Energy-efficient Q-learning-based routing in wireless sensor networks

Archana Chaudhari[1,2,*],
Vivek Deshpande[3] and
Divya Midhunchakkaravarthy[4]

[1]Lincoln University College, Petaling Jaya, Malaysia

[2]Instrumentation Engineering Department, Vishwakarma Institute of Technology, Pune, India

[3]Vishwakarma Institute of Information Technology, Pune, India

[4]Centre of Postgraduate Studies, Lincoln University College, Petaling Jaya, Malaysia

*E-mail: archana.kchaudhari@gmail.com

Received for publication December 12, 2024.

Abstract

Wireless sensor networks (WSNs) are extensively used to collect and transmit data from the environment through sensor nodes for various applications such as agriculture, medicine, and military. The sensor nodes are located in remote locations and harsh environmental conditions. Sensors operate on a limited power source. The lifetime of the entire WSN depends on the power of its nodes. Also, successful data transmission via routing from the source node to the destination depends on the battery life of the sensor node. It therefore becomes necessary to have an energy-efficient routing path to ensure successful packet delivery from source to sink nodes. The work proposes an energy-aware Q-learning-based routing algorithm in WSN. To establish an energy-efficient route, the work exploits the rewards in Q-learning based on the node energy. Experiments conducted using the proposed work demonstrate enhanced packet delivery, throughput, and networks lifetime. A comparison of the work with a basic Q-learning-based routing algorithm exhibits improved performance.

Keywords

wireless sensor networks, routing, reinforcement learning, Q-learning, node energy, energy-aware, energy aware Q-learning

I. Introduction

Wireless sensor networks (WSNs) are most extensively used to collect and analysis of data from the environment. The applications of WSNs include agriculture, military, and medicine. WSN deploys sensor nodes, such as temperature, humidity, and moisture, to collect data. These data are then transmitted to the destination or the base station [1–2]. A representative node consists of a battery, a sensor, and certain memory and microcontroller devices. The power needed for the node for data transmission is provided using a battery source. The memory device stores the algorithm along with the data sensed from the environment. The microcontroller device controls the data transmission and execution via algorithms. Figure 1 demonstrates a WSN with the architecture of a sensor node.

The sensor nodes are placed in remote locations and harsh environmental conditions, and they are often inaccessible. The sensor nodes depend on the battery life for energy. Also, the network life of WSN depends on the life of the sensors. Due to limitations of the battery, it is important to route the data transmission in such a way to consume energy in an optimal way [3–4]. This will help to extend the lifetime of the WSN.

The previous literature present the survey of energy-efficient or energy-aware-based routing protocols [5–9]. Two types of approaches are proposed: energy saving and energy balancing [10]. Many routing approaches have been proposed to save energy during monitoring of the environment and transmission of data. These are known as energy saving approaches. Another approach is the energy balancing approach. Both approaches prove useful to enhance the networks lifetime.
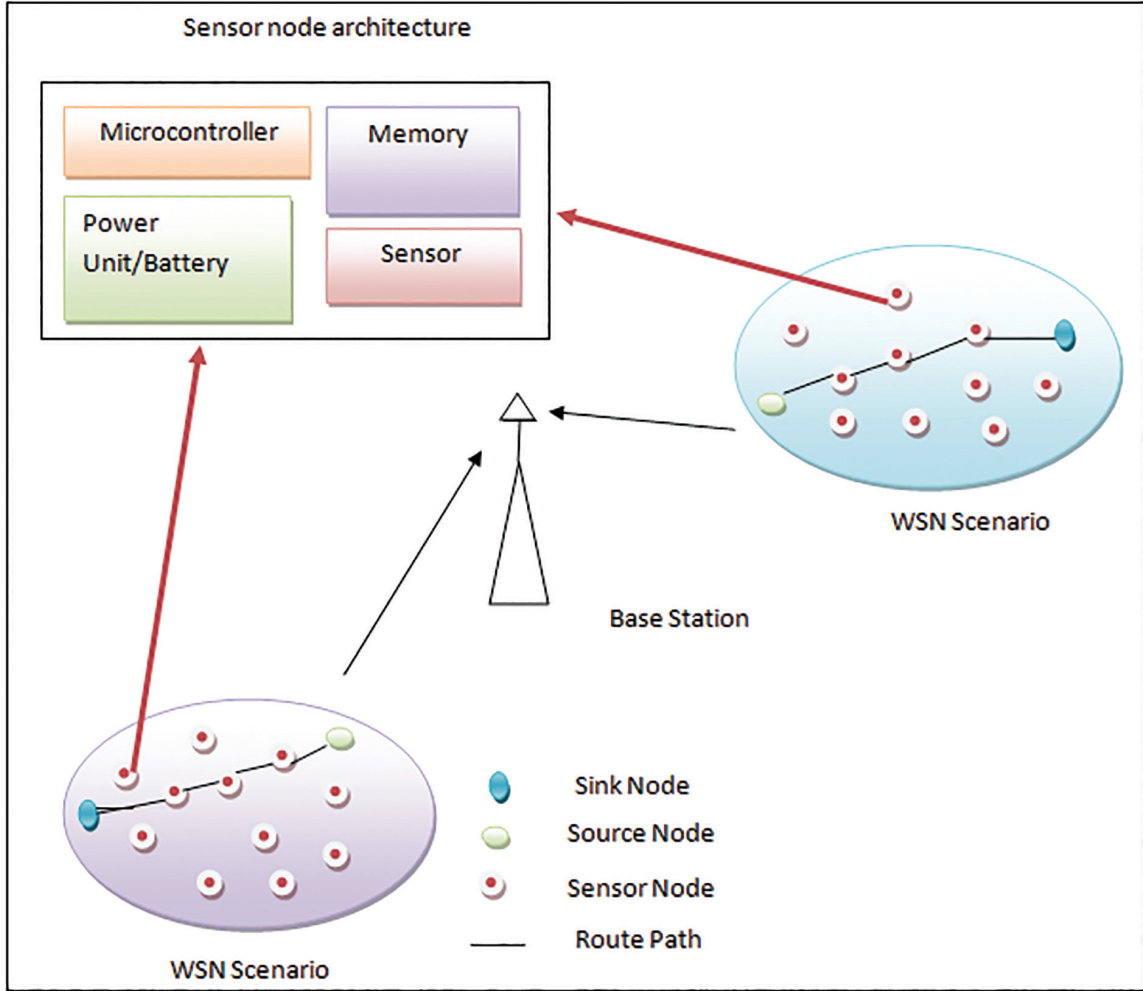
Figure 1: WSN with sensor node architecture. WSN, wireless sensor network.

Machine learning-based routing approaches have been used in the past decade. An overview of machine learning-based energy-efficient routing approach has been presented in Ref. [11].

Reinforcement learning (RL), a branch of machine learning, has been widely used in recent years for routing in WSN. RL offers a basis for WSN nodes to learn from earlier interactions with the environment and choose efficient actions for the future. It proposes two approaches: centralized and decentralized. In a centralized approach, the routing path is chosen based on information of the overall topology while in a decentralized approach, local information of every sensor and its neighbor is used for route selection [12].

An energy-efficient RL routing approach can be helpful to choose the ideal path to enhance the networks lifetime. In this work, Section 1 presents the introduction to the routing in WSNs and the Q-learning approach. Section 2 discusses state-of-the-art methods based on energy-efficient RL-based approaches for routing. Section 3 presents the proposed energy-aware Q-learning-based routing method. Experiments conducted using the proposed method are presented, along with results and discussions, in Section 4. Finally, conclusions based on the work are presented in Section 5.

II. Literature Review

Many routing protocols choose the next forwarder node having good quality of service (QoS). Hence the node energy with good QoS will exhaust. Jafarzadeh and Moghaddam [13] presented a QoS routing protocol in WSNs based on an energy-aware RL approach. Oddi et al. [14] proposed a Q-learning approach with

energy-efficient routing. In this approach, the residual batteries and control overhead of wireless sensors are considered. An intelligent routing protocol based on energy was proposed by Kiani et al. [15] to reduce energy consumption. In the method, initially clustering is applied to establish the network using a connected graph. Data are transmitted using the Q-value to improve networks lifetime. Li et al. [16] proposed an energy-efficient Q-learning-based routing for underwater WSN in which the link quality and energy contribute to Q-value. An energy-efficient routing approach based on RL was proposed by Guo et al. [17]. It explores the reward policy to compute the ideal route so as to enhance lifetime and energy efficiency. For balancing the energy and reducing the energy consumption of the WSN, Su et al. [12] selected the next forwarder based on the Q-values. A Q-learning algorithm depending on the data aggregation along with energy efficiency considerations was proposed by Yun and Yoo [18]. In this approach, they exploited the reward function based on data aggregation at every node and two types of energy: the residual node energy and communication energy to choose the optimum routing path. Cluster-based energy-efficient routing using the RL approach was presented by Yun and Yoo [18]. In the work, the initial energy of the nodes and hop count are used for the calculation of the initial Q-value for the choice of the cluster head (CH). For making the routing decisions, residual energy of the node, along with the hop count, is used to provide energy efficiency during data transmission [19].

For cognitive radio sensor networks, Joon and Tomar [20] proposed Ad hoc On Demand Distance Vector (AODV) routing based on energy-aware Q-learning. In the work, rewards are exploited depending on the Q-learning for choosing a cluster head. The route is established using AODV [20]. Su et al. [21] presented a Q-learning-based approach using the energy of information transmission and Q-values for extending the lifetime of the network. Maivizhi and Yogesh [22] presented Q-learning-based routing for in-network aggregation in WSNs. Q-values in the work are computed based on the node energy, link strength, and distance among nodes. Yun and Yoo [23] proposed Q-learning-based routing by exploiting the rewards in terms of sensor data aggregation, node residual energy, and communication energy.

III. Proposed Methodology

Q-learning is a model-free type of RL in which the RL agent learns from the environment from the past and current experiences. Q-learning consists of state-action pairs. To learn from the environment, a state is defined, and for every state, an action is defined. For each action, the agent is rewarded. The reward can be positive, negative, or neutral. Based on the reward, the Q-values are computed. The computed Q-values further choose the action to be executed to reach the goal.

For routing in WSN, the objective of the proposed Q-learning-based routing approach is to decide the next forwarder node for the transmission of data from the source node to the destination. In the proposed method, the source is taken as the current or present node. From the present node, the goal of the agent is to route data to the sink node via the best next neighbor.

The best next forwarder node is defined as that node chosen depending on the Q-value of each node. For the node on the path to be the best next forwarder node, the following conditions need to be met:

1. The node needs to be the neighbor node of the current node. A node can be a neighbor of the current node if it lies within the transmission range of the current node.
2. The Q-value of the node needs to be maximum among all the neighbor nodes of the present or current node.

From the above hypotheses, the Q-learning agent is defined in Eq. (1) as

$$Q(x,y) = (1-\alpha)*Q_{init}(x,y) \\ + \alpha*(Reward(x,y)+\Upsilon*\max(Q(x,y))) \tag{1}$$

where x is the present or current node, y is the neighbor of current node, α is the learning rate (taken as 0.5 for computations), γ is the discounted factor (chosen as 0.8 for computations), (x,y) is the Q matrix initialized randomly at the beginning, and $(Q(x,y))$ is the maximum of the Q-value chosen among all the neighbors of the present node.

For the proposed method, the reward exploits the energy of the neighbor nodes. It is defined in Eq. (2) as:

$$Reward(x,y) = \begin{cases} 100; if \ (node\,energy(y) \geq 0.1 Joules \\ 1; if \ (node\,energy(y) < 0.1 Joules \end{cases} \tag{2}$$

The neighbor node of the present node is defined based on the transmission range. The transmission

range for the WSN scenario is defined as 40 m. Eq. (3) is used to compute the neighbors of the current node.

$$node(y)$$
$$= neighbor\ of\ node\ x; if\ dist(x,y) \leq transmisison\ range$$
(3)

where $dist(x,y)$ is computed in Eq. (4) as:

$$dist(x,y) = \sqrt{(x_{x_{position}} - y_{y_{position}})^2 + (x_{y_{position}} - y_{y_{position}})^2} \quad (4)$$

The neighbor nodes update its node energy after every packet transmission using Eq. (5) as

$$energy\_node$$
$$= alpha1 * transmission\_rate * packet\_size * 8)$$
$$+ (alpha2 * transmission\_rate * packet\_size * 8)$$
$$* (dist(x,y) \wedge alpha)$$
(5)

where $alpha1 = 50e\text{-}9;\%J/bit$, $alpha2 = 0.1e\text{-}9;\%J/bit/m2$, $alpha = 2$, x = current node = next forwarder neighbor node, and $dist()$ is the function to estimate the Euclidean distance amid the nodes.

In the proposed method, from Eq. (2), if the energy of the neighbor node is <0.1 Joules, the reward is more, and if the energy of the node is <0.1 Joules, the reward is less. This is because the node having energy <0.1 Joules may not be able to transmit the packet to the next neighbor. The Q-value of the neighbor node is computed based on the reward from Eq. (1). The best next forwarder is chosen from the neighbor table of the current node as the node with the maximum Q-value.

The proposed method can communicate the packets from the source to the sink node. But the nodes adjacent to the sink nodes will experience added energy depletion because of congestion. Hence, after some time in the WSN scenario, no best next forwarder might exist as the energy of the nodes will deplete after packet transmissions. At this instance, the data packet will not be transmitted further, and the packet will be dropped.

Steps for the proposed energy-aware Q-learning-based routing algorithm are summarized as below:

1. Initialize the network nodes' positions with the number of nodes.
2. Choose the source node as the current node.
3. Obtain the neighbor nodes of the current node using the transmission range.

If the sink node exists in the neighbor table of the current node, check the node energy of the current node, and if the energy of the current node is >0.1 Joules, forward the packet to the sink node; else, drop the packet.

4. Compute the reward matrix *Reward*(*x,y*), where *x* is the current node and *y* is the node from the neighbor table from Eq. (2). Initialize the Q matrix $Q_{init}(x,y)$ randomly.
5. Compute the *Q*(*x,y*) matrix of each neighbor node from Eq. (1) and store the Q-value in the Q table of each node.
6. Choose the next forwarder node as the node with the max Q-value among the neighbor nodes of the current node.
7. Forward the packet to the next forwarder node and update the node energy of the current node.
8. If the packet has not reached the sink node, make the next forwarder node the current node and continue with Step 3 till the data packet reaches the sink.

IV. Results and Discussion

The experimental analysis of the proposed method is implemented using Matlab 2022 software. Two source nodes are defined, and Constant Bit Rate (CBR) traffic is attached to the source nodes. The duration of the CBR traffic is defined as 5 s. For experimental analysis, the number of nodes (N) is varied from 20 to 100 by keeping the packet size as 20 bytes and packet transmission rate as 20 packets/s. The proposed energy-aware Q-learning-based routing method is compared with the basic Q-learning-based routing algorithm [24].

The performance evaluation of the experimental analysis and comparison are conducted using the packet delivery ratio (PDR), packet loss ratio (PLR), throughput, and the networks lifetime. PDR is defined as the total number of packets sent from the source to the sink node. The PLR is the number of packets lost during transmission from the source to the destination. Throughput is defined as the total amount of data packets transmitted per given time. The networks lifetime is the total amount of energy consumed by the network and how many nodes are alive in the network.

Table 1 presents the comparison of the performance of the proposed energy-aware Q-learning-based routing algorithm with Q-learning-based routing.

Figure 2 presents the performance of the proposed energy-aware Q-learning-based routing method for the number of nodes against PLR. From

Table 1: Performance comparison of the proposed energy-aware Q-learning-based routing algorithm with Q-learning-based routing.

Number of nodes	PLR		PDR		Throughput		Networks Lifetime	
	Q-learning	Proposed energy-aware Q-learning	Q-learning	Proposed energy-aware Q-learning	Q-learning	Proposed energy-aware Q-learning	Q-learning	Proposed energy-aware Q-learning
20	0.3	0.58	0.7	0.42	0.175	21	0.2674	0.358
30	0.28	0.76	0.71	0.44	20.75	22	0.3155	0.344
40	0.28	0.51	0.725	0.49	23.25	25	0.3439	0.33
50	0.275	0.42	0.725	0.58	36.25	37	0.46233	0.33
60	0.17	0.01	0.725	0.99	41.35	49.5	0.5965	0.978
70	0.13	0.01	0.87	0.99	43.5	49.5	0.6865	1.29
80	0.13	0.01	0.87	0.99	43.5	49.5	0.8072	1.43
90	0.13	0.01	0.87	0.99	43.5	49.5	0.9189	1.47
100	0.13	0.01	0.87	0.99	43.5	49.5	1.032	1.596

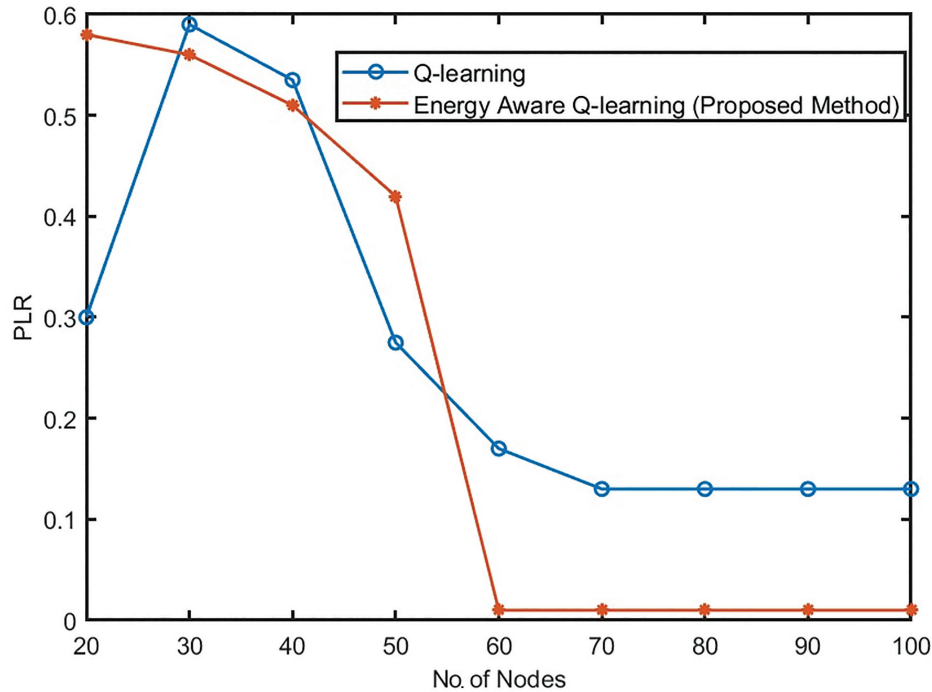PDR, packet delivery ratio; PLR, packet loss ratio.

Figure 2: PLR against number of nodes. PLR, packet loss ratio.

Figure 2, it is observed that due to the energy constraint in the reward function, the PLR has decreased compared with the basic Q-learning-based routing method.

Figure 3 represents the performance of the proposed method for the number of nodes against PDR. From Figure 3, PDR has increased compared with the basic Q-learning-based routing method.
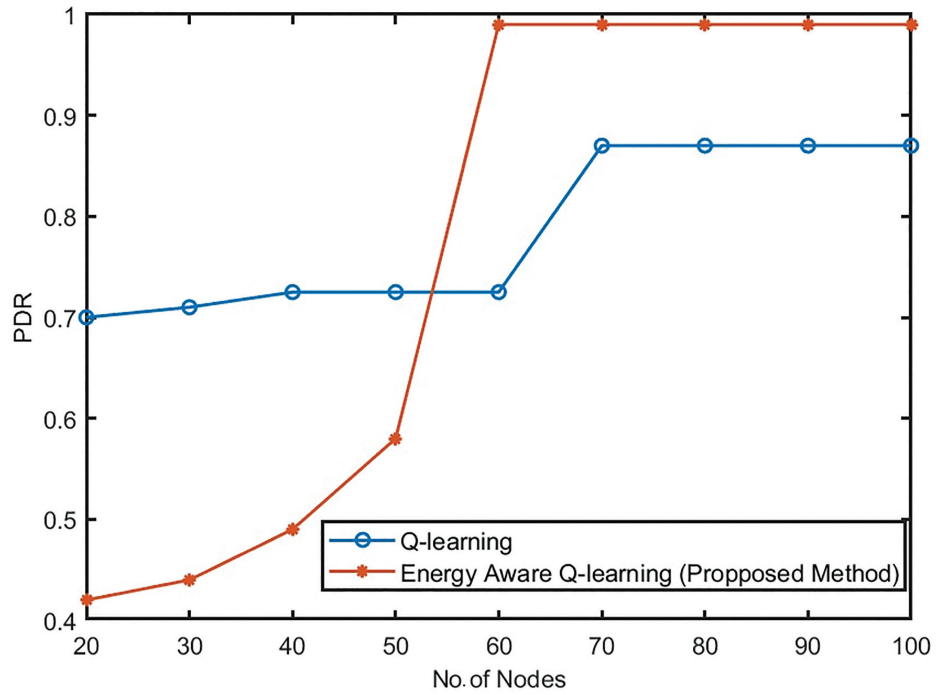
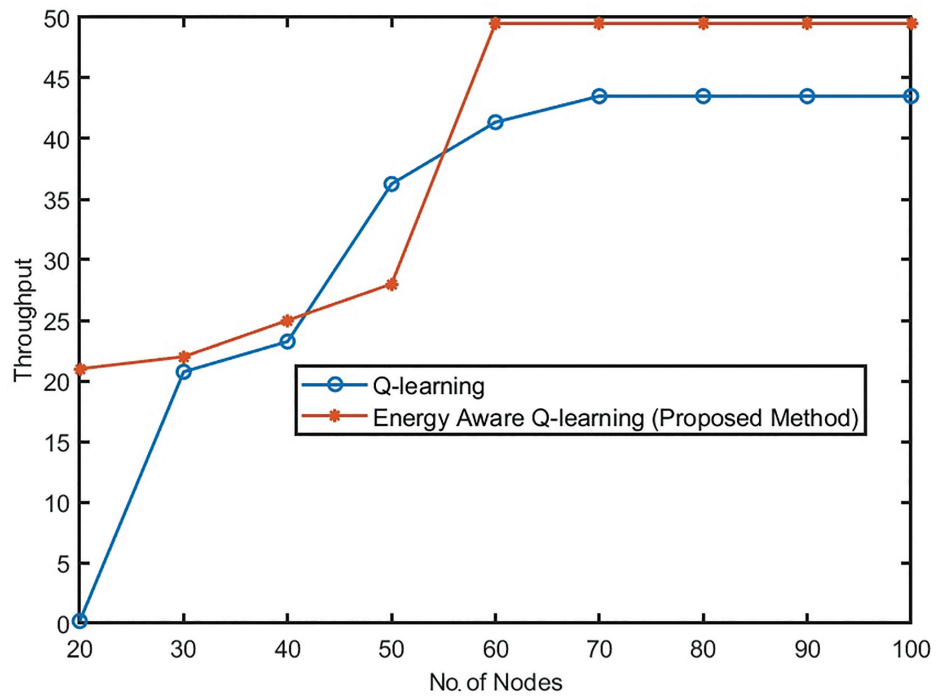Figure 3: PDR against number of nodes. PDR, packet delivery ratio.

Figure 4: Throughput against number of nodes.

Figure 4 exhibits the performance of the throughput of the proposed method. The throughput of the proposed method is seen to enhance compared with the Q-learning method.

Figure 5 demonstrates the networks lifetime of the proposed method and illustrates the enhanced networks lifetime using the proposed method in comparison to the Q-learning algorithm.
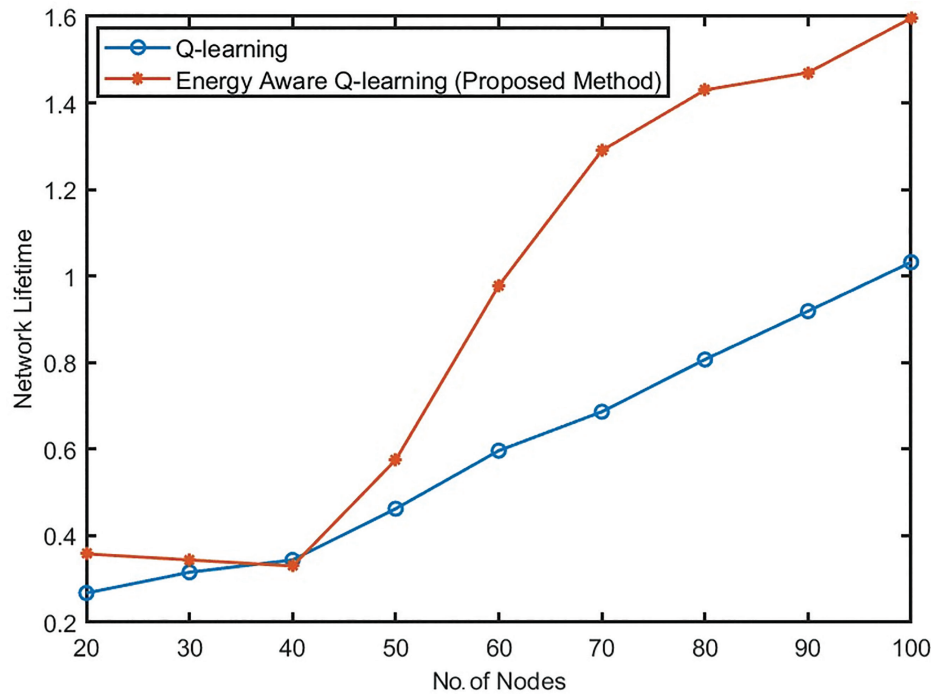
Figure 5: Networks lifetime against number of nodes.

V. Conclusions

Routing of data packets in WSNs ensures that packets transmitted from the source have reached the sink node without packet loss. In WSN, the nodes near the sink nodes experience congestion as most of the packets are routed via these nodes to the sink. This results in depletion of the nodes energy and packet loss. The work proposed an energy-aware Q-learning-based routing method. The work exploits the rewards in Q-learning in terms of the node energy for a better choice of the next forwarder node to the sink node along with experimental analysis. As stated in the literature review, many other works exploit the hop count of the nodes to the sink to choose the next neighbor. A few works have explored the hop count of the node along with the residual energy of the node for better routing. Experimental analysis of the proposed work demonstrates enhanced PDR, throughput, and networks lifetime and reduced PLR compared with basic Q-learning-based routing.

References

[1] Meena Pundir and Jasminder Sandhu, A systematic Review of Quality of Services in Wireless Sensor Networks using Machine Learning: Recent Trend and Future Vision, Journal of Network and Computer Applications, 188, (2021). DOI: 10.1016/j.jnca.2021.103084.

[2] Padmalaya Naya, G. K. Swetha, Surbhi Gupta, K. Madhavi, Routing in wireless sensor networks using machine learning techniques: Challenges and opportunities, Measurements, 178, (2021). DOI: 10.1016/j.measurement.2021.108974.

[3] Praveen Kumar D, Tarachand Amgoth, Chandra Sekhara Rao Annavarapu, Machine learning algorithms for wireless sensor networks: A survey, Information Fusion (2018).DOI: 10.1016/j.inffus.2018.09.013

[4] Raja Basha, A. A Review on Wireless Sensor Networks: Routing. Wireless Personal Communication 125, 897–937 (2022). DOI: 10.1007/s11277-022-09583-4

[5] Pantazis, N.A.; Nikolidakis, S.A.; Vergados, D.D. Energy-Efficient Routing Protocols in Wireless Sensor Networks: A Survey. IEEE Commun. Surv. Tutor. 2013, 15, 551–591. [CrossRef]

[6] Jingjing, Y.; Mengchu, Z.; Zhijun, D. Recent advances in energy-efficient routing protocols for wireless sensor networks: A review. IEEE Access 2016, 4, 5673–5686.

[7] Mohamed, R.E.; Saleh, A.I.; Abdelrazzak, M.; Samra, A.S. Survey on wireless sensor

network applications and energy efficient routing protocols. Wirel. Pers. Commun. 2018, 101, 1019–1055. [CrossRef]

[8] C. Nakas, D. Kandris, and G. Visvardis, ``Energy efficient routing in wireless sensor networks: A comprehensive survey,'' Algorithms, vol. 13, no. 3, p. 72, Mar. 2020, doi: 10.3390/a13030072

[9] M. Shaq, H. Ashraf, A. Ullah, and S. Tahira, ``Systematic literature review on energy efficient routing schemes in WSN: A survey,'' Mobile Netw. Appl., vol. 25, no. 3, pp. 882895, Feb. 2020, doi: 10.1007/s11036-020-01523-5.

[10] Ehsan Ahvar, Shohreh Ahvar, Gyu Myoung Lee, Noel Crespi, "An Energy-Aware Routing Protocol for Query-Based Applications in Wireless Sensor Networks", The Scientific World Journal, vol. 2014, Article ID 359897, 9 pages, 2014. DOI: 10.1155/2014/359897

[11] Ding, Q.; Zhu, R.; Liu, H.; Ma, M. An Overview of Machine Learning-Based Energy-Efficient Routing Algorithms in Wireless Sensor Networks. Electronics 2021, 10, 1539. DOI: 10.3390/electronics10131539.

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

[13] S. Z. Jafarzadeh and M. H. Y. Moghaddam, "Design of energy-aware QoS routing protocol in wireless sensor networks using reinforcement learning," *2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1-5, (2014), 10.1109/CCECE.2014.6900988.

[14] G. Oddi, A. Pietrabissa, and F. Liberati, "Energy balancing in multihop wireless sensor networks: An approach based on reinforcement learning," in Proc. NASA/ESA Conf. Adapt. Hardw. Syst. (AHS), 2014, pp. 262–269.

[15] Kiani F, Amiri E, Zamani M, et al. Efficient intelligent energy routing protocol in wireless sensor networks. Int J Distrib Sens N vol.5, issue 27, (2015), DOI: 10.1155/2015/618072.

[16] X. Li, X. Hu, W. Li, and H. Hu, "A multi-agent reinforcement learning routing protocol for underwater optical sensor networks," in Proc. IEEE Int. Conf. Commun. (ICC), 2019, pp. 1–7.

[17] W. Guo, C. Yan, and T. Lu, "Optimizing the lifetime of wireless sensor networks via reinforcement-learning-based routing," Int. J. Distrib. Sens. Netw., vol. 15, no. 2, 2019, Art. no. 155014771983354.

[18] W. -K. Yun and S. -J. Yoo, "Q-Learning-Based Data-Aggregation-Aware Energy-Efficient Routing Protocol for Wireless Sensor Networks," in IEEE Access, vol. 9, pp. 10737-10750, 2021, doi: 10.1109/ACCESS.2021.3051360.

[19] ViallyKazadi Mutombo, Seungyeon Lee, Jusuk Lee, and Jiman Hong, EER-RL: Energy-Efficient Routing Based on Reinforcement Learning, Mobile Information SystemsVolume 2021, DOI: 10.1155/2021/5589145.

[20] Ranjita Joon, Parul Tomar, Energy Aware Q-learning AODV (EAQ-AODV) routing for cognitive radio sensor networks, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 9, 2022, pp. 6989-7000, DOI: 10.1016/j.jksuci.2022.03.021.

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

[22] Maivizhi, R., Yogesh, P. Q-learning based routing for in-network aggregation in wireless sensor networks. Wireless Networks 27, 2231–2250, 2021. DOI: 10.1007/s11276-021-02564-8.

[23] W. -K. Yun and S. -J. Yoo, Q-Learning-Based Data-Aggregation-Aware Energy-Efficient Routing Protocol for Wireless Sensor Networks, IEEE Access, vol. 9, pp. 10737-10750, 2021. doi: 10.1109/ACCESS.2021.3051360.

[24] J. A. Boyan and M. L. Littman, Packet routing in dynamically changing networks: A reinforcement learning approach, in Proc. Neural Inf. Process. Syst. Conf. (NIPS), pp. 671–678, (1994), 10.5555/2987189.2987274

Authors' Biography

Archana Chaudhari: Post Doctoral Fellow at Lincoln University College, Petaling Jaya, Malaysia. Doctor of Philosophy (Engineering), Assistant Professor, of the Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India. Areas of scientific interests: Wireless Sensor Networks, Internet of Things, Machine Learning, Deep Learning, and Control System.
Email: archana.chaudhari@vit.edu

Vivek Deshpande: Post Doctor Fellow (Engineering), Director and Professor, Vishwakarma Institute of Information Technology, Pune, India. Areas of scientific interests: Wireless Sensor Networks, Distributed Systems, High Performance Computer Networks Real Time Embedded Systems, and Internet of Things.
Email: vivek.deshpande@viit.ac.in

Divya Midhunchakkaravarthy: Centre of Postgraduate Studies, Lincoln University College, Malaysia. Areas of scientific interests: Cybersecurity, Network and Systems, and Data Science.
Email: divya@lincoln.edu.my