I. J. Computer Network and Information Security, **2023, 4, 96-107**
Published Online on August 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2023.04.09

Reliable Data Delivery Using Fuzzy Reinforcement Learning in Wireless Sensor Networks

Sateesh Gudla*
Department of Computer Science and Engineering, JNTUK, Kakinada
Department of Computer Science and Engineering, Lendi Institute of Engineering and Technology (A), Vizianagaram, Andhra Pradesh, India
E-mail: sateesh.research@gmail.com
ORCID iD: https://orcid.org/0000-0001-8785-2190
*Corresponding Author

Kuda Nageswara Rao
Department of Computer Science and Systems Engineering, Andhra University, Andhra Pradesh, India
E-mail: knrao2038@gmail.com

Received: 28 January 2023; Revised: 19 April 2023; Accepted: 27 May 2023; Published: 08 August 2023

Abstract: Wireless sensor networks (WSNs) has been envisioned as a potential paradigm in sensing technologies. Achieving energy efficiency in a wireless sensor network is challenging since sensor nodes have confined energy. Due to the multi-hop communication, sensor nodes spend much energy re-transmitting dropped packets. Packet loss may be minimized by finding efficient routing paths. In this research, a routing using fuzzy logic and reinforcement learning procedure is designed for WSNs to determine energy-efficient paths; to achieve reliable data delivery. Using the node's characteristics, the reward is determined via fuzzy logic. For this paper, we employ reinforcement learning to improve the rewards, computed by considering the quality of the link, available free buffer of node, and residual energy. Further, simulation efforts have been made to illustrate the proposed mechanism's efficacy in energy consumption, delivery delay of the packets, number of transmissions and lifespan.

Index Terms: Wireless Sensor Networks, Q-learning, Reinforcement Learning Algorithm, Fuzzy Logic, Energy Efficiency, Reliable Paths, Network Lifetime.

1. Introduction

Recent advances in improving sensing technologies in Wireless Sensor Networks (WSNs) lead to expanding application scenarios such as environmental monitoring, health monitoring, battlefield surveillance, forest fire detection, and issuing tsunami alerts [1–4]. Tiny sensor nodes in a wireless sensor network have constrained sensing, battery life, memory, processing speed, and communication abilities [4]. The finite battery (energy) directly influences the wireless sensor network's lifetime. Sensor nodes utilize energy to sense, send, and receive data. The multi-hop communication approach consumes more energy for transmitting the data than sensing. During the data transmission, intermediate nodes consume much energy compared to the network's other nodes. It thus creates an energy hole problem, and packets may be lost due to the premature death of sensor nodes. A packet loss may happen due to the less available sensor nodes buffer and the link quality. Therefore, a sensor node drops the packet due to less residual energy, low available buffer memory, and poor link quality.

Since wireless sensor networks are typically deployed in isolated regions where frequent maintenance, such as battery replacement, is not possible, increasing the network's lifetime is a significant difficulty. Therefore, energy conservation in sensor networks is a considerable aspect. A packet in a WSN will be retransmitted to achieve reliable data delivery, which is essential in applications like health monitoring in chemical plants, nuclear plants, bridges, and other data-sensitive applications. Due to the data retransmission, the data delivery time is also prolonged. Hence, the packet drops lead to a rise in energy consumption and a delay in packet delivery. Therefore, reducing packet losses improves the network's energy efficiency and data delivery time. Finding energy-efficient, high-quality, and well-buffer-balanced routing pathways can significantly minimize packet loss.

This current work presents a routing technique based on a reinforcement learning mechanism to discover the energy-efficient route in a WSN. An optimal path or the energy-efficient route here in this context is defined as the route, where each node in the path is resourceful in terms of remaining energy, available buffer, link quality, and distance. Consequently, it reduces the packet dropping rate and improves reliability and lifetime. A sensor node's weight (reward) is determined using a fuzzy logic controller that considers the node's energy, available free buffer, distance, and link quality. These nodes' rewards have been used in Reinforcement Learning to achieve a high data transmission rate and to reduce packet drops. The proposed mechanism finds the highly resourceful nodes (in terms of remaining energy, available free buffer, and quality of link) as intermediate nodes in a routing path. Specifically, this study aims to accomplish the following.

- Design a routing system based on fuzzy reinforcement learning to increase energy conservation and reliability in data delivery in WSN.
- Design of a fuzzy reinforcement learning-based routing method with a fuzzy logic-based method for assessing the reward.
- The efficiency of the proposed mechanisms is demonstrated through simulation.

The remaining parts of this paper are arranged as follows. The related work and motivation are discussed in Section 2. Section 3 explains the suggested routing technique based on fuzzy reinforcement learning. In Section 4, the outcomes of the simulations are shown, and in Section 5, the conclusions are discussed.

2. Related Works and Motivation

This section discusses the existing routing approaches in WSNs, fuzzy-based, and reinforcement learning-related mechanisms in WSNs.

Akyildiz et al. [1,2] have discussed design factors, applications, protocol stack, and challenges of WSNs. WSN's data exchange protocol must be developed with energy efficiency in mind. A network's lifetime is proportional to its energy efficiency. A WSN's lifetime is presented in several ways throughout the literature. Lifetime and energy consumption models in WSN have been the subject of research. A model for Energy Efficient Cluster Based Routing (EECBR) was described by U. Hariharan et al.[3], the suggested model optimized the choice of Cluster Head (CH) via the Grey Wolf Optimization algorithm, with energy and distance serving as inputs. EECBR uses a Tree-based Remote Vector routing method and an advanced Multihop Dijkstras algorithm to perform intra-cluster routing instead of a traditional Base Station (BS). Using residual energy, available buffer, link quality, and distance as inputs, S. Gudla et al. [4] proposed a learning automaton and A*algorithm-based energy efficient routing approach in WSN, wherein the weight of the node is calculated using learning automata and the route is computed using A*algorithm.

Fuzzy logic systems have already been implemented in wireless sensor networks [5–15] to aid in cluster head selection, effective decision-making, maximizing network lifetime while minimizing energy consumption, and developing effective routing protocols. In [5], Neamatollahi et al. designed a fuzzy policy to lessen the re-clustering problem in WSNs. Ni et al. [6] examined how to extend the life of a sensor network by selecting cluster heads using particle swarm optimization and the idea of fuzzy logic. In order to extend the life of nonuniform WSNs, Baranidharan Balakrishnan et al. [7] have developed a Fuzzy Logic-based energy-efficient clustering hierarchy (FLECH); to maximize the lifespan of the network, FLECH makes use of fuzzy logic to intelligently combine factors like residual energy, node centrality, and distance to BS to elect the most optimal nodes to serve as cluster head. R.Yousaf et al. [8] proposed a method based on fuzzy logic to determine the amount of time sensor devices in an IWSN using the IEEE 802.15.4 protocol spend sleeping based on the current battery level and the ratio of throughput to Workload. An energy-efficient and long-lasting wireless sensor network (WSN) is suggested by Zhang Siqing et al. [9], along with a fuzzy-logic-based clustering protocol for multi-hop WSNs. Based on parameters such as leftover energy, free buffer space, distance, and link quality, S.K.Mothku et al. [10] employed a fuzzy logic approach to determine the best path a packet can take via a network.

To stop intruders from getting into WSN Neha Singh et al. [11] developed a fuzzy rule-based system. The proposed approach is based on classifying nodes into the "red," "orange," and "green" categories, respectively. In [12], the author presented an energy-efficient, Fuzzy-based cluster head selection method to help WSNs survive longer in service. Stephan et al. [13] used fuzzy logic to implement zone-based clustering and then flexibly select a cluster head (CH), which aimed to fix the unfair distribution of CHs' unused energy. S. Phoemphon et al. [14] suggested a four-step fuzzy logic-based clustering architecture: determining the competition radius (CR), electing a CH, having a CM enter the cluster, and deciding on the criteria for electing the next CH. Giri.A et al. [15] developed a routing protocol to transmit information to the base station and a fuzzy clustering algorithm to pick cluster heads. The optimized fuzzy clustering algorithm chooses cluster heads with type-1 fuzzy reasoning by considering residual energy, distance from the base station, and node concentration. Further particle swarm optimization determines the most energy-efficient route to the base station.

It is possible to model system uncertainties and make real-time judgments using fuzzy logic. The fuzzy logic controller can also use many inputs and apply several rules to arrive at a single output. Fuzzy logic is more straightforward and adaptable and uses less processing power than traditional logic.

Networks in traditional WSN approaches are slow to respond dynamically because they have been particularly programmed [16]. To successfully deal with situations like these, machine learning (ML) methods can be used [16]. At its core, ML can automatically learn from experience without the need for human input or re-programming When applied, the machine learning technique, reinforcement learning (RL), allows us to gain insight from previous experiences with its surroundings so that it can make more informed decisions about its future actions [17].

A new Q-learning technique (QLCLUSTER) is presented by Farzad K [18] to find the best routes from nodes to a faraway medical hub. Guo et al. [19] developed a routing protocol driven by a reinforcement learning algorithm using distance, hop count, and residual energy. A scheduling approach using the Q-learning technique is proposed by D.Kim et al. [20] for a real-time and dynamic IoT environment. Qin Yang et al. [21] suggested that the FANET routing protocol be built on Q-learning-style fuzzy logic. In terms of link and overall path performances, the suggested algorithm makes it easier to identify the routing paths to be processed. Each UAV uses a fuzzy system with a link- and path-level parameters to select its own best routing path to the destination. Parameters at the path level include hop count and successful packet delivery time, while those at the link level include transmission rate, energy condition, and fight status between neighbor UAVs. The reinforcement learning approach dynamically adjusts the path-level parameters. In [22], Maivizhi et al. presented an adaptive routing algorithm that uses the Q-learning reinforcement learning algorithm to obtain a routing tree with minimum data, such as energy levels, proximity, and link quality. W.K.YUN et al. [23] have introduced data aggregation and an energy-efficient Q-learning-based routing to discover the best path to boost the network's lifespan. In [23], a reward policy is defined based on energy, distance, and hop count. Tripti Sharma et al. [24] introduced a technique that builds on grouping and reinforcement learning. Clustering-based methods have been widely used in the development of energy-efficient protocols, and reinforcement learning is a collection of algorithms motivated by operant conditioning in animal behavior. However, the current work proposed a reliable data delivery approach based on Q-learning to find energy-efficient paths to decrease the number of packets dropped in WSNs. The reward in the proposed system is decided with the aid of the fuzzy logic system and called the node's fuzzy probability by considering link quality, free buffer size, distance, and residual energy.

In this work, fuzzy has been used to find the weight or reward of a node using operating distance, free buffer size, link quality, and available energy as the parameters. The reinforcement learning mechanism will use the reward to find energy-efficient routes in WSNs.

Motivation

In wireless sensor networks, reducing the packet drop rate diminishes the data retransmission rate, enhances energy efficiency, and reduces packets' delivery delay. Maintaining reliable data transmission, efficient use of energy, and stringent delivery delay of packets is very much required for WSNs applications such as health monitoring, battlefield surveillance, and forest fire detection. In WSNs, packets will be lost because of numerous factors, such as low energy levels, no available buffer, and poor link quality. Therefore, selecting a node (among the neighbor nodes) to serve as an intermediate node in a route from one place(source) to another (destination) is challenging. The current routing procedure selects an efficient node with high free buffer availability, increased energy levels, and better link quality as a relay node. The current research uses a reinforcement learning-based technique to discover the routes in WSNs. The node's weight or reward is calculated using fuzzy logic.

3. Proposed Mechanism

A model of a network is presented here. A fuzzy logic-based reward mechanism is explained for evaluating the reward of a node. Further, Routing Algorithm using Fuzzy Logic and Reinforcement Learning (FRLRA) is shown to find energy-efficient routes in WSNs.

3.1. Network Model and System Architecture

To represent a wireless sensor network, we can think of it as an undirected graph G (V, E); the nodes of a WSN are represented by the set V, whereas the edges (links) are represented by the set E, and S (sink node) ∈ V. Through sensing and multi-hop communication, data is gathered at every node (sensor) and sent to S. Here; data packets may be lost at intermediate nodes due to poor link quality, unavailability of remaining energy, and free buffer at a sensor node.

3.2. Fuzzy Logic-based Reward Mechanism for Calculating a Node's Reward

The fuzzy logic system is discussed in this section to evaluate sensor node weights (rewards). Most system characteristics rely on environmental conditions when using a WSN in a dynamic environment. Fig.1.shows a fuzzy logic system. The application of fuzzy logic in decision-making under conditions of uncertainty has led to significant advances in this area.

Choosing the perfect node in a routing path is a challenging problem due to the uncertain characteristics of a node. One way to use a fuzzy logic controller is to employ a combination of parameters and rules to arrive at the desired outcome [8,21]. In the current work, we use fuzzy logic to calculate the node's weight (reward) by considering the node's features, such as distance, link quality, leftover energy, and available buffer space.

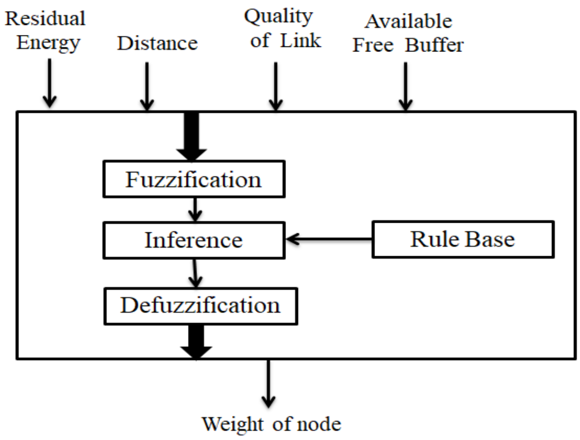

Fig.1. Fuzzy logic system

A fuzzy logic system produces a reward for each sensor node by combining buffer space, link quality, distance, and residual energy. The fuzzy logic system has four stages:[10] (i) fuzzification, (ii) member functions, (iii) fuzzy rule set, and (iv) defuzzification. The crisp values of the four input parameters are translated into linguistic variables in the fuzzification stage using the membership functions. The linguistic considerations (refer to Table 1) for the input variables residual energy and buffer are low, medium, and high (shown in Fig.2), for the input variable link quality, are poor, average, and good (shown in Fig.3). for the input variable distance are close, adequate, and far (shown in Fig.4). The fuzzy output is a reward, and its linguistic considerations are very low, low, low medium, high medium, and strong (refer to Fig.5). In addition, a node's reward (weight) is calculated using the fuzzy rule set, i.e., IF/THEN rules- out of the 81 rules for our fuzzy system Table 2 shows the sample rules. The membership functions and the fuzzy rule set combine to provide a fuzzy inferred output via the fuzzy inference system (reward). Lastly, in the defuzzification stage, the centroid method calculates the reward (crisp value) [14].

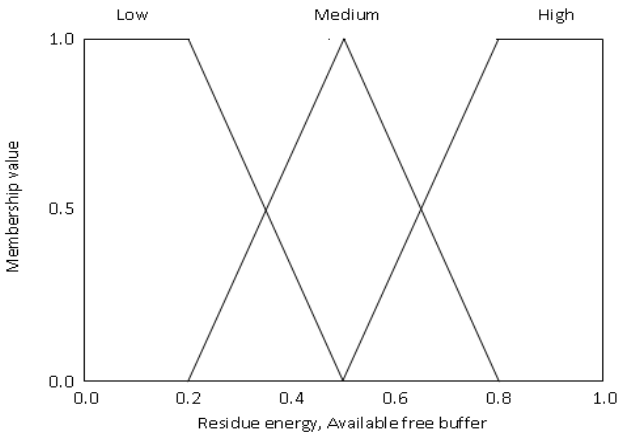

Fig.2. Fuzzy- membership-functions for residual energy

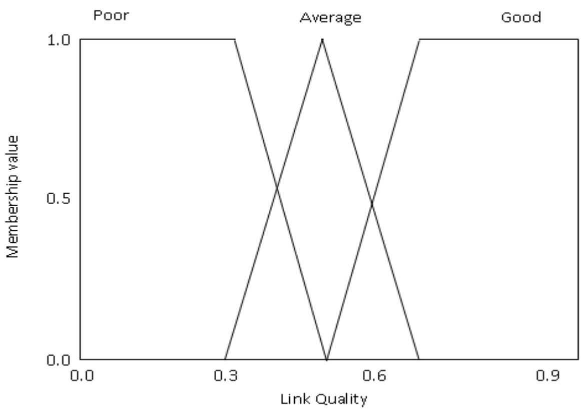

Fig.3. Fuzzy- membership function for quality of link and available free buffer

3.3. Fuzzy Reinforcement Learning-based Routing Algorithm (FRLRA)

This section explains Fuzzy Reinforcement Learning based Routing Algorithm (FRLRA) to find effective routing paths in a wireless sensor network.

Algorithm 1. Fuzzy Reinforcement Learning-based Routing Algorithm (FRLRA)

v is a source node, Initialize ERP (i.e., route) = { };
Make all node values 0 (i.e., Unvisited);
Construct the Q-table;
ERP= v;
Mark v to 1 (i.e., Visited);
do
For all possible neighbor nodes (i.e.,$w \in N_v$ w**)** **,do**
    Compute the Fuzzy Probability (reward) FP using the Fuzzy Inference System;
If FP> δ and dist(w, Sink) < dist(v,Sink) then
$N_{list} = N_{list} \cup w$; ▷ N_{list} is the new neighbor's list
End If
End For
If $\boldsymbol{N_{list}}$ is not null, then
    If random (0,1)<ξ **then** ▷ ξ **is exploration trade-off**
N_{chosen}= randomly choose a node from N_{list};
Else
N_{chosen}= Choose largest Q-value value node from N_{list} ($\in$-greedy mechanism));
End If
$ERP = ERP \cup N_{chosen}$;
Mark $\boldsymbol{N_{chosen}}$ to 1 (i.e., Visited)
End If
If N_{chosen} is not Sink, then
$R = RC$; (Constant reward)
Else If $Residual_{Energy}$ of $N_{chosen} < \theta$ **then**
$R = -RE$ (i.e.,Low Energy Penalty)
Else If $\boldsymbol{Buffer\ of\ N_{chosen}} < \hat{\gamma}$ **then**
$R = -RB$ (i.e., Low Buffer Penalty)
Else If N_{chosen} has no neighbors **then**
$R = -RN$ (i.e., No Neighbor Penalty)
Else $R = FP$ (i.e., Fuzzy Probability)
End If
Next state (i.e.v^1) =N_{chosen}
Update Q-Value:$Q(v, N_{chosen}) = R + \gamma * MaxQ(v^1, N_{chosen})$
v=v^1**;**
While N_{chosen}! = Sink

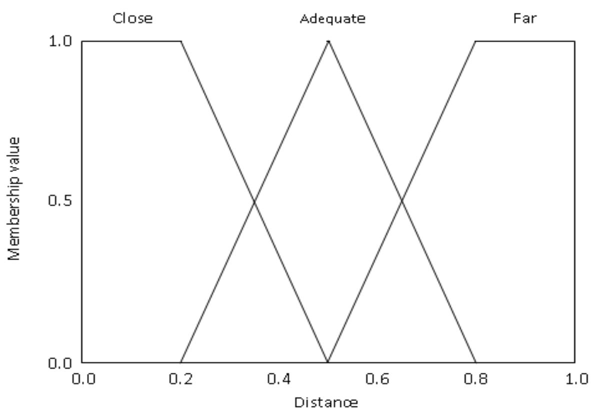

Fig.4. Fuzzy-membership-functions for distance

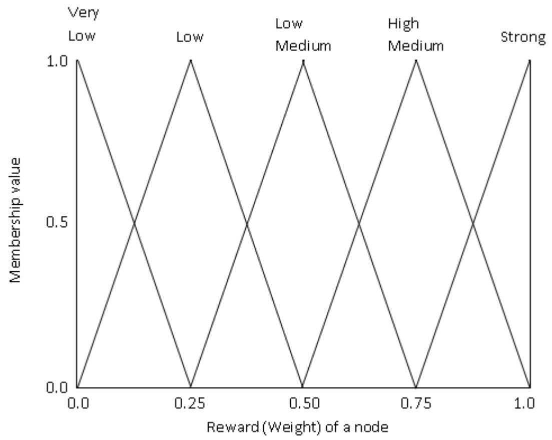

Fig.5. Fuzzy-membership-functions for reward

Table 1. Linguistic values for inputs and outputs

Input / Output variable	Linguistic values
Residual energy	Low, Medium, and High
Buffer	Low, Medium, and High
Link quality	Poor, Average, and Good
Distance	Close, Adequate, and Far
Reward	Very Low, Low, Low Medium, High Medium, and Strong

Table 2. Rules for making fuzzy decisions

Inputs				Outputs
Residual Energy	**Buffer**	**Distance**	**Link Quality**	**Reward of a node**
Low	Low	Far	Poor	Very Low
Low	Low	Adequate	Average	Low
Low	Medium	Adequate	Good	Low Medium
Low	High	Close	Good	High Medium
Medium	Medium	Far	Good	High Medium
Medium	High	Adequate	Good	Strong
Medium	High	Close	Poor	Low Medium
High	Low	Close	Average	Low
High	Medium	Close	Good	Strong
High	High	Close	Good	Strong

Substantial amount of data about the state of the network are required to work on the proposed problem. Since WSNs are inherently dynamic, conventional optimization mechanisms may not give optimal (in terms of reliable data transmission and energy efficiency) routing paths. In recent years, RL is a widespread technique and more suitable for dynamic networks, which motivated us to apply it to this problem. Rather than looking back at earlier states to figure out where to go, the route is calculated using only the current state. The proposed problem can be modeled as a Markov decision process; it is termed as a tuple $\{S, A, R, P, \gamma\}$. "S" denotes the state space, “A” is action space, “R” is system reward, “P” is immediate reward and "γ" is discount factor.

- All the network nodes represent a state space.
- Action represents a packet that traverses from one node to another node. A $=\{a_1, a_2, \ldots, a_n\}$. The action means the next hop node is in the optimal destination path. Actions are made based only on current state data. (i.e., Each state may take a different course of action)
- R is the reward, r_i immediate reward as a result of an action a_i performed on an environment in state S0. Penalties are built into the reward, if necessary, resources are not made available, and R denotes reward, a value from Table 3.

Table 3. Reward value

Reward (R)	Meaning
FP	Nchosen is not sink
RC	Nchosen is sink
-RE	Low Energy Penalty
-RB	Low Buffer Penalty
-RN	No Neighbour Penalty

- The transition probability is denoted by P. According to the possibilities of moving from one state to the next, the agent takes some action.
- γ, the discount factor can be any value from 0 to 1. An agent with γ = 1 considers the sum of all future rewards when deciding, while an agent with γ= 0 learns an action based on the immediate reward.

Q-Learning Mechanism

In this work, to resolve the routing issue, the best possible policy or value has been determined using a reinforcement learning (RL) strategy [16,17]. Most researchers prefer Q-learning [22,23] as the best model-free reinforcement learning technique. Q-learning is a method that uses experimentation and observation to decide the finest course of action to take. The Q-learning agent optimizes its policy by taking whatever action yields the highest reinforcement signal (reward). In the reinforcement learning process, the agent receives a reward for its actions and learns to conform to the environment based on its state. Each action will be given a Quality value (Q-value). Agent updates this Q-value with the received reward during the learning phase. Eventually, the agent learns the actual values and makes the most efficient path selections.

Table 4. Summary of notations

Symbol	Description
v	Source Node
ERP	Efficient Route Path
N_V	Set of neighbor nodes of v
FP	Fuzzy Probability or Reward
RC	Constant Reward
RE	Low Energy Penalty
RB	Low Buffer Penalty
RN	No Neighbor Penalty
ө	Residual Energy Threshold
$\hat{\gamma}$	Buffer threshold

Given the distance between nodes, link quality, node energy, and buffer capacity, finding the optimized route between two nodes (source and sink) is our goal. Finding the shortest path is a job of the Q-learning procedure. Before proceeding to the fuzzy reinforcement learning to find the optimal path, our notational conventions are as follows. The N_v is the set of neighbor nodes to source node $v \in V$ where $N_v \subset V$, ERP is the efficient route between the node v and sink node, and distance from w and sink is dist(w,Sink). In this problem, the packet is the agent, the WSN is the environment, and the nodes are the states. Upon receiving the neighbor node's information, each node executes the proposed mechanism to find the shortest route. In each time slot, the agent acts on the environment, obtains the reward, and moves to the next state (next node in the shortest path). The agent's goal is to maximize the sum of all available rewards. On the other hand, the agent does not have the choice of the highest instant reward for every action. Some rewards may take longer than expected. The agent will gain experience determining the optimal process for providing the highest long-term reward. Therefore, the objective of learning is to maximize the reward that can be expected, which is defined as a packet is routed from node-i to one of the neighbor nodes; let us say (node-j1), (assume node-j1, node-j2, ...node-jd are neighbor nodes, d is the degree of node-i) and it sends back a reward. The Q-value will be computed with this reward value. The new Q-value is defined as

$$Q_{new}(i, j_1) = R(i, j_1) + \gamma * Max\big(Q(i, j_1), Q(i, j_2), \ldots, Q(i, j_d)\big) \quad (1)$$

Where: $Q_{new}(i, j_i)$ denotes the new Q-Value.

The procedure for determining the shortest path formation for an agent is as follows. First, initialize route ERP as the empty set. For each source node v, find the neighbor set N_v. For each node v in the neighbor node set, construct the Q-table, add the node v as visited, and compute the reward using exploration and exploitation based on the fuzzy probability value.

In Algorithm.1, we get an overview of the proposed method for reliable data transmission, which uses fuzzy reinforcement learning. Every source node in the WSN proceeds to do so again. The set of all possible nodes is figured out for each node, and the best suitable next hop node is determined. Assume v is the source node, and ERP denotes an efficient route path, initially empty. The source node constructs the Q-table. The source node v will be added to ERP and mark source node v as visited (Lines 1-5, Algorithm 1). For source node v, find the neighbor nodes set (N_v) using hello packets. Compute the fuzzy probability value (reward or weight) (FP). For each neighbor node (refer to Section 3.2). If a neighbor node's ($w, w \in N_v$) fuzzy probability value is more than a threshold value(δ) and if it is close to the sink node than source node v, then add source node v to the new neighbor's list (Nlist) (lines 7-12). Lines 13-18 give the exploration and exploitation based on the thresholdξ. A node will be chosen as NChosen from the Nlist. The selected node (NChosen) will be added to the efficient route path (ERP) and made as visited (lines 19-20). Lines 22-32 compute the reward for the action performed by the agent (refer to Table 3). The Q-values are updated in Line 34 using Equation 1. The agent will move to the next state, i.e., to the next node NChosen. The process will repeat if NChosen is not a sink node (lines 33-36). The summary of the notations is described in Table 4.

4. Performance Evaluation

The effectiveness of the current mechanism, fuzzy logic, and reinforcement learning routing algorithm (FRLRA) in terms of consumption of energy, the number of transmissions (including retransmissions), average packet delivery delay, and lifetime is assessed here. The simulation is conducted via Network Simulator-3 (NS3). In the simulation, A comparison is made between the proposed routing technique, the RINA [22] and Q-DAEER [23] routing mechanisms. In this work, the data collection round is defined as each node generating one packet and sending it to the sink node. In the simulation, every packet is retransmitted until it reaches the sink node to accomplish reliability in the data delivery. By considering this assumption, we estimated the node's energy consumption, average delivery delay of the packet, total retransmissions, and network lifetime. Network lifetime is a data collection round in which the first node in the network dies [4].

4.1. Parameters for Simulation

The parameters considered for simulation are described in Table 5 [10]. The simulation runs with a range of network sizes, from 50 to 500 nodes, the initial energy of a node is 25K Joules, and the buffer capacity is 2.5K bytes. The values of the links' qualities are graded on a scale from 0 to 1. The following energy consumption parameters are considered. Es, Er, and Et denote energy consumption for sensing, receiving, and transmission operations [5,10].

4.2. Evaluation of Simulation Outcomes

For various applications of WSNs, it is essential to enhance energy efficiency and reliable data collection. Reliable data transmission depends on packet loss. Re-transmission of dropped packets does not satisfy the strict delivery delay requirement and increases energy consumption. Therefore, reducing the packet dropping rate improves energy efficiency and reduces the delay in reliable data collection. The main parameters, such as buffer overflow, unreliable link quality, and low energy levels (temporary disconnection of nodes due to low energy levels), are identified as the factors that cause the dropping of packets.

Table 5. Parameters-of-simulation

Parameter	**Value**
Number Nodes	50 to 500 nodes
Range of transmission	30 meters
Node's initial energy	25 KJ
Packet Size	960 bits
$E_S = \alpha_3$	$\alpha_3 = 50 \times 10^{-9}\, Joules/bit$
$E_r = \alpha_{12}$	$\alpha_{12} = 0.787 \times 10^{-6}\, Joules/bit$
$E_t = \alpha_{11} + \alpha_2 d^n$	$\alpha_{11} = 0.937 \times 10^{-6}\, Joules/bit$ $\alpha_2 = 10 \times 10^{-12}\, Joules/bit/\, meters^2$ d = 85 meters

Fig.6 compares the packet delivery delay among RINA[22], Q-DAEER[24], and the proposed mechanism FRLRA. Looking at Fig.6, the proposed mechanism has the shortest packet delivery time compared to the other two. In the current work, a reward of a node is computed using fuzzy logic; based on the available free buffer, remaining energy, link quality, and distance of the node. However, all the mechanisms use the reinforcement learning algorithm, but the reward calculation is different. The routing decisions are taken based on various parameters. In RINA [22], the reward is calculated (the routing decisions are) based on residual energy, distance, and quality of the link. Here, a node's current free buffer is not considered while routing. The packet dropping rate increases if a node does not have a free buffer. In Q-DAEER [23] mechanism, the routing paths are constructed based on energy level, distance, and hop count. Here, the main parameters like link quality and current free buffer are not considered while taking routing decisions. Hence, the packet loss is more in Q-DAEER [23] mechanism than in the other two mechanisms. If the packet loss increases, the number of retransmissions increases to accomplish highly reliable data delivery. If there are many nodes in a network, there will be more packet drops, slowing down data delivery.

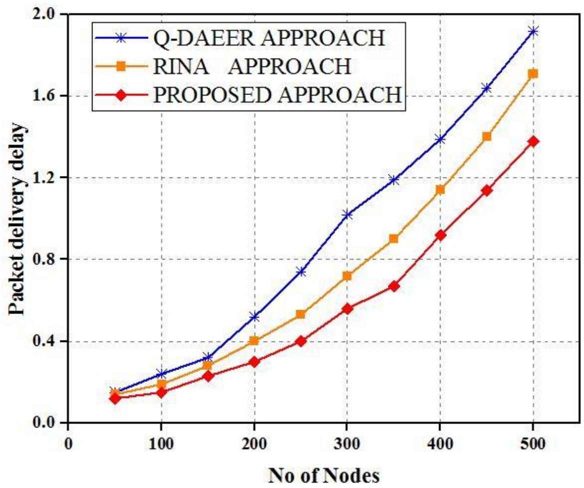

Fig.6. Comparison of average end-to-end packet delivery delay

Fig.7 compares the total number of transmissions among the three mechanisms. The rate at which packets are dropped in transmission raises the network's retransmission overhead. Via simulation, every lost packet is retransmitted to achieve better reliability in data delivery. The proposed method determines which network node should come next in each routing path with better-remaining energy, strong link quality, more available buffer, and shorter distance. This routing decision reduces packet loss in the network compared to the other two mechanisms. The data in Fig.7 states that our approach indicates a considerable transmission reduction.

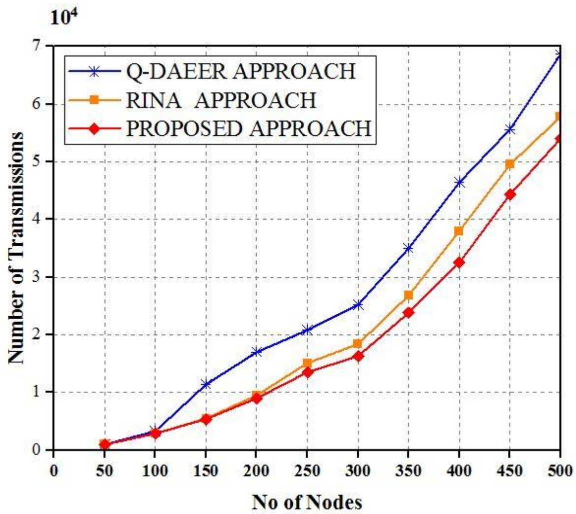

Fig.7. Comparison of the number of transmissions

A comparative evaluation of the energy consumption is presented in Fig.8. When compared to RINA [22] and Q-DAEER [23] processes; the proposed mechanism has lower network energy consumption. This is due to more packet loss in Q-DAEER and RINA mechanisms. The more the rate at which packets are lost, the more the rate they must be resent, hence the more energy requirements. The number of transmissions in the RINA[22] and Q-DAEER [23] is higher than in the proposed mechanism. Because of this, there was a corresponding rise in the amount of energy used by the network. Compared to the RINA [22] and the Q-DAEER[24], the proposed method resulted in up to 22% and 16% lower energy consumption, respectively.

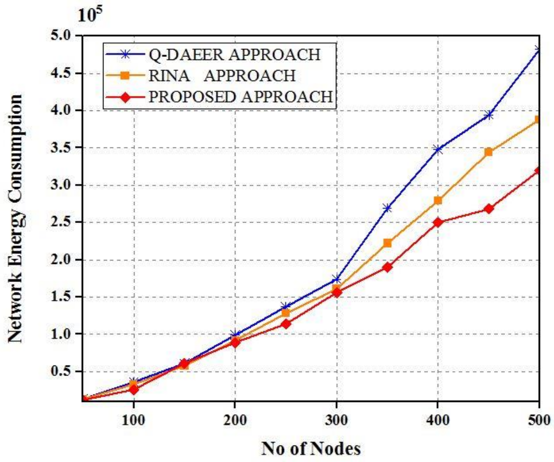

Fig.8. Comparison of network energy consumption

Fig.9 and Fig.10 show that first node died round (data collection round) and half nodes died round, respectively. As depicted in Fig.9, FND arrived earlier in Q-DAEER [23] and RINA [22] compared to the current approach. Fig.10 shows that as the number of nodes rises, the half number of nodes that died round arrived earlier due to higher network energy consumption. Fig.11 shows all nodes died round. When the number of nodes grows, the lifetime is reduced due to the rise in packet dropping rate. With the proposed approach, the lifetime of the network is greatly extended.

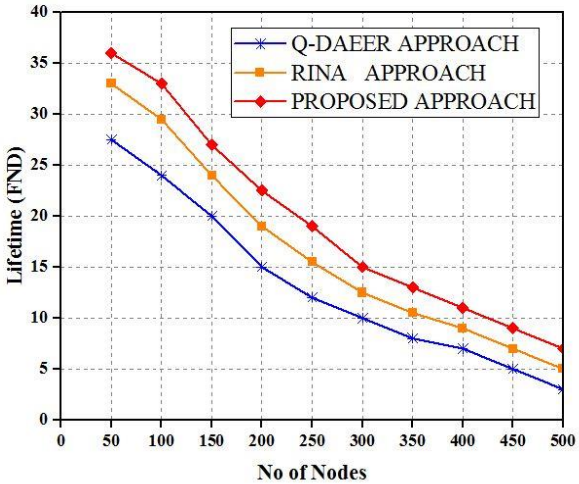

Fig.9. Comparison of network lifetime using fnd round

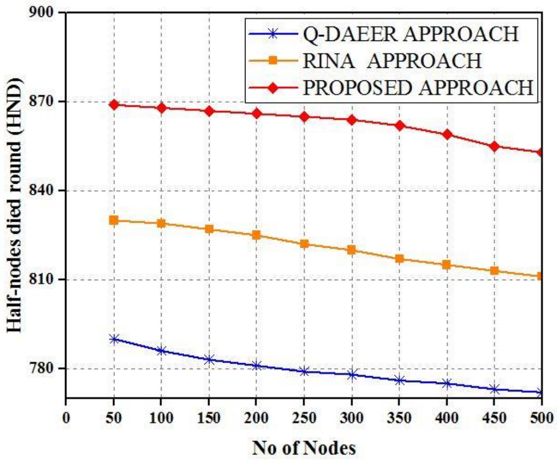

Fig.10. Comparison of network lifetime using hnd round

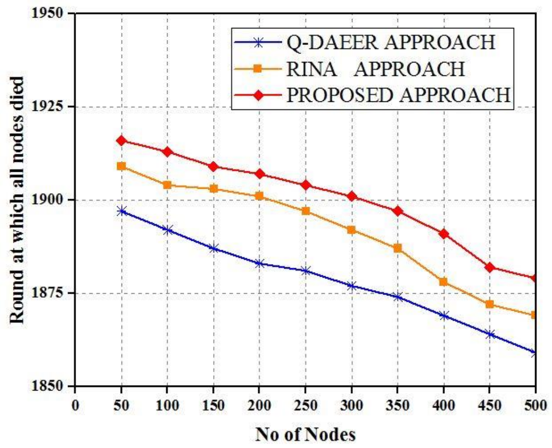

Fig.11. Comparison of network lifetime using lnd round

5. Conclusions

As wireless sensor networks are infrastructure-less and resource constrained, most of the works are dynamic in nature. Also, the parameters associated with the nodes changes dynamically based on the work progress. In order to take routing decisions in these scenarios, machine learning approaches are best suitable. Moreover, uncertainties in the input parameters are best resolved using Fuzzy Inference System. The proposed work considered the combination of a Fuzzy Inference based Q-Learning approach to finding the optimal path to enhance energy effectiveness and data delivery reliability. The Q-Learning technique has been used to find the optimal routing paths. It selects the best next-hop nodes using the rewards of nodes. The fuzzy inference system computes the reward (fuzzy probability value) using node characteristics such as residual energy, link quality, free buffer, and distance. The proposed approach helps to reduce packet dropping rates, thus further minimizing packet delivery latency and energy consumption. It is proved from the simulation that it has improved the network's lifetime by up to 55%. Future research could develop it further by concentrating on movable and rechargeable network nodes.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci,Wireless sensor networks: a survey, Computer Networks, Volume 38, Issue 4, 2002, Pages 393-422. https://doi.org/10.1016/S1389-1286(01)00302-4.

[2] Ian F. Akyildiz, Ismail H. Kasimoglu, Wireless sensor and actor networks: research challenges, Ad Hoc Networks, Volume 2, Issue 4, 2004, Pages 351-367. https://doi.org/10.1016/j.adhoc.2004.04.003

[3] U.Hariharan, K. Rajkumar, T. Akilan, A. Ponmalar, A multi-hop protocol using advanced multi-hop Dijkstra's algorithm and tree based remote vector for wireless sensor network, Journal of Ambient Intelligence and Humanized Computing, Volume 14,Issue 16,2023,Pages 6877–6895. DOI:10.1007/s12652-021-03548-4

[4] S. Gudla, N.R. Kuda, Learning automata-based energy efficient and reliable data delivery routing mechanism in wireless sensor networks, Journal of King Saud University- Computer and Information Sciences, Volume 34, Issue 8, September 2022, Pages 5759-5765. https://doi.org/10.1016/j.jksuci.2021.04.006

[5] P. Neamatollahi, M. Naghibzadeh, S. Abrishami, Fuzzy Based Clustering-Task Scheduling for Lifetime Enhancement in Wireless Sensor Networks, IEEE Sensors Journal , Volume: 17, Issue: 20, 15 October 2017, Pages: 6837 – 6844. DOI: 10.1109/JSEN.2017.2749250

[6] Q. Ni, Q. Pan, H. Du, C. Cao, Y. Zhai, A Novel Cluster Head Selection Algorithm Based on Fuzzy Clustering and Particle Swarm Optimization, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Volume: 14, Issue: 1, 01 Jan.-Feb. 2017, Pages: 76 – 84, DOI: 10.1109/TCBB.2015.2446475

[7] Baranidharan Balakrishnan and Santhi Balachandran, FLECH: Fuzzy Logic Based Energy Efficient Clustering Hierarchy for Nonuniform Wireless Sensor Networks, Hindawi Wireless Communications and Mobile Computing, Volume. 2017, 13 pages, 2017. https://doi.org/10.1155/2017/1214720

[8] R. Yousaf, R. Ahmad, W. Ahmed, A. Haseeb, Fuzzy Power Allocation for Opportunistic Relay in Energy Harvesting Wireless SensorNetworks,IEEEAccess,Volume:5,29-August-2017,Pages 17165 –17176,DOI: 10.1109/ACCESS.2017.2743063

[9] Zhang Siqing,Tao Yang ,Yang Feiyue, Fuzzy Logic-Based Clustering Algorithm for Multi-hop Wireless Sensor Networks, Volume 131, 2018, Pages 1095-1103. https://doi.org/10.1016/j.procs.2018.04.270

[10] S.K. Mothku, R.R. Rout, Adaptive Fuzzy-Based Energy and Delay-Aware Routing Protocol for a Heterogeneous Sensor Network, Journal of Computer Networks and Communications, Volume 2019, Article ID 3237623, https://doi.org/10.1155/2019/3237623

[11] Neha Singh, Deepali Virmani, Xiao-Zhi Gao, A Fuzzy Logic-Based Method to Avert Intrusions in Wireless Sensor Networks Using WSN-DS Dataset, International Journal of Computational Intelligence and Applications ,Volume. 19,Issue 3,2020 DOI: 10.1142/S1469026820500182.

[12] V. Rajaram, N. Kumaratharan, Multi-hop optimized routing algorithm and load balanced fuzzy clustering in wireless sensor networks, Journal of Ambient Intelligence and Humanized Computing,Volume12, Pages 4281–4289, 2021.

https://doi.org/10.1007/s12652-020-01827-0.
[13] Stephan, T., Sharma, K., Shankar, A., Fuzzy-Logic-Inspired Zone-Based Clustering Algorithm for Wireless Sensor Networks,International Journal of Fuzzy Systems , Volume 23, Pages 506–517, 2021. https://doi.org/10.1007/s40815-020-00929-3.
[14] S. Phoemphon, C. So-In, P. Aimtongkham, T.G. Nguyen,An energy-efficient fuzzy-based scheme for unequal multi-hop Clustering in wireless sensor networks, Journal of Ambient Intelligence and Humanized Computing, Volume 12,2021,Pages 873–895.https://doi.org/10.1007/s12652-020-02090-z
[15] Giri, A., Dutta, S. & Neogy, S., An Optimized Fuzzy Clustering Algorithm for Wireless Sensor Networks, Wireless Personal Communications, Volume 126, Pages 2731–2751, 27 June 2022. https://doi.org/10.1007/s11277-022-09839-z.
[16] D. Praveen Kumar, Tarachand Amgoth, Chandra Sekhara Rao Annavarapu, Machine learning algorithms for wireless Sensor networks:A survey,Information Fusion, Volume 49, September-2019, Pages1-25, https://doi.org/10.1016/j.inffus.2018.09.013
[17] Z. Mammeri, Reinforcement Learning Based Routing in Networks: Review and Classification of Approaches, IEEE Access , Volume: 7, 29 April 2019, Pages 55916 – 55950, DOI: 10.1109/ACCESS.2019.2913776
[18] Farzad Kiani, Reinforcement Learning based Routing Protocol for Wireless Body Sensor Networks, 7th International Symposium on Cloud and Service Computing., Pages 22-25, November 2017. DOI: 10.1109/SC2.2017.18
[19] W. Guo, C. Yan, T. Lu, Optimizing the lifetime of wireless sensor networks via reinforcement-learning-based routing, International Journal of Distributed Sensor Networks , Volume 15, Issue 2, February 26, 2019, https://doi.org/10.1177/1550147719833541
[20] D. Kim, T. Lee, S. Kim, B. Lee, H.Y. Youn, Adaptive packet scheduling in IoT environment based on Q-learning, Journal of Ambient Intelligence and Humanized Computing, Volume 11, June 2000, Pages 2225–2235. https://doi.org/10.1007/s12652-019-01351-w
[21] Q. Yang, S. Jang, S. Yoo, Q-Learning-Based Fuzzy Logic for Multi-objective Routing Algorithm in Flying Ad Hoc Networks, Wireless Personal Communications,Volume 113,July 2020, Pages115–138, https://doi.org/10.1007/s11277-020-07181-w.
[22] R.Maivizhi, P. Yogesh, Q-learning based routing for in-network aggregation in wireless sensor networks, Wireless Networks Volume 27, April 2021, Pages 2231–2250, https://doi.org/10.1007/s11276-021-02564-8
[23] W.K. Yun, S.J. Yoo, Q-Learning-Based Data-Aggregation-Aware Energy-Efficient Routing Protocol for Wireless Sensor Networks, IEEE Access , Volume: 9, 13 January 2021, Pages10737 – 10750, DOI: 10.1109/ACCESS.2021.3051360
[24] Tripti Sharma , Archana Balyan , Rajit Nair , Paras Jain , Shivam Arora , and Fardin Ahmadi, ReLeC: A Reinforcement Learning-Based Clustering-Enhanced Protocol for Efficient Energy Optimization in Wireless Sensor Networks, Hindawi Wireless Communications and Mobile Computing, Volume 2022, Article ID 3337831, 16 Pages, https://doi.org/10.1155/2022/3337831

Authors' Profiles

Gudla Sateesh is an Associate Professor in Computer Science and Engineering at the Lendi Institute of Engineering and Technology. He taught several subjects in UG and PG courses for over 18 years. His research and publication interests include Wireless Sensor Networks, the Internet of Things, Machine Learning, and Data Science. He has published 22 papers in various peer-reviewed journals. He received his M.Tech in Computer Science and Technology in 2008 from Andhra University and is pursuing a Ph.D. from JNTUK in Wireless Sensor Networks.

Dr. Kuda Nageswara Rao is working as a Professor at Andhra University, Visakhapatnam. He has 25 Years of Teaching, Research, and administrative experience. He received his Ph.D. in Computer Science and Engineering from JNTU Kakinada in 2010, M.Tech in Computer Science and Technology in 2001 from Andhra University, and B.E in Electronics and Communication Engineering from GITAM, Andhra University in 1998. He has published more than 60 research papers in various international peer-reviewed journals. He has organized and attended many International and National conferences, presented papers, chaired sessions, and delivered invited talks. His area of research includes Wired and Wireless Networks, Internet Technologies, Cyber Security, Cloud Computing, IoT, and Software Engineering. Under his esteemed guidance, more than ten scholars were Awarded Ph.D. degrees.

How to cite this paper: Sateesh Gudla, Kuda Nageswara Rao, "Reliable Data Delivery Using Fuzzy Reinforcement Learning in Wireless Sensor Networks", International Journal of Computer Network and Information Security(IJCNIS), Vol.15, No.4, pp.96-107, 2023. DOI:10.5815/ijcnis.2023.04.09