting that the increased technology converges to optimize the learning process more rapidly and effectively [39, 40]. At the same time, it is feasible to avoid repeating "test and mistake" on obstacles by starting the technique early on in the learning process. In brief training, enhanced route planning advantages may be achieved, indicating that the improved methodology converges faster and more effectively to optimize the learning process.

In the simulation experiment, the Q-Learning algorithm is upgraded to optimize route planning. Designing the optimal route entails getting sensor data on the unknown or partly known surroundings, including data on obstacles, using that data to locate, and identifying possible collision-free travel optimal and sub-optimal paths [41]. The nodes arrange their local route autonomously in real-time based on the computed information about the surrounding area at their present position. On the other hand, adaptation needs are increasing as a result of interferences such as complex variety and unknown elements in the natural environment. Different investigations have conducted related development studies to

address the cluster head steering issue in uncertain and changeable settings; however, related development approaches continue to have certain shortcomings.

In this study, an optimum route planning approach based on reinforcement learning is proposed. The Q-Learning algorithm, which is based on reinforcement learning, determines the ideal route for the cluster head node in complicated obstacle situations. Q-Learning is a well-known model-independent reinforcement learning algorithm [42]. When the cluster head node has a vast feature space, path optimization algorithms demand a considerable storage capacity. The research started with the Q-Learning algorithm, which has a poor learning efficiency and a sluggish convergence rate. It then builds on existing information to inspire the concept of guided search. The knowledge-based guidance proposed in this paper is based on an artificial potential energy field generated using the conventional Q-Learning method. The Q-Learning route planning technique initializes the Q-value by adding search heuristic information previous knowledge of the environmental state space. The cluster head node takes advantage of the auxiliary information to actively choose the next route point to resolve the blind search issue early on. However, it rapidly converges during the early stages of learning and significantly enhances the algorithm's learning efficiency and convergence speed.

3.1.1 Q-Learning

The proposed method is based on the Q-Learning algorithm, a popular technique in RL that uses Markov Decision Processes (MDPs) to detect learning problems by choosing the optimum energy-efficient routing path [43]. The core of reinforcement learning is how to attain objectives by interacting with the environment. The analytical composition of MDP consists of a tuple of $[M = S_t, A_t, P_a(s_t, s'_t), R_t(\text{reward}), \gamma]$, where $s_t \in S_t$ is finite set of state among all possible states and S_t represent the state collection at epoch t . $a_t \in A_t$ indicates a specific action among all actions and A_t represents a collection of actions. $P_{a_t}(s_t, s'_t)$ is the transition probability from state s_t to state s'_t under the action a_t . R_t is reward value of A_t action selection at epoch t depending on state s_t . $0 \leq \gamma \leq 1$ is the discount factor that indicates the degree of the agent's far-sighted future. Calculates the feedback reward value γ based on the agent's current state s_t and action a_t to completely characterize the reinforcement learning process. For each state s_t , strategy representation $\pi(s_t)$ to select the action a_t .

Every agent learns information and retains every action a_t , known as Q-value $Q_t(a_t)$, that correlates to the reward for determining the action a_t in at an epoch. Cluster head nodes execute action among all possible actions in states s_t that evaluate the result based on the immediate reward value of action a_t and the current state's estimation.

$$Q_t(a'_t) \leftarrow (1 - \alpha) \times Q_t(a_t) + \alpha \times [R_t(a_t + 1) + \gamma \times \max Q_t(a'_t)] \quad (1)$$

In (Eq. (1)), $0 \leq \alpha \leq 1$ indicate the learning rate, $0 \leq \gamma \leq 1$ is discount factor, and R_t is the current reward value for choosing an action at t , which varies according to the state at the time $t + 1$. $\max Q_t(a_t)$ is a maximum value of Q_t contained in the next epoch t . After reciprocating all actions, the cluster head learns the overall best behavior, and the value of the $Q_t(a_t)$ function will eventually converge to the optimal value is shown (Eq. (2)).

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha \times [R_{t+1} + \gamma \times \max Q_t^\pi \times (s'_t, a'_t) - Q(s'_t, a'_t)] \tag{2}$$

In our RL-based multi-hop routing model, each source node executes an action based on stored Q-values using a ϵ -greedy approach. It sends data to the selected action's next-hop(s) parent. The node transmits data along the activity's route. However, each source node computes and updates the action's Q-value using the action's nodes' remaining power and E2E latency during transmission.

Agent: In this proposed study, the agent is the object of learning; it keeps track of all dispersion nodes and potential transmission routes. The agent will continue to interact with its surroundings, selecting the right action depending on its current state, resulting in a new state and reward value for the agent. Each sensor node is treated as an autonomous agent that distributively learns from the surroundings to send data packets.

State: is described as a collection of states $S = \{s_1, s_2, \dots, s_i, \dots, s_j\}$ associated with each sensor node in the network. When a packet is transmitted from node n_i to node n_j , the node's status changes from s_i to s_j .

Actions: $A = \{a_1, a_2, \dots, a_i\}$, denotes a set of exploratory activities, where action a_i symbolizes the node n_i being selected as the next hop forwarder. Each node i is considered a self-contained entity with set of parent N_i . A node may choose to hop through one or more parent nodes. Each node selects actions using the ϵ -greedy method; before each node makes a selection, a random probability value $p \in [0, 1]$ is created. If $p \geq \epsilon$, node i choose action with the most significant $Q(a_t)$ value in the time slot t ; if $p \leq \epsilon$, node i will randomly choose an action.

Reward Function: The reward is the current state impact of the action. In the proposed study, the reward is determined by energy usage and data transfer latency. Every node determines the reward for an action using the node information transmitted by its parent node. Node i transmits a tuple of information comprising energy and data to neighboring nodes along its selected path during each time interval t . Initially, each nodes i computes the estimated remaining energy EE of parent $k \in N_i$ in time slot t , to forward information as illustrated in Eq. (3).

$$EE_k^t = E_k^t - dt_k^t \times E_{cost}^t \tag{3}$$

where E_k^t represents the node available energy, dt_k^t depicts the volume of data in the node buffer, E_{cost}^t represents the energy cost of sending a unit of data in a time slot t . The variables calculation entail the number of time slots required for data transfer at the ending stage of a learning process. The (Eq. (4)) represents $\min E(n)$ the minimum remaining energy of a node and $\sum E(n)$ remaining energy of the intermediate

nodes in the path p_{th} . The designed algorithm is intended to choose the route with the highest power weight, which has a more significant effect on path discovery and avoids attempting the path with a low power node. An action can contain single or multiple paths, the reward value R_a^t of the action a in time slot t containing path is defined as mathematically.

$$R_a^t = \frac{\sum(\min E(n) + \sum E(n))}{\sum n(p_{th})} - t \quad n \subseteq p_{th} \quad (4)$$

(**Algorithm 1**) shows that all sensor nodes are regarded as agents, and every agent's action will generate an $Q(a_t)$. There are single or multiple transmission paths from the source node to the sink node, and every optional path of the sensor node is combined and generated without repetition. At the beginning of the study, the agent uses *egreedy* to choose an action [44]. An operation entails numerous paths, and each sensor node selects single or multiple paths for data transmission. When the transmission is initiated, the sensor node only forward the information in proportion to the path in action. It computes the time slot used for whole data transmission toward the sink node, which is the *E2E* delay. A standard single-step Q-Learning algorithm is in explained in (**Algorithm 1**).

In this proposed multi-path routing based on Q-Learning, every source node choose an action based on stored *Q-value* through the ϵ greedy strategy. Then, the source node forward the data to the next single or multiple neighboring nodes in the exclusive action. During the transmission process, every source node computes and updates the energy and data information in the selected path in the time slot t . However, On the completion of the data transmission process, every sensor node computes and updates the action's *Q-value* based on the remaining energy and *E2E* delay of this learning. In the proposed algorithm, every sensor node is regarded as a whole as an agent. The source node is randomly combined into multiple non-repeated sets before the selection of the whole path. The agent computes every action's reward based on node information compared to the number of time slots needed for data transmission, computation, and updating.

Algorithm 1 Multi-hop routing algorithm based on Q-learning

Initialize $Q_t(s_t, a_t)$ for each node
 Repeat (for each wheel) Initialize s_t
 Repeat (for each step of wheel)
 Select a_t for s_t according to the approach
 Take action using ϵ -greedy according to $Q_t(s_t)$
 Use actions a to transfer data
 $reward \leftarrow R_t(s_t, a_t)$
 $Q_t(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha \times [reward + \gamma \times \max_{a'} Q_t(s'_t, a'_t) - Q_t(s_t, a_t)]$
 for all s_t belong to S_t , a_t belong to A_t
 $a_t \leftarrow a'_t$
 $s_t \leftarrow s'_t$
 Calculate the time slot t used to transmit data
 Update node information
 Until s_t is terminal

3.1.2 Selection of cluster head

The critical functions of cluster head selection algorithms are to elect and divide the scheduling order of each cluster member. The LEACH methodology selects the cluster head based on the (Eq. (5)). The cluster head is chosen randomly, but the sensor node collected and communicated different data in the network as a result of large monitoring activities. However, once a node with low energy is identified as a cluster head, it consumes rapidly and fails to communicate in the network. Regarding the key shortcomings, the node’s residual energy is an important factor to consider while selecting the cluster head node. If a node’s residual energy is sufficient, threshold $Th_{(n)}$ values improve the cluster head’s selection probability [3, 45]. If the node energy is low, threshold $Th_{(n)}$ value reduces the cluster head’s selection probability, which effectively extends the lifetime and load balancing of the network. This paper mainly considers the residual energy of nodes by modifying the value of threshold $Th_{(n)}$ as declare in (Eq. (5) and (6)).

$$Th_{(n)} = \left\{ \left(\frac{Prob(p)}{1 - Prob \times (rmod \times (\frac{1}{Prob}))} \right) \quad n \in G \right\} \tag{5}$$

In (Eq. (5)) $Prob(p)$ is percentage (probability) of cluster head nodes, r is the current number of rounds in-network, G is the last $1/p$ round. The node energy increases when the threshold $Th_{(n)}$ is increased. In (Eq. (6)), the $d_{crossover}$ is used to calculate the shortest distance from the base station.

$$d_{crossover} = \frac{4 \times \pi \times \sqrt{L} \times hr \times ht}{\lambda}, \quad (6)$$

where L is transmission loss, hr and ht is the receiving and transmitting antenna heights, respectively, and λ is the wireless signal's wavelength.

Each primary cluster head broadcasts hello packets containing the node identification number throughout the network when the cluster head is elected. The non-cluster-head node received the hello message according to the signal strength of the received message. When a member node identifies which node is the cluster head, it notifies the cluster head node through a request message. The distinction between LEACH-EFT and LEACH is in the request message, including the member node and cluster head node identification numbers. LEACH-EFT protocol includes information on the member node's remaining energy in the request message. The primary cluster head node will specify a cluster head node in a request message based on the remaining energy information of the nodes. The cluster head node has a high or low residual energy value, depending on the value specified in the request message. Sorting the cluster members' identification numbers and creating a TDMA Schedule table. The TDMA Schedule database stores the identification number of each element and the cluster member node's status information. As the standby cluster head node, the primary cluster head node will default to the first node in the TDMA Schedule table. When the primary cluster head node fails, the standby cluster head node is enabled. The primary cluster head assigns a time slot to each member node using the TDMA Schedule table. Finally, the cluster's member nodes send data in the sequence specified in the TDMA Schedule table.

3.1.3 Selection of cluster head member node

The process of replacing the failing cluster head node with the standby cluster head node encounters two difficulties [46, 47]. During cluster head election, the main concern of the cluster head node is obtaining information about the major group head section. In the instance of the primary cluster head node failing, the standby cluster head node constantly seeks for and replaces the primary cluster head node. The substitution light cluster head is moved to the primary cluster head position to replace the current primary cluster head node. The suggested procedure spares the cluster head, generated by resident cluster members who remain active during time slots and sleep during non-active time slots to conserve energy.

$$Th_{(n)} = \left\{ \frac{Prob(p)}{1 - Prob \times (rmod \times (\frac{1}{Prob}))} \times \frac{E_{res} \times N_i}{E_{ini} \times N_i} \right\} \quad (7)$$

In (Eq. (7)) $Prob(p)$ is percentage ((probability)) of cluster head nodes, r is the current number of rounds in-network, Rnd is the last $1/p$ round, E_{res} indicate the residual energy of node N_i , and E_{ini} is starting energy of node N_i . The node energy usage increases when the threshold $Th_{(n)}$ is increased. The suggested study resolves the issue by determining the value of the threshold $Th_{(nd)}$, which signifies the lowest amount of energy required for a packet to reach the base station. If the energy of the

primary cluster head node is less than the value of the cluster member's threshold $Th_{(hd)}$, the primary cluster head node is regarded to be about to fail.

Meanwhile, the primary cluster head alerts the backup cluster head that it is time for a replacement [8, 47]. If the base station drifts away from the sensor node, the threshold value $Th_{(hd)}$ remains unchanged and is identical to the data packet received. The specific value of $Th_{(hd)}$ can be obtained from (Eq. (7)).

$$T_{(hd)} = k \times E_{req} + \epsilon_{amp} \times k \times d_{to-BS}^2 \tag{8}$$

In (Eq. (8)), k message size, represents the number of bits contained in a data packet, E_{req} energy consumed in forwarding/receiving of each bit of data, ϵ_{amp} the energy

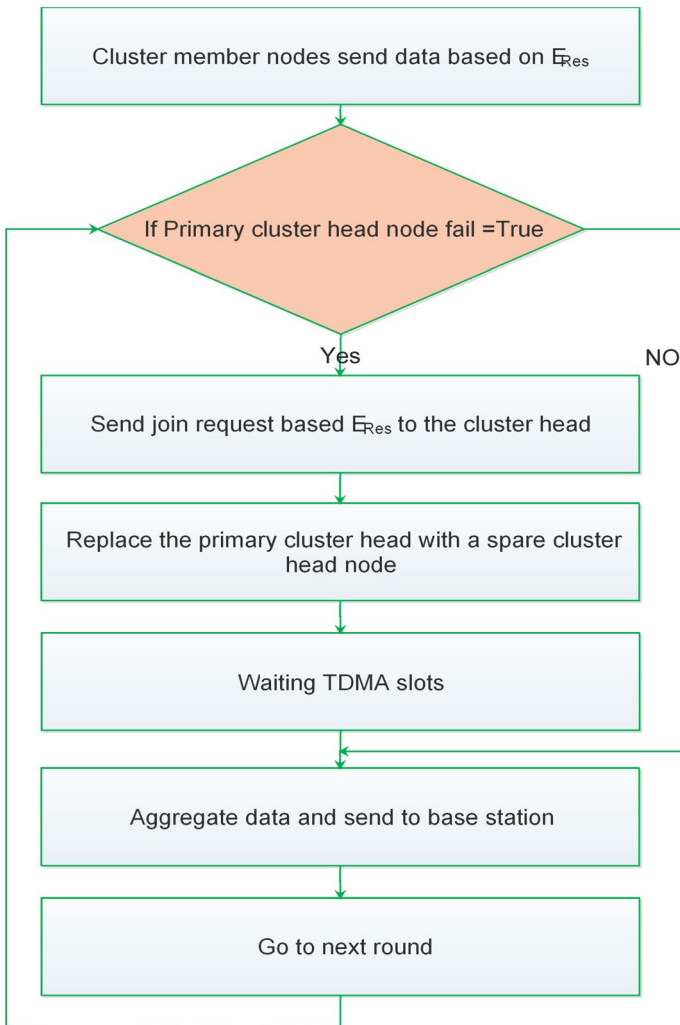


Fig. 4 Workflow of the cluster head node. The cluster head fault-tolerant algorithm is used to replace a non-functional primary cluster head node with a spare (standby) cluster head node

required to amplify the transmission signal per bit of data per unit area. d_{io-BS} is the distance from the area's geometric centre where all wireless nodes are located near the base station (BS). This paper proposed an algorithm that includes the following steps to overcome the facing challenges:

- Step 1: The primary cluster head maintains the TDMA Schedule table coupled with node residual energy and transfers it to the spare cluster head, where it is modified.
- Step 2: The main cluster head was acknowledged for data transmission and operates as a backup for fusion processing.
- Step 3: In this cluster, the spare cluster head takes over as the cluster head.

The backup cluster head node uses the TDMA Schedule table and broadcasts new scheduling information to all cluster members. Each cluster member collecting data by the updated TDMA Schedule table and transmits it to the base station. The LEACH-EFT method's cluster head fault-tolerant algorithm is illustrated in (Fig. 4).

3.1.4 Non-cluster head node selection method

This method evaluates and analyses the non-cluster head node of LEACH-EFT Fault-tolerant algorithm. Non-cluster head node failure is not targeted for fault tolerance in the TL-LEACH and LEACH algorithms. The frequent failures of non-cluster-head nodes in traditional techniques caused by each node's energy consumption over time result in increased losses. The LEACH-EFT protocol proposes a fault tolerance algorithm for non-cluster head nodes' failure includes the following steps [48]:

- Step 1: The cluster head node TDMA Schedule database contains information on failed non-cluster head nodes.
- Step 2: After configuring the TDMA Schedule table of cluster head nodes, the failed non-cluster head nodes were removed, and the new members were allocated cluster time slots.
- Step 3: The revised scheduling order was assigned to cluster member nodes, and nodes transmit collected data to the cluster head following the revised schedule.

By comparing the non-cluster head nodes with TL-LEACH and LEACH routing protocols, it is observed that the LEACH-EFT protocol improves the efficiency of the cluster head node. To update the TDMA Schedule table, a small amount of computation cost is required, and transmission cost is utilized to broadcast a new TDMA Schedule table. Consequently, the cost decreases the efficiency of data transmission within the cluster and increases the network's fault tolerance performance.

4 Performance analysis of proposed schemes

LEACH-EFT protocol's performance is analyzed and evaluated compared to the TL-LEACH and LEACH routing protocols in the same environment setting. The base station is located a significant distance from the sensor node and is immobile, using a considerable amount of energy. All sensor nodes in the network have the same architecture and initial energy. The transmission power of the sensor node is adjustable and proportional to the transmission distance.

4.1 Simulation metrics

A wireless sensor network's performance directly impacts its availability, making it a critical topic that demands in-depth investigation. Various metrics are utilized to validate the performance of the LEACH-EFT, TL-LEACH, and LEACH routing protocols in terms of the network life cycle, benefits, and drawbacks [49, 50].

- Energy efficiency: often known as energy savings, refers to the number of requests a network can serve with a limited amount of energy.
- Life Cycle: A WSN life cycle is defined as the time necessary to produce needed information. Aspects impacting the life cycle of a wireless sensor network are including both hardware and software.
- Delay Time: When an observer delivers a request to its recipient, the delay time of a WSN is measured. The time necessary for the answer message to be received. Numerous variables influence the time delay of WSN. Inter-delay time is intrinsically linked to the application and directly impacts the applicability and breadth of wireless.
- Perception accuracy: The perception accuracy of wireless sensor networks pertains to the perception information acquired by observers. Sensor accuracy, information processing techniques, and network communication protocols have all impact on sensor perception accuracy.
- Scalability: WSN scalability is represented in the number of sensor nodes and network coverage, as well as scalable domain limitations, life cycle, time delay, and perception accuracy.
- Fault tolerance: Sensors in WSN often fail to owe to environmental factors, lack of power, and other factors. This necessitates that the sensor network's software and hardware be incredibly fault-tolerant.

When the network's software and hardware failure, the system automatically rebuild to fix mistakes and assure the network's regular functioning. The performance indicators are not just standards for assessing sensor networks, but they are also standards for wireless transmission. A significant amount of research work is performed to attain the predicted objectives.

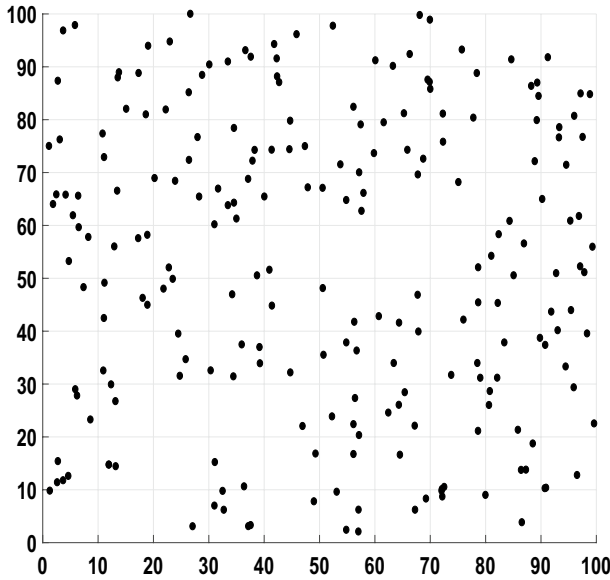


Fig. 5 Wireless sensor nodes are distributed randomly over the simulation area to monitor the environment

Table 2 Different simulation parameters are used to evaluate the effectiveness of proposed routing protocols regarding network life cycle, energy consumption, E2E delay, and packet delivery ratio

Network parameter	Primary values
Simulation area	$(0, 0) \times (100, 100)$
Base station coordinates (BS_x, BS_y)	(50, 175)
Node deployment	Random
Packet size	500 Bytes
Number of sensor nodes	100
Number of rounds	3000
ϵ_{amp}	50 nJ/bit
$d_{crossed}$	87 m
Packet header size	25 Bytes
Node initial energy	2 Joule
Energy consumption rate	50 nJ/bit
Simulation Duration	Variable
Network bandwidth	1 Mbps
Signal amplifier magnification	10 pJ/bit/m ²
Number of divided clusters	5
Data fusion energy consumption rate	5 nJ/bit

4.2 Simulation environment

The (Fig. 5) shows that the 100 sensor nodes are sparsely randomly distributed in a $100m \times 100m$ region with their coordinates ranging from (0, 0) to (100, 100),

and the base station is situated at (50, 175) using the MATLAB simulator [50]. The locations are relatively close, and the comparable nodes have deemed a cluster with a decent clustering effect. Sensor nodes placed randomly across the network and base station (50, 175) are represented in the figure. The network capacity is 1 Mbps, and data packets are 500 bytes in size. The details of the parameters used in the simulation are illustrated in (Table 2). The LEACH-EFT protocol primarily simulates the cluster head election procedure and non-cluster head node failure. The significant cluster head section is chosen and broadcast through the function application to advertise cluster head through simulation of (Eq. (8)).

Each non-cluster head node invokes the application function to select the optimal cluster and strongest signal based on the received signal from the primary cluster head. The primary cluster head node was chosen by the algorithm that sorts the cluster's member nodes based on remaining energy and selects the backup cluster head node. The function application is executed on the primary cluster head node. It generates the TDMA Schedule database and schedules for each cluster member node. A fault-tolerant method for non-cluster head node failure in function application is created during non-cluster head failure. According to the TDMA Schedule table, the primary cluster head calls this function. The data were captured by obtaining the ID numbers of all unsuccessful non-cluster head nodes and calling the function.

Finally, the function application must build a timetable that reschedules the cluster's member nodes. When the cluster head node fails, the function application confirms the failure by analyzing the cluster head alive responses. The node executes the application function to update and create a new TDMA Schedule table. Cluster members restart data collection inside the cluster by the revised scheduling sequence and then send it to the base stations.

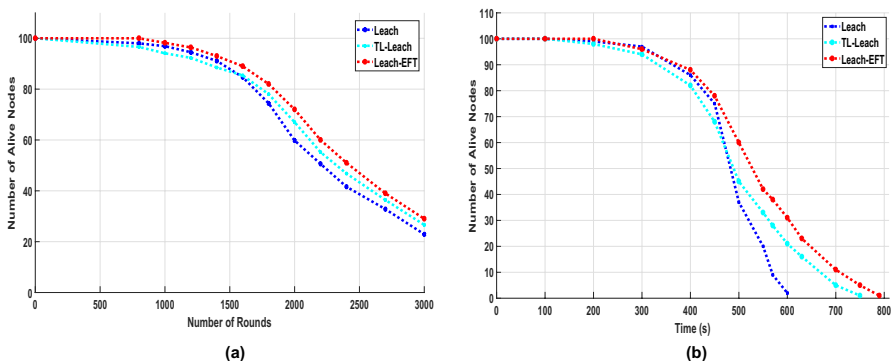


Fig. 6 Figure (a) and (b) demonstrate the total number of living nodes for each of the three approaches. It is observed that the suggested approach is proficient in extending the network's life and outperforms existing protocols in terms of performance. The main reason for current protocols' poor performance is that cluster heads are selected randomly and ad-hoc, leaving vulnerable areas of the network unprotected

4.3 Performance evaluation state-of-the-art routing protocols

According to the simulation function, LEACH-EFT, TL-LEACH, and LEACH protocols have been simulated 20 times based on the network lifetime, real-time data packets, and real-time throughput. In inappropriate clustering, the impact is more apparent in areas of the network with an unequal distribution of nodes. The cluster head in the network chooses the node with the most energy as the cluster head. The simulation results provide a real-time comparison of data delivered through the LEACH-EFT, TL-LEACH, and LEACH protocols.

The (Fig. 6a and b) illustrate the number of alive nodes for each of the three protocols LEACH, TL-LEACH, and LEACH-EFT, respectively. This parameter is essential in determining the network's longevity and the frequency of node mortality. The presented algorithm can prolong the life of the network and outperform current protocols in terms of performance. The suggested method has a network lifetime of 789 seconds, while the TL-LEACH protocol has a network lifetime of 761 and LEACH has a network lifetime of 613 seconds. The findings indicate that the LEACH-EFT technique has a 1.4% longer lifespan than the TL-LEACH and LEACH protocols. After 400 seconds, the number of surviving nodes in the LEACH-EFT protocol is significantly greater than TL-LEACH and LEACH protocol. Under the same number of rounds, the LEACH-EFT algorithm retained a more significant number of surviving nodes than the LEACH method. The prime reason for the LEACH and LEACH-TL protocols' poor performance is that nodes so imminent to one another are elected as cluster heads, resulting in an uneven distribution of cluster heads that may not cover the whole network.

The (Fig. 7a) depicts the entire amount of data signal collected at the BS over time. However, the LEACH-EFT protocol transmits a massive amount of data to the base station than the conventional LEACH, resulting in a reduced latency. This

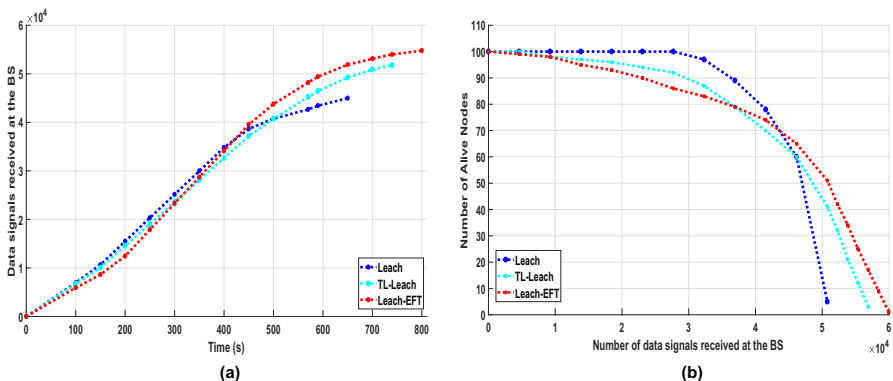


Fig. 7 Figures show the network life cycle of the proposed algorithm in contrast to that of existing algorithms. **a** The findings indicate that the LEACH-EFT protocol acquired between 10% more data packets than the other protocols at the base station. The reduced packet loss rate shows the enhanced fault tolerance of the LEACH-EFT protocol when a network node fails. **b** The number of surviving nodes in the network reduces as the number of rounds increases. The suggested design provides a more extended network lifetime than existing approaches due to the increased number of alive nodes

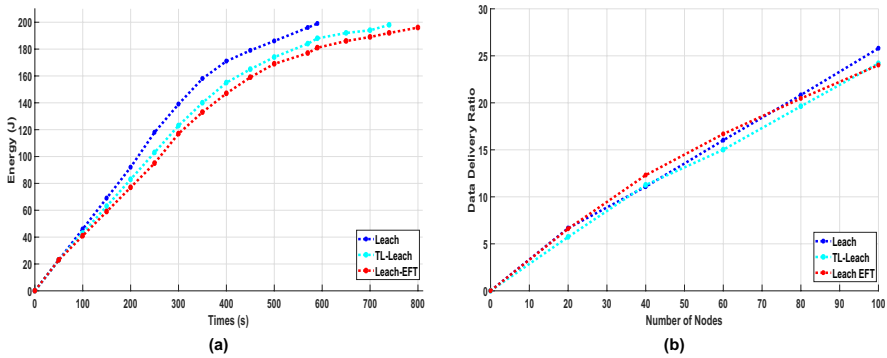


Fig. 8 Figures illustrate **a** the cumulative energy used by the nodes as the simulation proceeds. The proposed system consumes less energy than existing systems due to the energy estimation of the cluster head node and cluster member node. The suggested technique is found to be capable of increasing the network lifespan by up to 800 s. **b** describes the suggested method's performance in terms of packet delivery ratio. The proposed techniques transfer packets faster than existing approaches owing to the adoption of the TDMA schedule

outcome may be explained by implementing an adequate TDMA scheduling policy, which allowed for increased data transmission frequency. The (Fig. 7b) demonstrates the distribution of alive nodes based on data received at the base station. As the number of rounds increases, the number of remaining surviving nodes in the network decreases. The results reveal that the LEACH-EFT protocol obtained 5% – 10% more data packets at the base station and a 20% increase in network lifespan than other approaches. The decreased packet loss rate demonstrates the LEACH-EFT protocol's improved fault tolerance when a network node fails.

The figures (Fig. 8a) illustrate the network's energy consumption pattern throughout the simulation period. Power efficiency and network life's longevity are critical considerations. Every node is equipped with 2 Joules energy ($2 \text{ joule/node} \times 100 \text{ nodes} = 200 \text{ Joule}$) in the simulation environment. The energy of nodes reduces as the simulation proceeds owing to sensing, data transfer, and communication with different nodes. The suggested technique is found to be capable of increasing the network lifespan by up to 800s. It demonstrates clearly that the suggested method's energy consumption pattern is more effective than other variations. The high energy consumption of TL-LEACH and the LEACH protocol reduces the network's lifetime, which is improved in LEACH-EFT. Initially, node failure, particularly the failure of cluster head nodes in the network, leads to significant packet loss. Analyses were carried in a dynamic environment, and the total remaining energy consumption of the remaining nodes and the network modified in a round-by-round approach. The cluster-head selection based on thresholds conserves energy and balances the network's load, prolonging the network's life.

However, in (Fig. 8b), the network throughput performance of LEACH-EFT is larger than TL-LEACH and LEACH protocols after 400 seconds. The simulation demonstrates that the LEACH-EFT network throughput is about 8% greater than current protocols. It demonstrates that the LEACH-EFT protocol is incapable of fully sustaining consistent throughput before a network node fails. Routing

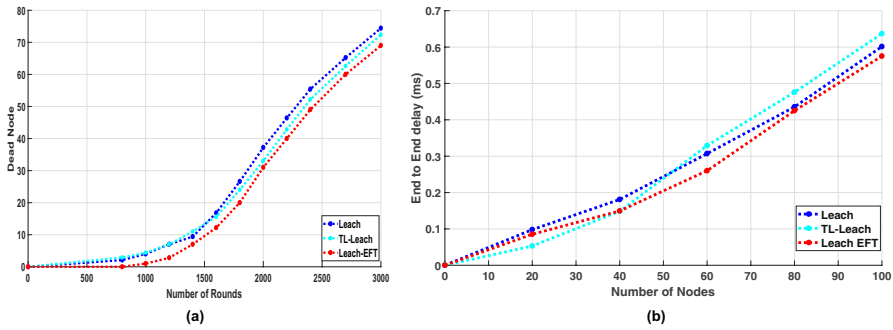


Fig. 9 Figure illustrates the number of dead nodes and E2E latency for both traditional and novel method. **a** Dead Node per number of rounds, **b** Due to the TDMA scheduling and optimal selection of cluster heads, the proposed approach shown minimal latency

Table 3 Suggested model’s performance is compared against state-of-the-art techniques using a variety of parameters

Parameters	Leach	TL-Leach	Leach-EFT
Network life time in second	613	761	789
First node dead at round	730	761	789
Dead node at 3000 round	73	69	64
Live node at 3000 round	27	31	36
Energy consumed per time (s)	199/611	193/767	183/823
Data deliver ratio (pkt)	27	24	23
E2E delay in (ms)	0.67	0.62	0.58

protocols should prioritize package delivery and energy efficiency and ensure that the maximum number of nodes is available at any one time. Reduces the effect of a node failure, which results in decreased packet loss, increasing computing power.

The (Fig. 9a) depicts the number of dead nodes against the number of rounds in all presented routing protocols. Due to the enhanced cluster selection technique, the first dead node becomes apparent after 730 rounds, which is more visible than the LEACH method. It is observed that 73 nodes are dead in LEACH after 3000 rounds, in TL-LEACH 69 node after 3000 rounds, and in LEACH-EFT 64 after 3000 rounds. We optimized the performance parameter to evaluate the IoT efficiency and observed that LEACH-EFT performs better with an increasing number of nodes and network area.

The curve changes in the (Fig. 9b) indicating throughput of LEACH-EFT is relatively stable from TL-LEACH and LEACH protocol that are limiting the E2E delay in the network. After 400 seconds, the TL-LEACH and LEACH protocols’ throughput curves exhibit a declining tendency. The network begins to experience node problems. In some cases, data packets are dropped, resulting in a decrease in network performance.

The (Table 3) demonstrates that the suggested model's overall performance is more significant than conventional algorithms.

5 Discussions

The Internet of Things is used for various purposes, including environmental monitoring, smart grid, energy management, medical care, smart manufacturing, smart buildings, intelligent transportation, logistics, and smart cities. These applications demand a dependable, robust, instantaneous, and quick network for device support and communication. It is advantageous in practice and has significant research implications in various sectors, including national security, military, medical and health, environmental monitoring, and building monitoring. Micro WSN establishes an internet connection between personal computers, home appliances, and other everyday requirements, enabling remote tracking and control. Wireless sensors are used to save energy and regulate security.

The proposed study develops a Q-Learning-based routing algorithm capable of monitoring the status of nodes in real-time. The enhanced technique for route discovery eliminates undesirable local repercussions. Nodes will investigate other nodes depending on the current optimum path's favored selection. The algorithm continuously modifies the routing strategy based on the node's power consumption to minimize data loss due to node failure and lower the average energy consumption under the premise of reliable data collecting. Nodes that use dispersed routing protocols may learn to pick the ideal set of parent nodes iteratively. In contrast, source nodes that use central routing algorithms can learn to discover the perfect route combination. Finally, this study conducts simulation experiments from four aspects: different number of source nodes, the different total number of nodes, different ϵ greedy, and random routing algorithms. The experimental findings demonstrate that this technique has a shorter E2E latency than traditional data packet forwarding techniques. The probability of a node being elected as a cluster leader increases the network's life and improves its robustness. Adjusting the TDMA Schedule table generated during the clustering phase resolves the cluster head link failures.

However, the techniques used in WSNs are increasing node energy and lengthen the life of the sensor network, considerably improving network failure tolerance. Additionally, the proposed research addresses energy conservation and fault-tolerance issues integrated into the LEACH technique. The LEACH protocol's random cluster head election process enables dynamic load balancing; no consideration is made for the nodes' remaining energy level. When a low-energy node is selected as the cluster head, its energy gradually depletes. The algorithm continuously modifies the routing strategy based on the node's power consumption to minimize data loss due to node failure and average energy consumption under the assumption of continual data collection project execution. The experimental findings show that the developed algorithm has reduced end-to-end latency and uses less energy. Routing and link scheduling will be researched together in the future to reduce end-to-end latency. We will also work

on real implementation and deployment to assess performance outside of simulation's intrinsic limits.

6 Conclusion

WSN technologically improved quickly and became one of the most impactful technologies of the modern period. The extensive use of wireless networks has significantly aided in advancing people's social interactions and industrialization. The in-depth application of server networks, energy conservation, and fault tolerance has come to prominence.

This research aims to evaluate the energy conservation and fault-tolerance problems inherent in the classic LEACH-based technique. The random cluster head voting method in the proposed method achieves dynamic resource utilization; minimal consideration is required for the nodes' remaining energy level. The probability of a node getting selected as a cluster header extends the network's life and strengthens its resilience. The proposed research selected the cluster head using fault-tolerance technology. The cluster head node failure is resolved by updating the TDMA Schedule table created during the clustering phase.

Consequently, the Q-Learning algorithm, which is based on reinforcement learning, determines the ideal route for the cluster head node in complicated obstacle situations. The cluster head node takes advantage of the auxiliary information to actively choose the next route point to resolve the blind search issue early on. However, it rapidly converges during the early stages of learning and significantly enhances the algorithm's learning efficiency and convergence speed. Although the rate of technological advancement is unpredictable, the IoT will play a significant role in the changing world. As a result of the proposed study, node residual energy was enhanced, and the network life cycle was extended, resulting in significantly higher network failure tolerance. Routing and connection scheduling will be investigated collaboratively in the future to reduce E2E latency.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Ko YB, Vaidya NH (2000) Location-aided routing (lar) in mobile ad hoc networks. *Wirel Netw*, 6(4):307–321
2. Biswas S, Das R, Chatterjee P (2018) Energy-efficient connected target coverage in multi-hop wireless sensor networks. *Ind Interactive Innov Sci Eng Technol*, pp 411–421. Springer
3. Sotheara S, Aso K, Aomi N, Shimamoto S (2014) Effective data gathering and energy efficient communication protocol in wireless sensor networks employing uav. In: 2014 IEEE wireless communications and networking Conference (WCNC), pages 2342–2347. IEEE

4. Amutha J, Sharma S, Sharma SK (2021) Strategies based on various aspects of clustering in wireless sensor networks using classical, optimization and machine learning techniques: review, taxonomy, research findings, challenges and future directions. *Comput Sci Revi*, 40:1–45
5. Diallo O, Rodrigues JJ, Sene M (2012) Real-time data management on wireless sensor networks: a survey. *J Netw Comput Appl*, 35(3), 1013–1021
6. Kanoun O, Bradai S, Khriji S, Bouattour G, El Houssaini D, Ben Ammar M, Naifar S, Bouhamed A., Derbel F, Viehweger C, (2021) Energy-aware system design for autonomous wireless sensor nodes: a comprehensive review. *Sensors*, 21(2): 1–25
7. Papadimitratos P and Haas Z (2002) Secure routing for mobile ad hoc networks. In: communication networks and distributed systems modeling and simulation Conference (CNDS 2002), number CONF, pp 1–13. SCS
8. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd annual Hawaii international Conference on system sciences, pages 1–10. IEEE
9. Chawla N, Jasuja A (2014) Algorithm for optimizing first node die (fnd) time in leach protocol. *Int J Curr Eng Technol*, 4(4): 2748–2750
10. Tiberti W, Cassioli D, Di Marco A, Pomante L, Santic M (2021) A model-based approach for adaptable middleware evolution in wsn platforms. *J Sensor Actuator Netw*, 10(1): 1–22
11. Qiu M, Ming Z, Li J, Liu J, Quan G, Zhu Y (2013) Informer homed routing fault tolerance mechanism for wireless sensor networks. *J Syst Archit*, 59(4–5): 260–270
12. Mahmood T, Akhtar F, Ur Rehman K, Ali S, Mokbal FM, Daudpota S (2019) A comprehensive survey on the performance analysis of underwater wireless sensor networks (uwsn) routing protocols. *IJACSA*, 10(5):1–11
13. Littman Michael L (2015) Reinforcement learning improves behaviour from evaluative feedback. *Nature*, 521(7553): 445–451
14. Wang D, Liu J, Yao D, Member IEEE (2020) An energy-efficient distributed adaptive cooperative routing based on reinforcement learning in wireless multimedia sensor networks. *Comput Netw*, pp 1–12
15. Banerjee A, Sufian A (2020) Reinforcement learning based transmission range control (rl-trc) in sd-wsn with moving sensors. arXiv preprint [arXiv:2005.08215](https://arxiv.org/abs/2005.08215), pp 1–27
16. Sivakumar P, Radhika M (2018) Performance analysis of leach-ga over leach and leach-c in wsn. *Procedia Comput Sci*, 125:248–256
17. Azharuddin M, Jana Prasanta K (2015) A distributed algorithm for energy efficient and fault tolerant routing in wireless sensor networks. *Wirel Netw*, 21(1), 251–267
18. Shi F, Tuo X, Yang SX, Lu J, Li H (2019) Rapid-flooding time synchronization for large-scale wireless sensor networks. *IEEE Trans Ind Inf*, 16(3): 1581–1590
19. Mahmood T, Akhtar F, Rehman KU, Azeem M, Mudassir A, Daudpota SM (2020) Introducing robustness in dbr routing protocol. *Int J Commun Netw Distrib Syst*, 24(3): 316–338
20. Tripathi M, Gaur MS, Laxmi V, Battula RB (2013) Energy efficient leach-c protocol for wireless sensor network. pp 1–4
21. Ali B, Mahmood T, Mirza MA, Memon S, Rashid M, Ajebesone EF (2019) Study and analysis of delay sensitive and energy efficient routing approach. *IJACSA*, 10(8):14–20
22. Moon SY, Cho TH (2009) Intrusion detection scheme against sinkhole attacks in directed diffusion based sensor networks. *Int J Comput Sci Netw Secur*, 9(7): 118–122
23. Loscri V, Morabito G, Marano S (2005) A two-levels hierarchy for low-energy adaptive clustering hierarchy (tl-leach). In: IEEE vehicular technology conference, 62, pp 1809–1813. IEEE
24. Ko YB, Choi JM, and Kim JH (2004) A new directional flooding protocol for wireless sensor networks. In: international Conference on information networking, pp 93–102. Springer
25. Zhe H, Maiko S (2009) New bounds on the minimum number of calls in failure-tolerant gossiping. *Netw Int J*, 53(1):35–38
26. Dehkordi SA, Farajzadeh K, Rezaeadeh J, Farahbakhsh R, Sandrasegaran K, Dehkordi MA (2020) A survey on data aggregation techniques in iot sensor networks. *Wirel Netw*, 26(2): 1243–1263
27. Lindsey S and Raghavendra CS (2020) Pegasus: power-efficient gathering in sensor information systems. In: Proceedings, IEEE aerospace conference, 3: 1125–1130. IEEE
28. Arati M and Agrawal Dharma P (2001) Teen: arouting protocol for enhanced efficiency in wireless sensor networks. In *ipdps*, 1: 1–7
29. Yao Y, Gehrke J, et al. (2003) Query processing in sensor networks. In *Cidr*, pp 233–244

30. Younis O, Fahmy S (2004) Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans Mobile Comput*, 3(4):366–379
31. Qiu M, Liu J, Li J, Fei Z, Ming Z, Edwin HM (2011) A novel energy-aware fault tolerance mechanism for wireless sensor networks. In: 2011 IEEE/ACM international Conference on green computing and communications, pp 56–61. IEEE
32. Liu X, Cao J, Bhuiyan M, Lai S, Wu H, Wang G (2011) Fault tolerant wsn-based structural health monitoring. In: 2011 IEEE/IFIP 41st international Conference on dependable systems and networks (DSN), pp 37–48. IEEE
33. Cheraghlou MN, Haghparast M (2014) A novel fault-tolerant leach clustering protocol for wireless sensor networks. *J Circuits Syst Comput*, 23(03): 1–17
34. Mehdi M, A Al-Fuqaha, M Guizani, JS Oh (2017) Semisupervised deep reinforcement learning in support of iot and smart city services. *IEEE Internet Things J*, 5(2): 624–635
35. Tang F, Zhou, Y, Kato N (2020) Deep reinforcement learning for dynamic uplink/downlink resource allocation in high mobility 5g hetnet. *IEEE J Sel Areas Commun*, 38(12): 2773–2782
36. Ajmi N, Helali A, Lorenz P, Mghaieth R (2021) Speech-mac: special purpose energy-efficient contention-based hybrid mac protocol for wsn and zigbee network. *Int J Commun Syst*, 34(1): 1–18
37. Zhao T, Xu XB, Wang SG (2020) Centralized q-learning based routing in eh-wsns with dual alternative batteries. *J Phys Conf Series*, 1544: 1–10. IOP Publishing
38. Seah MWM, Tham CK, Srinivasan V, Xin A (2007) Achieving coverage through distributed reinforcement learning in wireless sensor networks. In: 2007 3rd international Conference on intelligent sensors, sensor networks and information, pp 425–430. IEEE
39. Wang CW, Xia Q, Yao X, Wang W, Jornet JM (2018) Multi-hop deflection routing algorithm based on q-learning for energy-harvesting nanonetworks. In: 2018 IEEE 15th international Conference on mobile ad hoc and sensor systems (MASS), pp 362–370. IEEE
40. Bouzid SE, Serrestou Y, Raouf K, Omri MN (2020) Efficient routing protocol for wireless sensor network based on reinforcement learning. In 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp 1–5
41. Sun P, Lan J, Guo Z, Xu Y, Hu Y (2020) Improving the scalability of deep reinforcement learning-based routing with control on partial nodes. In: ICASSP 2020–2020 IEEE international Conference on acoustics, speech and signal processing (ICASSP), pp 3557–3561. IEEE
42. Raj RN, Nayak A, Kumar MS (2020) A survey and performance evaluation of reinforcement learning based spectrum aware routing in cognitive radio ad hoc networks. *Int J Wirel Inf Netw*, 27(1): 144–163
43. Yau KLA., Komisarczuk P, Teal PD (2012) Reinforcement learning for context awareness and intelligence in wireless networks: review, new features and open issues. *J Netw Comput Appl*, 35(1), 253–267
44. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT press
45. Junping H, Yuhui J and Liang D (2008) A time-based cluster-head selection algorithm for leach. In: 2008 IEEE symposium on computers and communications, pp. 1172–1176. IEEE
46. Ali MS, Dey T, Biswas R (2008) Aleach: advanced leach routing protocol for wireless microsensor networks. In: 2008 international Conference on electrical and computer engineering, pp. 909–914. IEEE
47. Batra PK, Kant K (2016) Leach-mac: a new cluster head selection algorithm for wireless sensor networks. *Wirel Netw*, 22(1): 49–60
48. Subhashree VK, Tharini C, Swarna LM (2014) Modified leach: a qos-aware clustering algorithm for wireless sensor networks. In: 2014 international Conference on communication and network technologies, pages 119–123. IEEE
49. Kumar GS, Vinu PMV, Jacob KP (2008) Mobility metric based leach-mobile protocol. In: 2008 16th international Conference on advanced computing and communications, pages 248–253. IEEE, 2008
50. Sharma V, Alam B, Doja MN (2019) An improvement in dsr routing protocol of manets using anfis. In applications of artificial intelligence techniques in engineering, pages 569–576. Springer

Authors and Affiliations

Tariq Mahmood^{1,2}  · **Jianqiang Li**^{1,3}  · **Yan Pei**⁴  · **Faheem Akhtar**⁵  ·
Suhail Ashfaq Butt²  · **Allah Ditta**²  · **Sirajuddin Qureshi**¹

¹ Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

² Department of Information Sciences, Division of Science and Technology, University of Education, Lahore 54000, Pakistan

³ Beijing Engineering Research Center for IoT Software and Systems, Beijing 100124, China

⁴ Computer Science Division, University of Aizu, Aizuwakamatsu, Fukushima 965-8580, Japan

⁵ Department of Computer Science, Sukkur IBA University, Sukkur 65200, Pakistan