problems in WSN using RL approach. Initially, the RLBCA algorithm is designed to remand energy consumption. On the other hand, ODMST has been designed to aggregate data as per requirement from the CH. The record of incoming requests has been stored into a table, and according to the recorded data next step has been performed [38].

3. Materials and Methods

The entire network is divided into two sections that are clustering and Route formation using RL based approach.

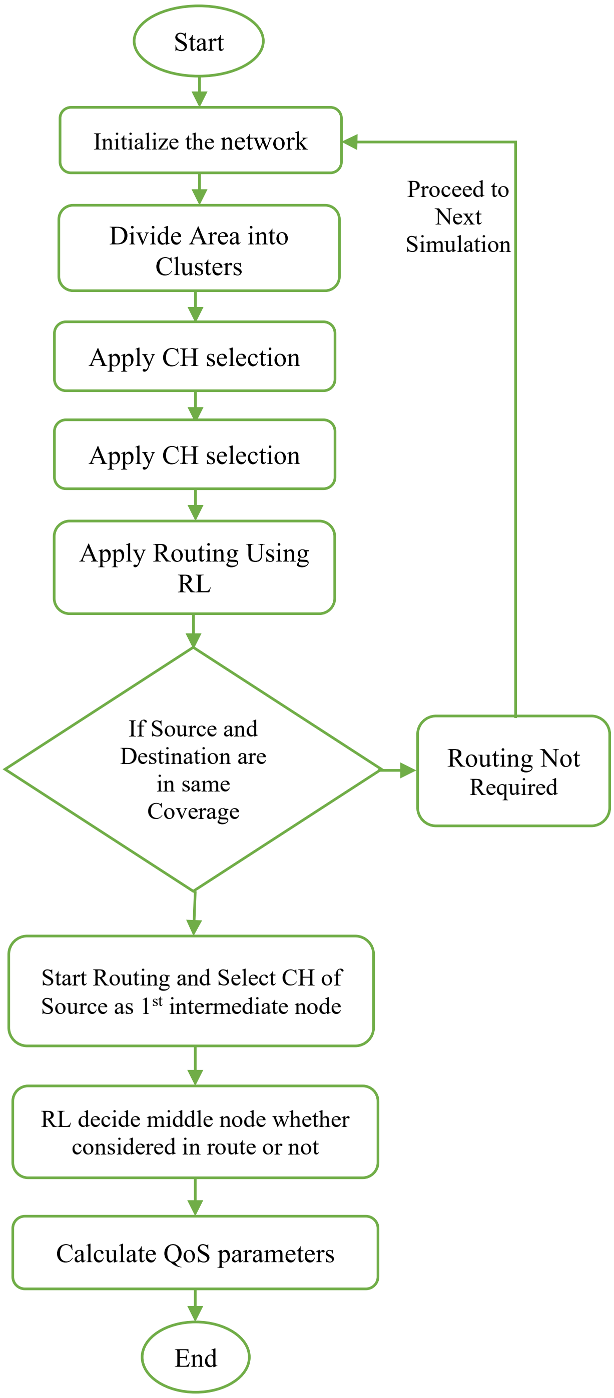

Figure 6. Flow of Proposed Work

Clustering is the process of dividing the entire network into different zones known as clusters.

Each cluster is monitored by a CH node, which collects and forwards data to the BS. In a large network area, the selection of CH becomes an important issue.
Since, with the increase in the network size distance between the CH and BS increases and hence decrease nodes lifetime and the failure of CH, it increases the network overhead. Also, the process of route formation has been done through the RL algorithm using the neural network concept. The overall flow is shown in Figure 6
An RL based energy efficient LEACH routing protocol (RL-LEACH) is designed to select appropriate CH based on the revert Q point.

3.1. Clustering

Initially, N numbers of nodes are deployed in the network with a dynamic size of the network. Here five different node variation with network size has been considered, which are presented in Table 2.

Table 2. Network Considerations

Network Number	Number of deployed Nodes	Network Area Range
Net-1	100	1000×1000 to 3000×3000
Net-2	200	
Net-3	300	
Net-4	400	

The nodes are deployed with defined properties such as energy, delay, co-ordinates, and collision rate. The algorithm followed for the node deployment is written below:

Algorithm 1: Node Deployment

WSN = Node Deployment (Nodes, N_{height}, N_{Width})

Where,
Nodes → Number of Nodes
N_{height} → Height of the WSN
N_{Width} → Width of WSN
WSN → A simulation area of WSN

Start
N_{height}=1000, 1200, 1400, 1600
N_{Width}=1000, 1200, 1400, 1600
Define N number of Nodes for the simulation of Network
For iin range of (1 to N)
Plot_Node (i)=coordinate (X, Y)
    Define Node = N (i)
Source_Node = random (N)
Destination_Node= random (N)
If Source_Node (S_N) == Destination_Node (D_N)
Source_Node = rand (N)

Destination_Node= rand (N)
Else
S_N=S_N
D_N=D_N
End –If
Define S_N as a source
Define $D_N$as destination
End - For
Return: WSN
End – Function

- **Sub Division of Network**

The entire network is sub-divided into a number of clusters, and the node having the highest energy behaves as a CH.For example, if we partitioned the network of dimension 1600×1600, and the size of the cluster is 400×400, then the network is partitioned into 16 clusters, each consists of unique ID as presented in Table 3.

Table 3. Cluster Identification for 1600×1600 Network Area

ID-1	ID-2	ID-3	ID-4
ID-5	ID-6	ID-7	ID-8
ID-9	ID-10	ID-11	ID-12
ID-13	ID-14	ID-15	ID-16

The CH's that have been formed in the network area are determined using equation (8)

$$T_{CH} = \sqrt{\sum_{i=1}^{N} \frac{d_{N-Sink} \times \ln(N)^{I_A}}{\frac{A}{I_A}}} \quad (8)$$

$I_A \rightarrow$ Attraction index and is calculated using Equation (9)

$$I_A = \begin{cases} 2 & \text{if } \frac{N}{\text{Total Area}} < .10 \ \& \ N > 50 \\ 1 & \text{Otherwise} \end{cases} \quad (9)$$

Using T_{CH} and I_A, the selection of Q points for RL learning has been decided.

3.1. RL Learning

RL collects information by continuously interacting with the surrounding environment and improving the performance of the system by achieving optimal results by performing all the necessary actions to draw conclusions. Q-learning is a category of RL scheme that creates a sequence of observations in terms of reward point. The visualization of the RL approach is illustrated in Figure 7.

The RL model presented in Figure 7 can be implanted inside each sensor node or on the outer surface of the sensor node. For example, each sensor node holds a reward point associated with an adjacent sensor node and each network point in the surrounding environment. In WSN, its main use is to learn about the route and hence provide the best route for data transmission. State sensing is used to indicate the action performed by the destination node. The action indicates the next communicating node to which data is to be forwarded. The reward point indicates the distance between the CH and the destination node. The increase in the reward point decreases the distance of CH towards the destination node and hence increase the network performance [39].

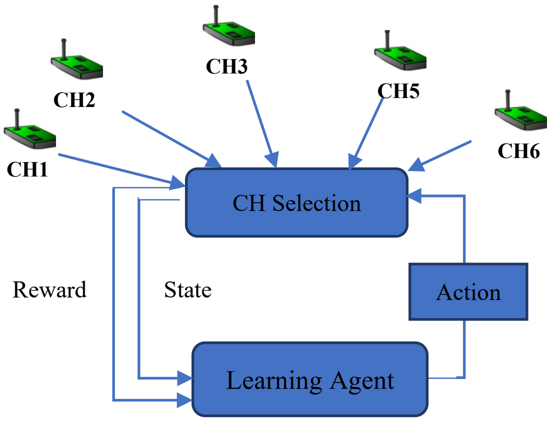

Figure 7. RL Learning Approach

RL-based routing protocol has been used to optimize the life of WSN networks in all specific aspects. Using this technique, the next CH is selected based on historical learning and available evaluated information, and properties of nodes like residual energy, transmission distance from CH to BS, and delay are taken into consideration to learn the best ways. In this way, the correct linkage between the sensor nodes and BS is performed with balanced energy consumption.

Q learning approach is used in the RL approach in which agent monitored the WSN environment of state (U). Accordingly, there is an action (C), and reward points are (P) are generated. Based upon the experience, the strategy is modified. The action with the highest reward point is selected for a certain state. In case if the transition probability is unknown, the next node selection using simple RL approach is not possible. Therefore, Q learning approach is provided, which calculates the accurate reward point and hence decides the best path [40]. Let the state of each node is represented by Q(u,c), which is being updated after every iteration by using equation (10)

$$Q(u_t, c_t) = (1-\alpha)Q(u_t, c_t) + \alpha\left[p_{t+1} + \beta \max_{\forall c_{t+1}} Q(u_{t+1}, c_{t+1})\right] \quad (10)$$

$Q(u_t, c_t) \rightarrow$ Accumulated reward generated by the node at state u, action c, which is taken at time t.

$p_{t+1} \rightarrow$ immediate reward point generated for action c, when the state transition takes place from stateU_t to U.

$\forall c_{t+1} \rightarrow$ action performed in the set of action $U_{t+1}(C)$

$\alpha \rightarrow$ learning factor

β→rate of discount lies between [0,1]

Algorithm: CH Selection using RL approach

Route = RL (Nodes, S, D, CHs, $E_{CONSUPTION}$, C_{AREA}, Distance)

Where,
Nodes → Number of Nodes
S → Source Node
D → Destination Node
CHs All Cluster Heads
$E_{CONSUPTION}$ Energy Consumption of Nodes
C_{AREA} Coverage Area of Nodes
Distance →Distance of node from neighboring nodes
Route →Route from S to D via CHs

Start
Route=[]
The starting point of Route is Source Node so, Route(1)=S
Define a flag for the destination, D_{FOUND}=0
While $D_{FOUND}\neq 1$
    Reward=[] // To store the current CH reward points
    For I in range of S(C_{AREA})
        Reward(I, 1)= $E_{CONSUPTION}$ (S(C_{AREA}, I))
        Reward(I, 2)= C_{AREA} (S(C_{AREA}, I))
        Reward(I, 3)= Distance (S(C_{AREA}, I))
    End – For
Max Reward = [Reward(I, 1) Reward(I, 2) Reward(I, 3)]
    If Max Reward ==True
        Route (Next Nodes)= CH (Max Reward)
    Else
        Selected Next Node in S(C_{AREA})
    End – If
    If D_{FOUND} = 1
        Route (Next Nodes)= D
    Else
        Repeat Process
    End – if
End – While
Return: Route as a path from S to D via CHs
End – Function

3.3. Proposed Reinforcement learning-based LEACH protocol (RL-LEACH)

Using the RL-LEACH approach, the selection of CH in WSN has been performed by considering three factors, such as Energy consumption, coverage area, and distance of CH from BS. Initially, the energy level of each node is compared with the defined threshold; if energy is higher, then the selected nodes' communication range is checked and then the distance is checked. Based on the QoS parameters, reward points are generated. The nodes having a higher reward point is selected as CH.

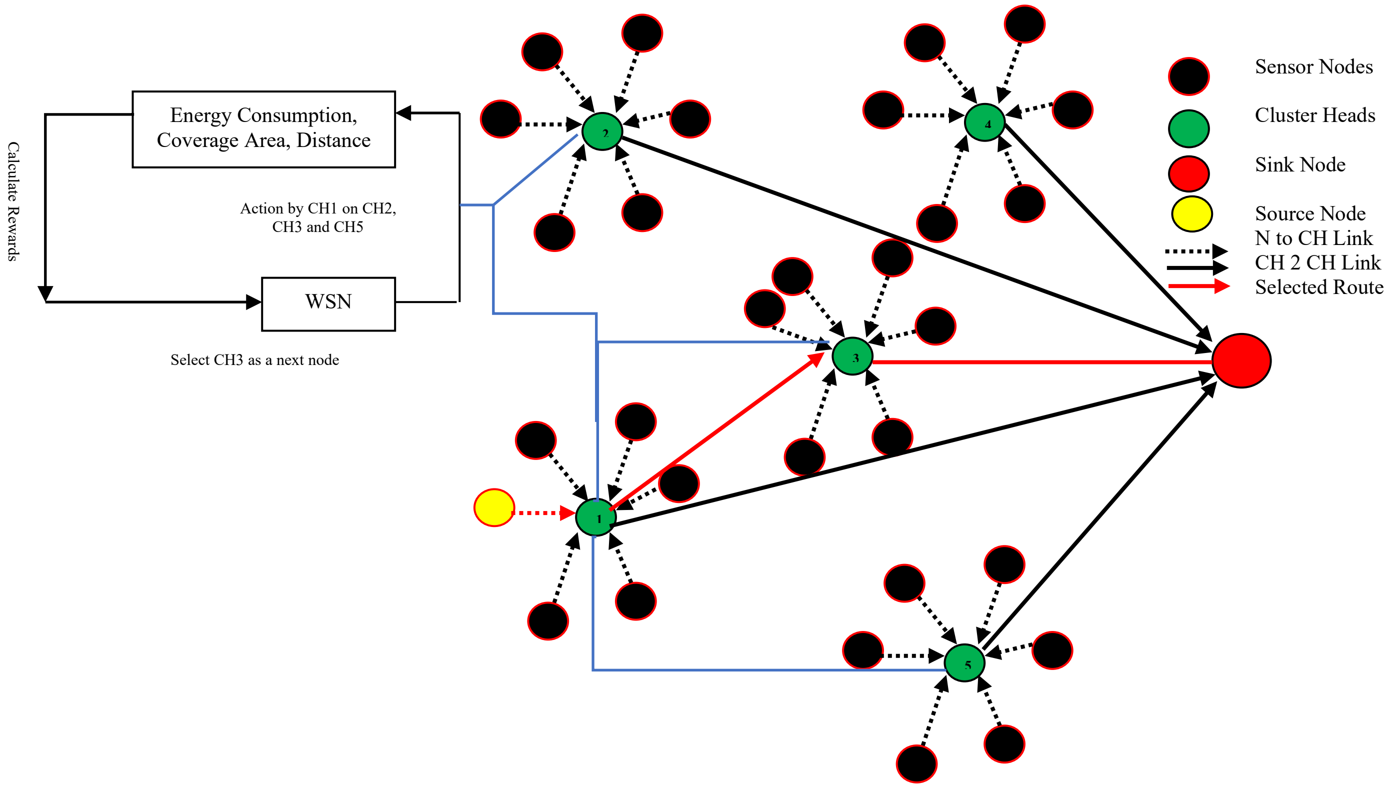

Figure 8. Proposed RL-LEACH approach Working scenario

Figure 8 describes the working of the RL-LEACH routing protocol that is used to select the CH from nodes based on their basic properties like Energy Consumption, Coverage Area, and Distance from Sink node. In the figure, 10 nodes are selected as CH (CH1, CH2……CH5). Source node wants to transmit the data packets, then CH1 is selected as a next intermediate node after that RL decides which CH nodes are selected as the next intermediate nodes. The selection of CH is performed based on the written algorithm of RL-LEACH:

OCHs= RL-LEACH (Nodes, Nodes-Prop)

Where,

Nodes → Number of Nodes

Nodes-Prop → Nodes properties like Energy Consumption, Coverage Area and Distance from Sink

OCHs→Optimized Cluster Heads(CHs) having better efficiency to monitor entire cluster

1 **Start**

2 During cluster formation, when the nodes required being CH, it may select an arbitrary random number (rNum) among zero and one.

3 Calculate threshold, $T_{new}(m) = \begin{cases} \frac{Prob}{1-Prob\left|mod\left(\frac{Round}{Prob^{-1}}\right)\right|} \times \frac{E_{RESIDUAL}}{E_{INITIAL}} & if\ m \in Grp \\ 0 & else \end{cases}$

   a. Where, m is number of nodes, $E_{RESIDUAL}$ is residual energy of nodes and $E_{INITIAL}$ is initial energy for nodes being generated during the initialization of WSN.

4 Call and Apply RL for CH selection

5 Reward=[] // To store the reward points of Nodes

6 **For I in range of m**

7 Reward(I, 1)= Nodes-Prop ($E_{CONSUPTION}$ (I))

8 Reward(I, 2)= Nodes-Prop (C_{AREA} (I))

9 Reward(I, 3)= Nodes-Prop (Distance (I))

10 **End – For**

11 Max Reward = [Reward(I, 1) Reward(I, 2) Reward(I, 3)]

12 **If rNum<$T_{new}(m)$ and Max Reward = True then,**

13 OCH = Node not selected as CH

14 **Else**

15 OCH = Node selected as CH

16 **End – If**

17 **Return:** Optimized CHs = $Node_{CH}$

18 **End– Function**

The protocols are tested under different topologies, which are generated randomly. The performance parameters are discussed later. The designed network with N number of nodes is shown in Figure 9.

The Q based learning has been implemented in MATLAB and the designed network for an area of 1600×1600, which is divided into 16 different clusters {ID1, ID2, ID3…………….ID16}, the identification of each clusters is represented in table 3. The pictorial representation is shown in Figure 9.

As shown in Figure 10, the source "S" is defined in the cluster (ID1) that wants to transmit data to destination "D" located in cluster ID 15. Now the route is created based on

4. Results and Discussions

This section describes the efficiency of the proposed RL-LEACH algorithm in terms of throughput, Packet Delivery Ratio (PDR), and energy consumption.

The proposed algorithms select optimal CH using the RL approach and enhance the lifetime by reducing the consumption of energy. In the simulation process, we have considered nodes (N=50, 100, 150, 200) deployed in a different area of (A=1000×1000, 1200×1200, 1400×1400, and 1600×1600) square meters. Initially, the parameters of network nodes such as energy consumed by each node, transmission delay, and collision rate and co-ordinated of each node are taken into consideration. The performance has been examined using simple LEACH, RL based Q learning approach by considering the same parametric values of the nodes. The detailed simulation parameters are presented in Table 4.

Table 4. Simulation Parameters for RL-LEACH protocol

Parameters	Value
Number of Nodes	100,200,300,400
Number of the destination node	1
Area of network	1000×1000
	1200×1200
	1400×1400
	1600×1600
Initial Energy of sensor nodes	2 J
Deployment mode of sensor nodes	Random
Transmission range	20 m
Sensing radius	10m
Packet Size	512 bits
Distance Threshold	12m

the directional orientation of destination node (D); that is, the data traveled from the source node (S) considered the path in the direction of its destination node as shown in figure 10. Source node 'S' have four nearby clusters in its coverage range named as ID2, ID 3, ID4 and ID5, each having different CH's such as CH1, CH2, CH3, CH4 and CH5 respectively. Now, to which CH, the S node forward data, depends upon the reward point generated by RL approach. For CH1, CH2, CH3, CH4 and CH5, the generated reward point is depicted in Table 6. According to that, the reward point of CH5 is higher compared to the other four CHs. Hence, S node forwards data to CH5. Now, CH 5 also, checks its nearby clusters that are CH7, CH9, CH10, and CH12. RL approach is applied to generate a reward point based on energy consumption, coverage area, and distance. Since, the reward point for CH12 is highest among all, hence CH5 forward data to CH12. Again, the same process is repeated by CH 12 node, which finds that CH15 have the highest reward point and passes data to CH15 node, which is the destination node. In this way, data travels from the S node to the D node.

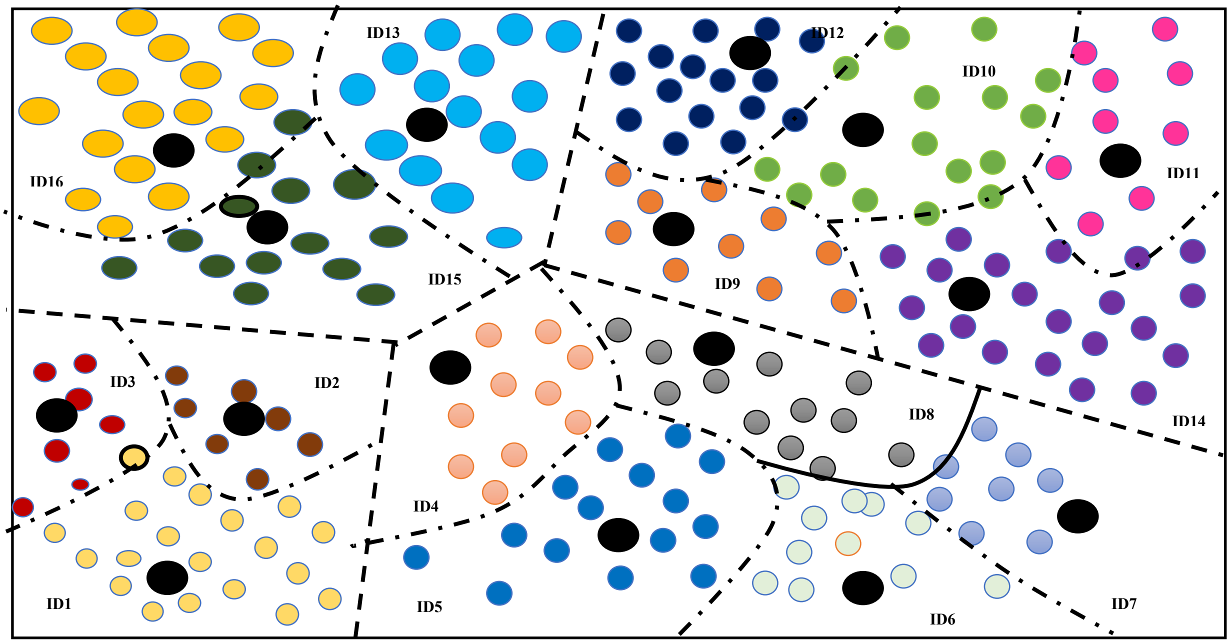

Figure 9. Clustering in WSN

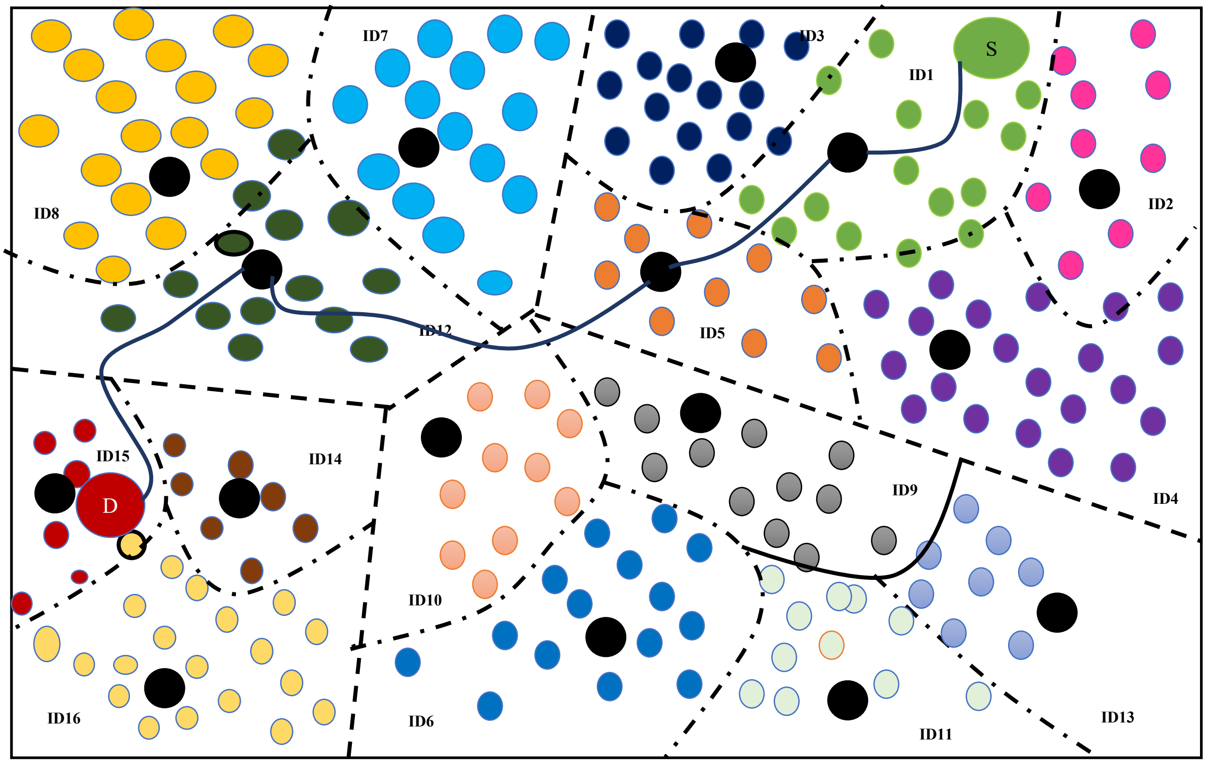

Figure 10. Route creation using R-LEACH

During Q based RL learning, each cluster member nodes pass data to their CH, which is in its communication range as mention in Table 4. The data is passed to the single CH instead of checking each and every cluster. In this way, an optimal cluster is selected, which consumed less energy and also near to the BS.

The simulation has been performed in which both the member nodes and the CH nodes are participated and examined the energy consumed by both nodes with different cluster sizes in the network. The examined network energy consumption value for 16 different clusters formed in Figure 10 is presented in Table 5.

Table 5. Network Energy Consumption Vs. Cluster Size

Cluster size	Network Energy Consumption (j)
1	512
2	504
3	420
4	435
5	424
6	412
7	408
8	405
9	409
10	406
11	407
12	405
13	404
14	406
15	408
16	442

As shown in Figure 11, the network energy consumption for fixed number of nodes deployed in each cluster (ID1, ID2……ID16), for small and large cluster size network, the consumed energy is high. For cluster size one in network are 1600× 1600, the consumed energy is 512 J, whereas for large cluster size of 16, the consumed energy is 442 J. This is because, for small cluster size network, the cluster member nodes have to communicate with other nodes that are placed at long distance and hence consume high energy due to the intracluster distance communication.

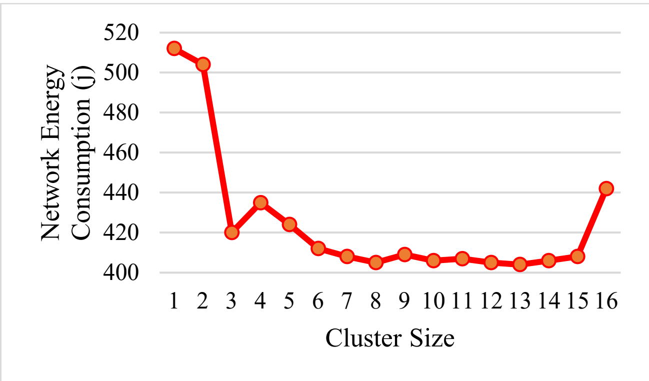

Figure 11. Network Energy Consumption Vs. Cluster Size

On the other side, with large cluster size (see cluster size=14), the energy consumption is low. But, in this case, with the increase in the cluster size, the communication between inter CH increases, which consumes more energy than cluster members. Therefore, it is necessary to select the appropriate size of clusters in the network so that energy consumption can be minimized. Here, this is performed by selecting appropriate CH and hence balance the energy consumption for both inter and intracluster communication in WSN. The graph depicted in Figure 11 shows that the optimal cluster size, where the energy consumption is minimum, is 8.

To obtain an optimal solution, we applied Q-learning to examine the performance of the proposed algorithm in terms of learning and adaptation to a dynamic environment. The CH selection using Q-learning is performed due to its quick convergence for shorter spam. The performance of the RL approach has been determined by computing the reward point for different member nodes in the CH based on the energy consumed and distance of the CH to the BS.

Route = [S CH1 CH5 CH12 CH15 D]

Here, S is source node and D is the destination node

To create a route from S to D, the RL approach is used for better path selection by avoiding the affected node within the route as an intermediate node. During the transmission, S transfer the data packets to own cluster head (CH1), and then now CH1 searches next nodes as CH5 based on the reward points using the RL technique. The selection process of CH5 as a next intermediate node is shown in the figure10. The corresponding reward point based on energy consumption, coverage area, and distance are listed in Table 6.

Table 6. Reward Point for Scenario-1

Reward Point			
CH2	CH3	CH4	CH5
0.4	0.83	0.85	0.9
0.8	0.84	0.87	0.94
0.6	0.83	0.93	0.98

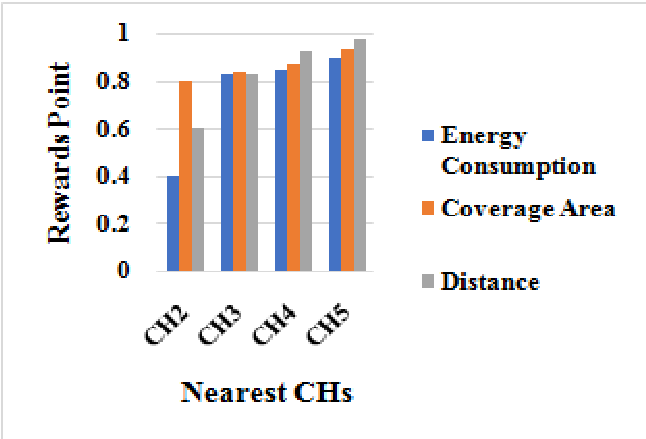

Figure 12. Reward point for Energy Consumption, Coverage Area, and Distance (Scenario-1)

In 2nd Scenario, route from S to D, using RL approach by considered CH5 as cluster head is performed. Let in the coverage area of CH5, clusters CH7, CH 9, CH10 and CH 12 occurs. Now, to which CH, the data is transmitted by the selected CH5 depends upon the reward point generated by RL approach as listed in Table 7 and graphically shown in Figure 13.

Table 7. Reward Point for Scenario-2

Reward Point			
CH7	CH9	CH10	CH12
0.52	0.88	0.87	0.92
0.91	0.91	0.9	0.95
0.73	0.93	0.96	0.98

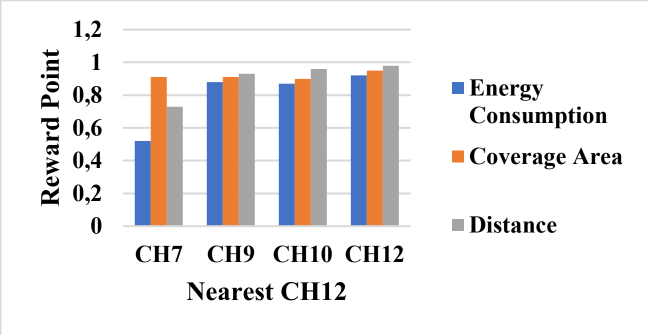

Figure 13. Reward point for Energy Consumption, Coverage Area, and Distance (Scenario-2)

The reward point of CH12 for all three considered parameters that are energy consumption, coverage area, and distance is high compared to the reward point for CH7, CH9, CH 10. Therefore , CH5, forward data to CH 12. Now, CH12, also check the reward points to its nearby clusters and pass data to only those CH having the highest reward point. Let, CH8, CH14, CH16, CH15 comes in the coverage range of CH12 cluster. Now, the reward point generated based on the three parameters as energy consumption, coverage area, and distance are checked. From Figure 13, it is clearly seen that the reward point for CH15 is highest, which is the destination in our example. Thus, CH12 passes data to CH15 as the destination node. By using the above strategy, the data from the source CH 1 is reached to destination CH15.

Table 8. Reward Point for Scenario-3

Reward Point		
CH8	CH14	CH15
0.55	0.68	0.84
0.87	0.9	0.98
0.78	0.82	0.89

After the selection of CH15 as the next intermediate nodes, CH15 checked its own coverage area and founded the destination (D).

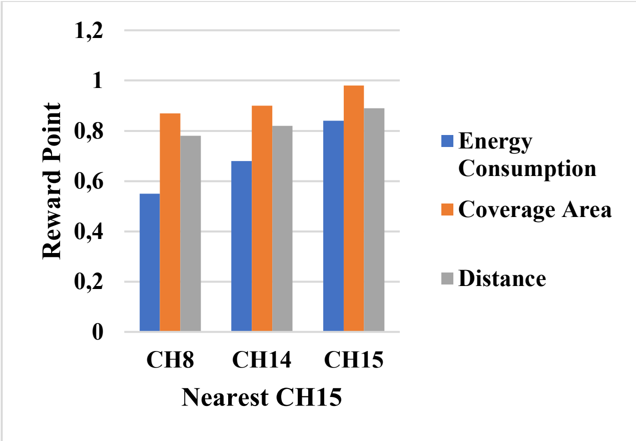

Figure 14. Reward point for Energy Consumption, Coverage Area, and Distance (Scenario-3)

4.1 Simulation Results

After the completion of route formation, data transmission takes place, and QoS parameters are calculated for the designed model using both approaches like RL and without RL. The computed values for network lifetime with respect to several nodes for three different scenarios that are (i) without RL, (ii) with RL, and (iii) Kiani et al. is presented in Table 9.

Table 9. QoS for 100 Nodes

NETWORK AREA (SQUARE METER)	NETWORK LIFETIME (MIN)			PDR			ENERGY CONSUMPTION (MJ)		
	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.
1000×1000	10.257	19.2685	13.7586	77.3258	89.7582	Kiani et al.	324.245	300.1245	322.157
1200×1200	9.8452	20.369	12.3658	78.2698	88.2672	80.1425	328.167	304.257	326.547
1400×1400	9.2578	18.2574	13.7569	80.269	86.328	80.3651	330.147	307.264	328.164
1600×1600	8.2567	17.36587	14.3658	82.365	85.742	82.2569	322.157	305.264	320.754
1800×1800	7.2594	19.0257	15.6987	85.0369	86.120	84.0142	320.157	312.0147	317.267
2000×2000	10.2575	17.8954	12.0875	75.0369	84.367	87.367	324.561	315.627	322.751

2200×2200	11.3678	20.367	13.6758	74.9851	87.0257	79.258	333.157	317.845	328.975
2400×2400	9.5678	12.5742	17.0258	76.358	86.3975	84.217	326.487	317.859	324.185
2600×2600	8.6572	13.6571	16.0987	74.128	87.364	78.265	329.175	314.751	327.154
2800×2800	6.3587	15.6987	15.0365	73.986	89.367	74.219	330.654	324.351	328.175
3000×3000	9.5687	20.3698	14.2581	76.285	81.0236	72.369	332.0145	327.698	329.457

Table 9 represents the values of network lifetime, PDR, and energy consumption computed for three scenarios (i) without RL, (ii) with RL, and (iii) Kiani et al. From the above-obtained values, it has been examined that proposed approach RL-LEACH outperformed among all and the average computed values for network lifetime without RL, RL-LEACH and Kiani et al. for 100 nodes are 9.15, 17.71, and 14.37 respectively. Similarly, for PDR, the average values for without RL, RL-LEACH, and Kiani et al. 77.64, 86.52, and 80.15, respectively. In the same way, the average values computed for energy consumption are 327.35m J, 313.36mJ, and 325.05mJ, respectively. The average network lifetime examined for the proposed work in contrast to the existing work is 24.34 min, 19.2min, respectively. Thus, there is an enhancement of 26.77 % has been obtained. Similarly, for the PDR, the average PDR for RL-LEACH and Kiani et al. are 88.87 and 86.78, respectively. In the same way, the average energy consumption for RL-LEACH and Kiani et al. are 308.09 mJ, 319.05 mJ, respectively. Therefore, improvement in PDR and energy consumption compared to Kiani et al. is 2.41 % and 3.44 %, respectively.

Table 10. QoS for 200 Nodes

NETWORK AREA (SQUARE METER)	NETWORK LIFETIME (MIN)			**PDR**			**ENERGY CONSUMPTION (MJ)**		
	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.
1000×1000	15.1253	24.3657	20.0325	79.2475	90.457	88.3687	324.236	300.0214	320.254
1200×1200	17.2548	23.0147	18.2571	78.365	91.3265	87.2547	324.365	304.125	318.257
1400×1400	18.3697	22.03147	20.365	77.248	94.0325	85.6472	327.0215	301.025	319.754
1600×1600	20.3651	25.0367	21.0365	76.258	89.265	86.3657	328.0642	304.025	318.657
1800×1800	24.3657	26.3574	25.0364	75.269	87.0257	87.3694	325.842	304.657	317.254
2000×2000	14.0236	27.3658	16.0327	74.2895	88.365	88.2571	327.168	307.265	316.85
2200×2200	17.0369	23.157	18.0267	73.657	90.257	89.325	326.942	305.124	319.347
2400×2400	12.3651	22.1571	14.3657	72.314	91.365	84.3657	328.367	312.542	320.154
2600×2600	16.3282	21.2675	17.3684	71.028	88.257	85.2361	330.251	318.025	321.012
2800×2800	17.3658	28.0365	18.3672	70.258	84.0257	84.0321	327.845	315.261	318.674
3000×3000	21.3657	25.0367	22.3658	70.369	83.2481	88.365	327.942	317.025	319.257

Table 11. QoS for 300 Nodes

NETWORK AREA (SQUARE METER)	NETWORK LIFETIME (MIN)			PDR			ENERGY CONSUMPTION (MJ)		
	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.
1000×1000	25.125	32.321	30.1253	80.1273	92.1528	88.258	347.124	333.254	340.285
1200×1200	24.152	33.251	29.156	78.958	93.258	87.685	348.157	325.147	342.251
1400×1400	23.267	30.125	31.024	79.157	90.27	87.659	350.27	326.487	348.215
1600×1600	26.248	28.264	32.045	76.258	91.365	84.325	351.657	342.158	342.165
1800×1800	27.036	32.452	33.157	82.157	89.567	86.975	356.271	348.25	345.128
2000×2000	25.03	33.267	29.567	81.247	94.235	90.258	350.275	345.28	349.157
2200×2200	22.036	35.68	28.751	82.365	93.574	87.258	349.752	340.28	347.125
2400×2400	29.167	36.157	32.015	81.365	92.365	84.258	348.627	342.158	346.215
2600×2600	28.0176	31.254	30.672	80.25	91.257	89.365	345.742	341.258	344.852

2800×2800	22.015	29.157	30.758	79.268	89.368	88.147	350.21	346.758	348.752
3000×3000	21.425	27.265	29.458	78.154	90.257	90.258	352.487	349.785	350.275

Table 12. QoS for 400 Nodes

NETWORK AREA (SQUARE METER)	NETWORK LIFETIME (MIN)			PDR			ENERGY CONSUMPTION (MJ)		
	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.	Without RL	With RL/RL-LEACH	Kiani et al.
1000×1000	28.25	43.48	40.02	70.85	93.69	78.69	339.425	301.145	330.25
1200×1200	28.37	43.6	40.34	70.97	93.81	78.81	339.545	303.265	329.47
1400×1400	34.65	49.88	46.72	77.25	94.09	85.09	345.825	312.545	340.25
1600×1600	26.73	41.96	38.5	69.33	92.17	77.17	317.905	304.625	315.24
1800×1800	29.79	41.74	37.56	68.39	91.23	76.23	336.965	313.685	330.25
2000×2000	32.75	47.64	44.52	75.35	98.19	83.19	323.925	310.645	321.05
2200×2200	23.95	39.18	35.72	66.55	89.39	74.39	335.125	301.845	334.258
2400×2400	35.54	46.78	43.39	74.14	96.98	81.98	342.715	289.435	339.457
2600×2600	26.48	38.71	35.25	66.08	88.92	73.92	334.655	291.375	332.576
2800×2800	29.64	44.87	41.44	72.24	95.08	80.08	340.815	307.535	338.756
3000×3000	27.74	42.97	39.51	70.34	93.18	78.18	338.915	305.635	335.169

The results examined while deployed node count is N=200 nodes with different simulation area is depicted in Table 11. The proposed approach shows better results compared to existing work. The average percentage enhancement of network lifetime, PDR, and energy consumption of RL-LEACH protocol compared to the existing Kiani et al. is examined as 3.69 %, 5.69 %, and 1.67 %, respectively.

The analysed values for 200 nodes in different network area dimension are depicted in Table 12. It is clearly observed that the proposed approach with RL. RL-LEACH performed better compared to the existing approach. For 200 nodes, the percentage increase in the network lifetime, PDR, and energy consumption compared to existing work is 8.57 %, 18.32% and 8.36%.

4.2. Comparative Analysis

This section describes the comparative analysis for different nodes varies as N= 100, 200, 300, and 400 respectively for three designed scenarios such as Without RL, with RL and Kiani et al. the comparison has been performed on the basis of three parameters as Network Lifetime, PDR and energy Consumption.

Table 13. Network Lifetime

Number of nodes	Without RL	With RL	Kiani et al.
100	9	18	13
200	15	24	20
300	25	32	30
400	30	45	40

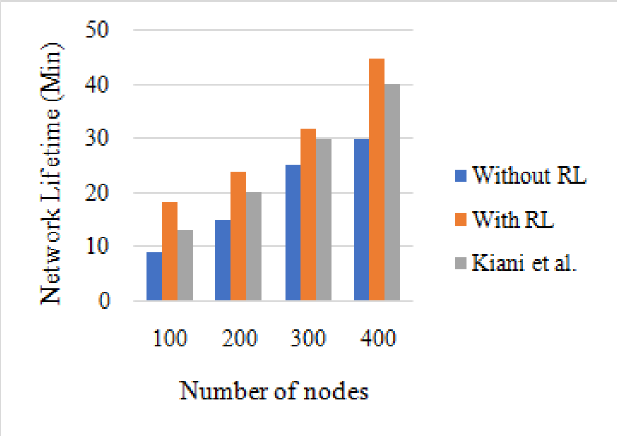

Figure 15. Network Lifetime (min)

The computed values of network lifetime for different number of nodes deployed in with various network size as depicted in Figure 15. From the figure it is clearly seen that with the increase in the number of nodes the network life increases. The comparative graph consists of three types of bar represented by the blue, the orange and the grey colour depicting the values of network life for without RL, with RL and Kiani et al. Among all, the proposed approach provides better network life. With the increase in the number of nodes (N), the rate of receiving data packet increases with the reduction of data packet transmission rate. This is due to the affecting nodes properties like as hop count, nodes failure, nodes density.

The examined values for second parameter that is Packet Delivery ratio (PDR) with respect to number of nodes for three different scenarios that are (i) without RL, (ii) with RL, and (iii) Kiani et al. is presented in Table 14.

Table 14. PDR

Number of nodes	Without RL	With RL	Kiani et al.
100	77	89	80
200	79	90	88
300	80	92	88
400	81	93	89

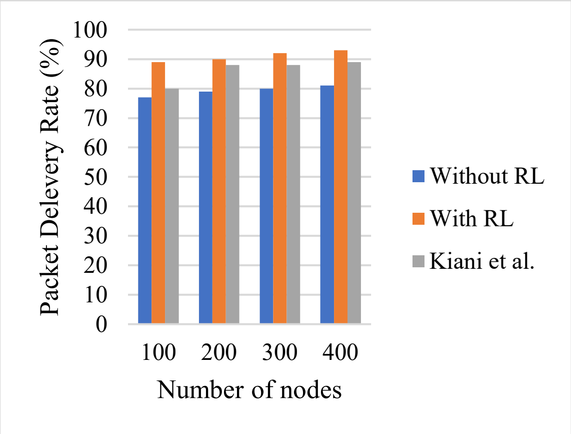

Figure 16. Packet Delivery Rate

Figure 16 compares the PDR for three different scenarios that are CH selection without RL, with RL and using Kiani et al. approach. The graph illustrated the PDR of the designed WSN when three different techniques works with 100, 200, 300, and 400 number of nodes. Proposed approach performs well compared to existing work and the maximum PDR is obtained when (N=400) are deployed. The increase in the rate of packet delivered increases the network reliability as well as the fault tolerance. In case, if the communicating node is found dead, then the proposed algorithm selects previous node and pass its data by considering it as CH node. The simulation shows that PDR increases with the increase in the number of nodes. The percentage increase in the PDR obtained using RL compared to existing Kiani et al approach for 100, 200, 300, and 400 nodes is 11.25%, 13.92%, 15%, and 14.81 % respectively.

Table 15. Energy Consumption (mJ)

Number of nodes	Without RL	With RL	Kiani et al.
100	324	300	325.05
200	336	325	319.05
300	347	333	345.85
400	358	348	331.52

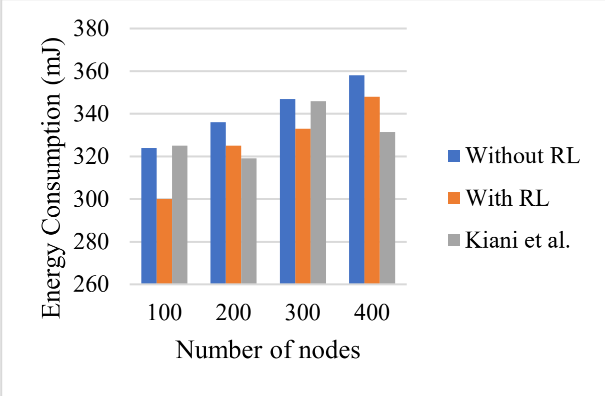

Figure 17. Energy Consumption

The energy consumed by three different approaches with different number of nodes is illustrated in Figure 17 along with the values presented in Table 15. The graph shows the improvement of our RL based cluster selection approach in contrast to the without RL approach and Kiani et al. The energy consumed by the cluster members is determined based on of Mean Square Error (MSE). Efficient network is designed with minimum MSE value examined for particular cluster. From the graph depicted in Figure 17, it is clearly seen that the proposed approach using RL algorithm have less MSE compared to the without RL approach and hence consume less energy. The percentage reduction in the energy consumption rate in contrast to the without RL approach observed for 100, 200, 300, and 400 nodes is 7.41%, 3.27%, 4.03%, and 2.79 % respectively.

In short, the main purpose of the designed protocol is to reduce energy consumption and to increase network lifetime with packet delivery rate. All parameters are obtained as per requirement and hence fulfil the need to design a balanced WSN in dynamic environment.

5. Conclusion

Similar to other existing networks, WSN is developed for specific applications like as military, rescue operation, agriculture and many more; each requires different features as per their need. Depending upon the network scenario, each requires new communication protocols. In addition, network design factors must be taken into account to achieve the expected performance. Among all factors, energy is the most important parameter in WSN and must be controlled by appropriate algorithm. In this research, we focused on to save energy by selecting the appropriate CH using RL approach and the designed algorithm provides better network life with improved energy. Initially, optimal number of clusters in a network is evaluated by analysing the network energy consumption of both inter and intra clustering communication. Also, the process of selecting appropriate CH using RL approach based on reward point is discussed. From the simulation results, proposed algorithm performed better in terms of network lifetime, packet delivery ratio, and energy consumption. The maximum enhancement in the PDR of 15% and its improvement in the energy saving is of 2.79 % has been obtained.

References

[1] Ahmadinia M, Meybodi MR, Esnaashari M, Alinejad-Rokny H. Energy-efficient and multi-stage clustering algorithm in wireless sensor networks using cellular learning automata. IETE Journal of Research. 2013 Nov 1;59(6):774-82.

[2] Arjunan S, Sujatha P. Lifetime maximization of wireless sensor network using fuzzy based unequal clustering and ACO based routing hybrid protocol. Applied Intelligence. 2018 Aug 1;48(8):2229-46.

[3] Shamsan Saleh AM, Ali BM, Rasid MF, Ismail A. A survey on energy awareness mechanisms in routing protocols for wireless sensor networks using optimization methods. Transactions on Emerging Telecommunications Technologies. 2014 Dec;25(12):1184-207.

[4] Ballav B, Rana G, Pattanayak BK. Investigating the effect of Black Hole attack on Zone Based Energy Efficient Routing Protocol for Mobile Sensor Networks. In2015 International Conference on Information Technology (ICIT) 2015 Dec 21 (pp. 113-118). IEEE.

[5] Esnaashari M, Meybodi MR. A novel clustering algorithm for wireless sensor networks using irregular cellular learning automata. In2008 International Symposium on Telecommunications 2008 Aug 27 (pp. 330-336). IEEE.

[6] Elhoseny M, Elleithy K, Elminir H, Yuan X, Riad A. Dynamic clustering of heterogeneous wireless sensor networks using a genetic algorithm, towards balancing energy exhaustion. International Journal of Scientific & Engineering Research. 2015;6(8):1243-52.

[7] Alarifi A, Tolba A. Optimizing the network energy of cloud assisted internet of things by using the adaptive neural learning approach in wireless sensor networks. Computers in Industry. 2019 Apr 1;106:133-41.

[8] Birajdar DM, Solapure SS. LEACH: An energy efficient routing protocol using Omnet++ for Wireless Sensor Network. In2017 International Conference on Inventive Communication and Computational Technologies (ICICCT) 2017 Mar 10 (pp. 465-470). IEEE.

[9] Xu J, Jin N, Lou X, Peng T, Zhou Q, Chen Y. Improvement of LEACH protocol for WSN. In2012 9th International Conference on Fuzzy Systems and Knowledge Discovery 2012 May 29 (pp. 2174-2177). IEEE.

[10] Yulan Z, Chunfeng J. Research about improvement of LEACH protocol. InThe 2nd International Conference on Information Science and Engineering 2010 Dec 4 (pp. 2281-2284). IEEE.

[11] Jin Y, Jo JY, Wang L, Kim Y, Yang X. ECCRA: An energy-efficient coverage and connectivity preserving routing algorithm under border effects in wireless sensor networks. Computer Communications. 2008 Jun 25;31(10):2398-407.

[12] Arumugam GS, Ponnuchamy T. EE-LEACH: development of energy-efficient LEACH Protocol for data gathering in WSN. EURASIP Journal on Wireless Communications and Networking. 2015 Dec;2015(1):1-9.

[13] Alsheikh MA, Lin S, Niyato D, Tan HP. Machine learning in wireless sensor networks: Algorithms, strategies, and applications. IEEE Communications Surveys & Tutorials. 2014 Apr 24;16(4):1996-2018.

[14] Hu Z, Wang Z, Wang J, Zeng T. Routing Protocol for Wireless Sensor Network Based on Automatic Cluster Optimization. International Journal of Online and Biomedical Engineering (iJOE). 2018 Dec 23;14(12):150-63.

[15] Zhang Y, Li W. Modeling and energy consumption evaluation of a stochastic wireless sensor network. EURASIP Journal on Wireless Communications and Networking. 2012 Dec 1;2012(1):282.

[16] Wang X, Han S, Wu Y, Wang X. Coverage and energy consumption control in mobile heterogeneous wireless sensor networks. IEEE Transactions on Automatic Control. 2012 Oct 18;58(4):975-88.

[17] Javaid N, Rasheed MB, Imran M, Guizani M, Khan ZA, Alghamdi TA, Ilahi M. An energy-efficient distributed clustering algorithm for heterogeneous WSNs. EURASIP Journal on Wireless communications and Networking. 2015 Dec;2015(1):1-1.

[18] Logambigai R, Kannan A. Fuzzy logic based unequal clustering for wireless sensor networks. Wireless Networks. 2016 Apr 1;22(3):945-57.

[19] Azharuddin M, Jana PK. PSO-based approach for energy-efficient and energy-balanced routing and clustering in wireless sensor networks. Soft Computing. 2017 Nov 1;21(22):6825-39.

[20] Jamal T, Butt SA. Low-energy adaptive clustering hierarchy (LEACH) enhancement for military security operations. Proc. Of Journal of Basic and Applied Scientific Research, ISSN. 2017:2090-4304.

[21] Palan NG, Barbadekar BV, Patil S. Low energy adaptive clustering hierarchy (LEACH) protocol: A retrospective analysis. In2017 International conference on inventive systems and control (ICISC) 2017 Jan 19 (pp. 1-12). IEEE.

[22] Bian X, Liu X, Cho H. Study on a cluster-chain routing protocol in wireless sensor networks. In2008 Third International Conference on Communications and Networking in China 2008 Aug 25 (pp. 964-968). IEEE.

[23] Xiaoping W, Hong L, Gang L. An improved routing algorithm based on LEACH protocol. In2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science 2010 Aug 10 (pp. 259-262). IEEE.

[24] Peng W, Edwards DJ. K-means like minimum mean distance algorithm for wireless sensor networks. In2010 2nd International Conference on Computer Engineering and Technology 2010 Apr 16 (Vol. 1, pp. V1-120). IEEE.

[25] Tong M, Tang M. LEACH-B: an improved LEACH protocol for wireless sensor network. In2010 6th international conference on wireless communications networking and mobile computing (WiCOM) 2010 Sep 23 (pp. 1-4). IEEE.

[26] Yu J, Qi Y, Wang G, Gu X. A cluster-based routing protocol for wireless sensor networks with nonuniform node distribution. AEU-International Journal of Electronics and Communications. 2012 Jan 1;66(1):54-61.

[27] Nehra V, Sharma AK. PEGASIS-E: Power efficient gathering in sensor information system extended.

Global Journal of Computer Science and Technology. 2013 Dec 7.
[28] Amgoth T, Jana PK. Energy-aware routing algorithm for wireless sensor networks. Computers & Electrical Engineering. 2015 Jan 1;41:357-67.
[29] Al-Humidi N, Chowdhary GV. Energy-aware approach for routing protocol by using centralized control clustering algorithm in wireless sensor networks. InComputing, communication and signal processing 2019 (pp. 261-274). Springer, Singapore.
[30] Le TT, Moh S. An energy-efficient topology control algorithm based on reinforcement learning for wireless sensor networks. Int. J. Control Autom. 2017;10(5):233-44.
[31] Savaglio C, Pace P, Aloi G, Liotta A, Fortino G. Lightweight reinforcement learning for energy efficient communications in wireless sensor networks. IEEE Access. 2019 Mar 4;7:29355-64.
[32] Banerjee PS, Mandal SN, De D, Maiti B. RL-Sleep: Temperature Adaptive Sleep Scheduling using Reinforcement Learning for Sustainable Connectivity in Wireless Sensor Networks. Sustainable Computing: Informatics and Systems. 2020 Jun 1;26:100380.
[33] Boyan JA, Littman ML. Packet routing in dynamically changing networks: A reinforcement learning approach. InAdvances in neural information processing systems 1994 (pp. 671-678).
[34] Wang P, Wang T. Adaptive routing for sensor networks using reinforcement learning. InThe Sixth IEEE International Conference on Computer and Information Technology (CIT'06) 2006 Sep 20 (pp. 219-219). IEEE.
[35] Hu T, Fei Y. QELAR: A machine-learning-based adaptive routing protocol for energy-efficient and lifetime-extended underwater sensor networks. IEEE Transactions on Mobile Computing. 2010 Feb 8;9(6):796-809.
[36] Razzaque MA, Ahmed MH, Hong CS, Lee S. QoS-aware distributed adaptive cooperative routing in wireless sensor networks. Ad Hoc Networks. 2014 Aug 1;19:28-42.
[37] Kiani F, Amiri E, Zamani M, Khodadadi T, Abdul Manaf A. Efficient intelligent energy routing protocol in wireless sensor networks. International Journal of Distributed Sensor Networks. 2015 Mar 2;11(3):618072.
[38] Soni S, Shrivastava M. Novel learning algorithms for efficient mobile sink data collection using reinforcement learning in wireless sensor network. Wireless Communications and Mobile Computing. 2018 Aug 16;2018.
[39] Hu T, Fei Y. QELAR: A machine-learning-based adaptive routing protocol for energy-efficient and lifetime-extended underwater sensor networks. IEEE Transactions on Mobile Computing. 2010 Feb 8;9(6):796-809.
[40] Chen H, Li X, Zhao F. A reinforcement learning-based sleep scheduling algorithm for desired area coverage in solar-powered wireless sensor networks. IEEE Sensors Journal. 2016 Jan 12;16(8):2763-74.