

Reinforcement Learning to Improve QoS and Minimizing Delay in IoT

Mahendrakumar Subramaniam^{1,*}, V. Vedanarayanan², Azath Mubarakali³ and S. Sathiya Priya⁴

¹Department of Electronics and Communications Engineering, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India

²Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

³College of Computer Science, King Khalid University, Abha, Saudi Arabia

⁴Department of Electronics and Communications Engineering, Panimalar Institute of Technology, Chennai, Tamil Nadu, India

*Corresponding Author: Mahendrakumar Subramaniam. Email: mahen.sub@yahoo.com

Received: 16 May 2022; Accepted: 06 July 2022

Abstract: Machine Learning concepts have raised executions in all knowledge domains, including the Internet of Thing (IoT) and several business domains. Quality of Service (QoS) has become an important problem in IoT surrounding since there is a vast explosion of connecting sensors, information and usage. Sensor data gathering is an efficient solution to collect information from spatially disseminated IoT nodes. Reinforcement Learning Mechanism to improve the QoS (RLMQ) and use a Mobile Sink (MS) to minimize the delay in the wireless IoT s proposed in this paper. Here, we use machine learning concepts like Reinforcement Learning (RL) to improve the QoS and energy efficiency in the Wireless Sensor Network (WSN). The MS collects the data from the Cluster Head (CH), and the RL incentive values select CH. The incentives value is computed by the QoS parameters such as minimum energy utilization, minimum bandwidth utilization, minimum hop count, and minimum time delay. The MS is used to collect the data from CH, thus minimizing the network delay. The sleep and awake scheduling is used for minimizing the CH dead in the WSN. This work is simulated, and the results show that the RLMQ scheme performs better than the baseline protocol. Results prove that RLMQ increased the residual energy, throughput and minimized the network delay in the WSN.

Keywords: Wireless sensor network; quality of service parameters; Internet of things; reinforcement learning; mobile sink

1 Introduction

The main challenge of the Internet of Thing (IoT) network is the battery lifespan of the limited IoT devices. IoT devices utilize large amounts of energy to transmit information to their neighbours. These transmissions will rapidly lead to the device's battery reduction. IoT and Wireless Sensor Network (WSN) are joined in an inventive method for a smart observing system in this approach. Sensor nodes are organized in fields that collect information about dissimilar attributes [1].



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

WSN is playing a significant part in the arrival of IoT and the simplification of the usage of the web. The IOT cloud indicates the platforms which permit making web services suitable for the objects combined on the internet [2]. This work utilizes the chosen methodology for smart agriculture with IoT. This approach recognizes the significant devices, network approaches, platforms, processing information technologies, and the applicability of smart agriculture with IoT. Precision agriculture uses IoT and machine learning methods that are directed to distinguish from the growers about the trending machinery and their effect on the accuracy of cultivation. Smart environment sensors combined with IoT can offer a new idea in tracking, recognizing, and observing objects in a situation. This can offer possible assistance leading up to the opportunity of attaining a green world and a maintainable routine [3].

The clustering method is initially executed to minimize energy utilization. A node which is associated together is known as a cluster [4]. Cluster Head (CH) is an essential cooperative function for a network. In a clustered WSN, a node is selected as a CH that collects its cluster data and then it transmits to the Base Station (BS) through other CHs [5].

IoT permits ecological sensors to attach to other systems via Bluetooth or Wi-Fi to forward huge quantities of information to the network and can permit us to take a better accepting of our environments and discover appropriate solutions for nowadays ecological issues [6]. Reinforcement Learning (RL) makes up an uncomplicated non-linear facility that converts from lower layer to higher layer to reach a superior resolution. It is encouraged via transmission plan and information treatment in nerve preparations. RL method is utilized to improve the rate of accuracy rate. RL is adopting a learning method that discharges agents take process to intensify the opinion of rewards. An RL method has an agent that examines the state notices and makes a decision [7]. The benefit of this learning is eliminating great-level features from the information; also, it can be qualified to achieve many aims. It is applied in many fields such as social networks, business intelligence, handwriting recognition, Bioinformatics, speech detection, and image processing [8]. This kind of learning has many compact issues like routing, data quality evaluation, energy-saving, and attacker detection [9].

Problem Statement:

An Optimal Path from Source to Sink Node (OPSSN) is introduced to achieve packet delivery and improve network performance. This method provides automatic variation to dynamic linking and low processing overhead through selecting the default path. This method increases the number of extra neighbors that are conserved data communication. However, this approach increases the energy utilization in the network [10]. To solve this problem, the RL to improve Quality of Service (QoS) and minimizing delay in IoT is introduced. This approach enhances power-efficient and QoS augmentation using machine learning concepts. The contribution of this approach has the following statement.

- Initially, we form the clusters by the node distance, and the CH node is selected by the RL technique. RL technique assists in discovering an optimal rule and enhances the expected outcomes based on the value of the incentives.
- The incentives value is computed by the QoS parameters such as minimum energy utilization, minimum bandwidth utilization, minimum hop count, and minimum time delay. The highest QoS node is selected as the best incentive node. Hence, the RL method improves the QoS and energy efficiency.
- Then, we use Mobile Sink (MS) to collect the CH data from a number of clusters of IoT devices. The MS constant speed movement from one CH to the next CH. We use sleep and awake scheduling is used to improve energy efficiency.
- The MS is used to collect the data from CH, thus minimizing the network delay. The sleep and awake scheduling also minimize the CH dead in the WSN.
- Finally, the MS collects the surrounding information from IoT sensor nodes and forwards the data to the BS.

2 Related Works

Energy-efficient Q-learning and approach is utilized for controlling congestion and traffic-aware routing. This approach measures the node-link quality through Q-Learning Routing. Hence, this approach minimized the routing overhead and reduced the time expressively. KullbackLeibler sparse auto encoder congestion-control method, with the assistance of the reconstruction loss operation via departure, reduces the congestion, thus minimizing the losses of the packet and increasing the received ratio in the network [11]. QoS and energy-aware cooperative routing is applied with RL via opponent modelling, optimizing a cooperative transmission method by received signal strength and node energy [12].

The Dynamic Rerouting with Cooperative Communication (DRCC) approaches to retransmit the losses and updated packets. The cooperative opportunistic approaches reduce the prolonged routing and evade the network holes [13]. In this approach, the greedy algorithm communicates the data packets from the sender to the receiver. Furthermore, the opportunistic routing evades the void regions by creating backward communication to discover reliable routes [14]. The relay cooperation approach is utilized for reliable information distribution; while the signal to noise ratio of the obtained signal is not inside the predefined threshold next, the maximal ratio combination is applied as various techniques to enhance the SNR of the obtained signals [15].

An energy competent, cooperative routing approach recognized as Region-Based Courier-nodes Mobility through Incremental Cooperative routing. This approach applies the transmission nature of sensor nodes and executes an incremental cooperative routing. A severe assessment and verification method with present state-of-the-art yield enhanced the energy efficiency [16]. An enhanced ant colony optimization method for solving the packet drop issue. This approach contains an update of local and global pheromone to improve path examination and manipulation [17]. Traffic aggregation and multi-hopping WSN approach can create a great load. In this approach, the aggregation applying per-node energy depletion and routing method dynamically selecting a traffic-aware distance-vector from several topologically, negligible hop-count also the quality of the link to equate the load, as a result, minimize the maintenance of the network [18].

The entire traditional approach described the sensor node's localization. Though, they necessitate widespread calculation power that rises with the enhancement in computational complexity. Link quality evaluation is used to express the link quality. However, this approach is unstable [15]. Normally, the localization techniques are separated into two parts: distance applying Angle of Arrival, Residual Signal Strength Indication (RSSI), Time of Arrival, etc. RSSI-based link quality measurement is used to enhance the quality. However, it creates additional computational complexity [19]. In a lightweight genetic algorithm, every sensor is aware of the data traffic rate to observe the congestion. In this approach, the node fitness function is intended from the middling and the standard deviation of the traffic rates. The dominant gene sets method chooses the appropriate transmitting sensor nodes data to evade the highest traffic congestion [20]. A distributed traffic-aware routing approach with the capability to adjust the information communication rate is introduced for multi-sink, efficiently allocating traffic from the sender to sink nodes [21]. Machine learning is intelligent learning for providing an efficient result. [22] Support Vector Machine is used to improve the efficiency of the network [23]. Convolution network with feature optimization method minimizes the network parameters. The sample-based self-learning can enhance the recognition accuracy [24].

3 RL Mechanisms to Improve the QoS and Using MS for Minimizing the Delay

A WSN contains numerous IoT nodes distributed arbitrarily and these nodes form several clusters of IoT devices, is shown in Fig. 1. In this figure, we form the clusters by the node distance, and the CH node is selected by the RL method. Here, we use MS to collect CH data from a number of clusters of IoT

devices. The MS is a movement agent and it moving constant speed from one CH to the next CH. The MS main function collects the environment information from CH's and forwards this information to the BS. Sleep and awake scheduling is a power on and power off schedule. Sleep represents the sensor node monitoring is off and awake indicates the sensor node monitoring state is on. Hence, minimizes unwanted energy consumption. We use sleep and awake scheduling is used to improve energy efficiency. Each node's energy is 1 joule, and the energy value is below 0.3 joules; this node automatically goes to sleep condition. The BS near CH is easily a burden. This approach using MS for collecting data from the CH, as a result, minimized the CH burden in the WSN.

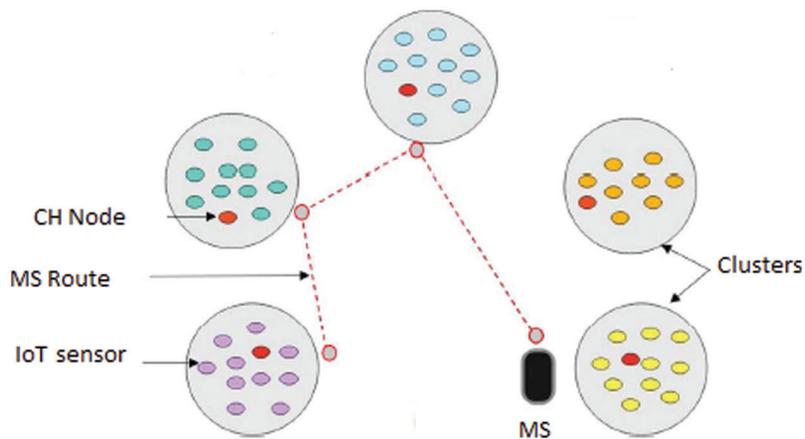


Figure 1: Example diagram of RLMQ

In this approach, we determine an optimal data gathering rule which enhances the data gathering and reduces the energy utilization of the sensor node. This approach attributes as follows $\{ST, AC, In, TP\}$. The Markov Decision Process issue (MDP) defines the function of the State (ST), Action (AC), Incentive (In) and Transmission Possibility (TP). Here, the RL technique solves the MDP issue. At time t , the state represents $ST = L, G, In, CH$.

Here, L indicates the node location, G represents the MS, RE denotes the remaining energy and CH depicts the CH. The RL selects the CH.

I_{ED} indicates an incentive agreed upon if the data gathering is more significant than a specified threshold and the energy utilization is a smaller amount than a specified threshold;

- I_{SD} represents an incentive agreed upon if the data gathering is greater than a specified threshold, but the energy utilization is greater than a specified threshold;
- I_{ee} represents incentive agreed if the data gathering is smaller than a specified threshold, but the energy utilization is smaller than a specified threshold;
- I_{ne} represents a penalty agreed upon if the data gathering is smaller than a specified threshold, but the energy utilization is extra than a specified threshold;
- I_{move} is a penalty that is agreed upon every time the MS moves lack data collecting.

In this approach, the MS is used for data gathering purposes. The MS collects the observing information from each CH and forwards this information to the BS. The movable devices are called as MS. The MS examples are, drones, movable vehicles etc. The MS-based data collection improves the QoS. The MS transfers the data to BS at a specific time. In addition, it improved energy efficiency and minimizes the routing load in the WSN.

Fig. 2 explains the structure of the Reinforcement Learning Mechanism to improve the QoS (RLMQ) approach. RL technique assists in discovering an optimal rule and enhances the expected value of the whole incentive from the present state and each value of the total incentive from the present state and the following states. An agent monitoring a state s of its situation at time stage $S_n \in S$ and S denotes the states set shown in Fig. 3. The agent acts as an $\in A$, where A represents the feasible actions set. Assume the present state denotes the S_n , the next state denotes the S_{n+1} , and the possible action indicates the interaction among neighbor nodes. This action updates the environment state to a next state S_{n+1} , and the agent will obtain an incentive (In) consequently. The situation is probable to be non-deterministic as taking similar action in a similar state on a dissimilar instance may effect different states and various incentives. The agent’s objective is to discover a rule π which maps a state S_n to a possibility of selecting action a_n signify as following in Eq. (1).

$$\pi : S \rightarrow P(A) \tag{1}$$

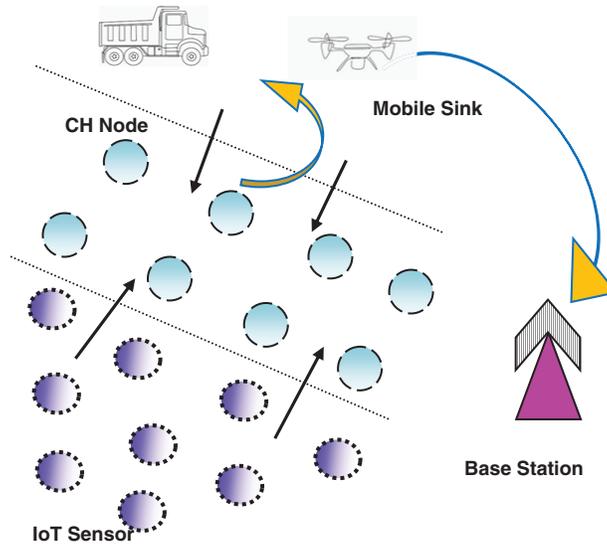


Figure 2: Structure of RLMQ approach

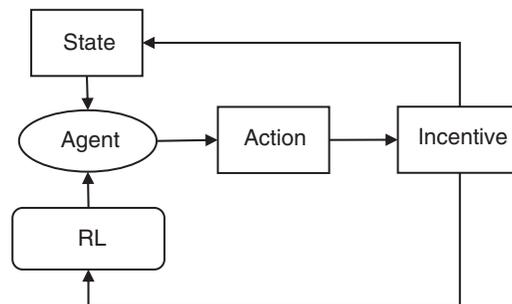


Figure 3: State Diagram of RLMQ

The optimal rule π^* is distinct as

$$\pi^* = \arg \max E[In|\pi] \tag{2}$$

This optimal rule has an optimal state-action value operation which suits Bellman's optimality. It is depicted the following

$$Q^*(s, a) = E[I_n + 1 + B \max Q_*(S_{n+1}, a_{n+1})] \quad (3)$$

RL is an effective technique while the situation is restricted to small state spaces. However, computing and adapting values of QoS table for every state-action pair by QoS factor. Here, we determine the QoS value by minimum energy utilization (EU), minimum Bandwidth Utilization (BU), minimum Hop Count (HC) and minimum time Delay (D) for data transmission. The Incentive value is computed based on the QoS table value. The highest QoS node is selected as the best incentive node. Here, the value Incentive (In) computation is given below.

$$I_n = \beta^{HC} \times (EU + BU + D) \quad (4)$$

The incentive is the total of discounted upcoming incentives. To compute the incentive we utilize β here $\beta \in [0, 1]$ is a discount feature that decides the effect of the prospect rewards on the present one. Here, the CH is elected by the highest incentive nodes. Then, the MS collects the surrounding information from IoT sensor nodes and forwards the data to the BS.

4 Simulation Results

In this work, the network simulator –2.35 tool for measuring the network performance of OPSSN and RLMQ approaches. Here, we use 100 sensor nodes, and this sensor node's transmission range is 150 m. The RLMQ and OPSSN approaches execution simulation time is 200 s. The function of the RLMQ is measured by throughput, delay, residual energy, packet loss ratio and routing load.

• Throughput

The throughput determines the success rate of the data transmissions in a WSN. The throughput values of the existing OPSSN and proposed RLMQ approaches are shown in Fig. 4. From this figure, the RLMQ approach throughput ratio is 16.47% greater than the RLMQ approach since it forms the route by RL. It enhances routing efficiency.

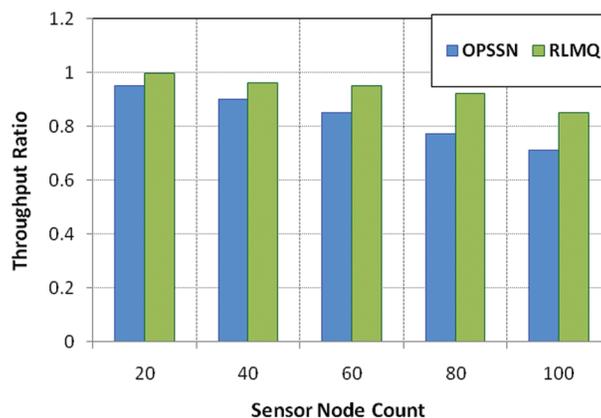


Figure 4: Throughput Ratio of OPSSN and RLMQ

But, the OPSSN minimizes the throughput when it increases the sensor node count since it can't use the machine learning concept in the WSN.

- **Delay**

Fig. 5 displays the Delay of OPSSN and RLMQ Approaches based on sensor node count. From this figure, the OPSSN approach illustrates the 7.85% highest delay time equated to the RLMQ approach. The proposed RLMQ approach is screening the highest throughput and the least delay since it uses the MS for data gathering. The OPSSN approach can't use MS for data gathering. As a result, CH dead easily and increases the delay in the WSN.

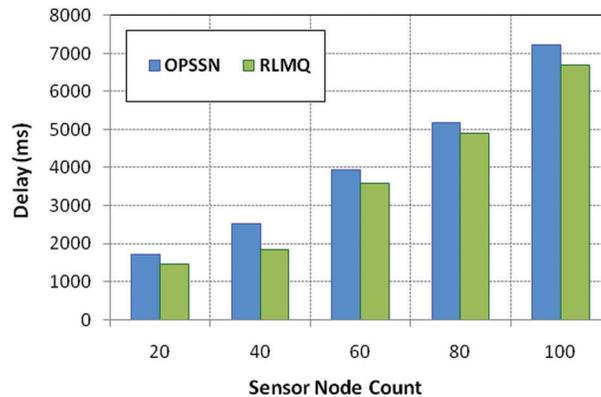


Figure 5: Delay of OPSSN and RLMQ

- **Packet Loss Ratio (PLR)**

It decides the ratio of packets lost in a WSN during data communication. The PLR also helps in deciding the network performance of OPSSN and RLMQ approaches. Fig. 6 illustrates the comparisons of the PLR of OPSSN and RLMQ approaches based on sensor node count.

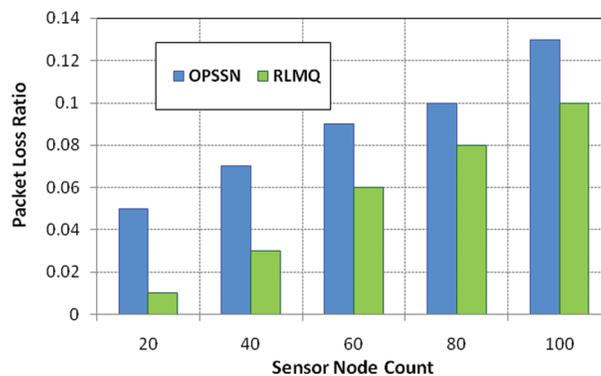


Figure 6: Packet Loss Ratio of OPSSN and RLMQ

Fig. 6 increases the sensor node count; the PLR of OPSSN is 30% highly raised compared to the RLMQ approach. But, the RLMQ approach raises the PLR value while raising the sensor node count in the WSN since it uses the RL technique to discover the route efficiently.

- **Residual energy**

Residual Energy is directly linked to the function of the WSN, and thus energy is a main contribution to the examination. The OPSSN and RLMQ approaches have been compared for residual energy in Fig. 7.

It illustrates that the RLMQ approach has the 0.13 Joule highest residual energy because using MS minimizes both the CH dead problem and the CH load. But, OPSSN presents the cluster burden and CH dead issues in the WSN.

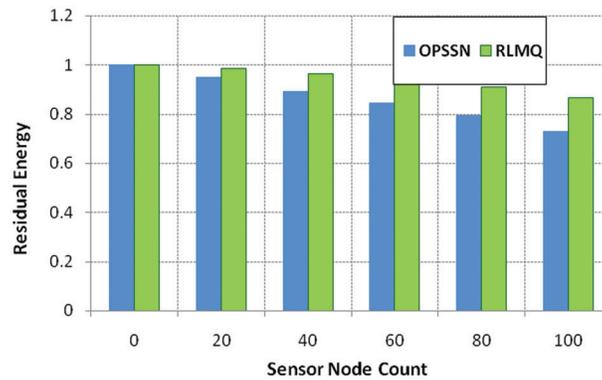


Figure 7: Residual Energy of OPSSN and RLMQ

• Routing Load

Routing Load is explained as the amount of control packets communicated for yielding a data pathway per data packet. It is defined as the ratio between the amounts of control packets forward to the amount of control packets received. Fig. 8 represents the routing load for OPSSN and RLMQ.

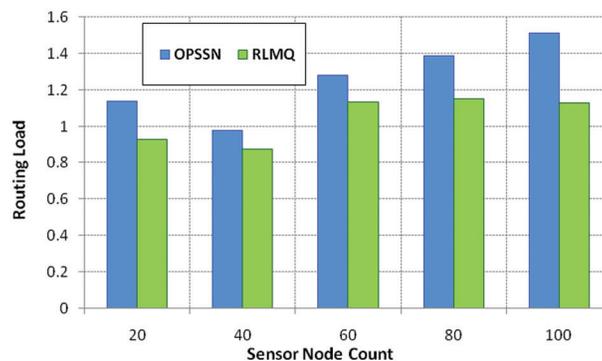


Figure 8: Routing Load of OPSSN and RLMQ

Fig. 8 displays that the proposed method RLMQ has the 25.4% least routing overhead equated to the existing method OPSSN in the WSN. The figure illustrates that the RLMQ approach has the least routing load in the WSN.

5 Conclusions

Machine Learning concepts provide very realistic outcomes to make possible the necessary QoS in the IoT environment. From the current studies, the QoS and machine learning to enhance the WSN performance are emphasized more with some vital challenges. Therefore, this paper considers RL Mechanism to improve the QoS and use MS to minimize the wireless IoT delay. The sleep and awake scheduling is used for minimizing the CH dead in the WSN. Here, the MS is used to collect the CH data, and the RL incentive values select CH. The incentives value is computed by the QoS parameters like minimum energy

utilization, minimum bandwidth utilization, minimum hop count and minimum time delay. The MS is used for collecting the data from CH; thus, it minimizes the network delay. Finally, the MS forwards the data to BS. From the results, the proposed method attains the required QoS and improves the energy efficiency in the WSN. Simulation results demonstrate that the proposed approach minimizes the network delay and diminishes the CH routing load.

Funding Statement: This research is done with the financial support by the Deanship of Scientific Research at King Khalid University under research grant number (RGP.2/241/43).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] N. A. Alduais, J. Abdullah, A. Jamil and L. Audah, "An efficient data collection and dissemination for IOT based WSN," in *IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conf.*, Vancouver, BC, Canada, pp. 1–6, 2016.
- [2] K. E. Khujamatov and T. K. Toshtemirov, "Wireless sensor networks based Agriculture 4.0: Challenges and apportions," in *Int. Conf. on Information Science and Communications Technologies*, Tashkent, Uzbekistan, pp. 1–5, 2020.
- [3] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat *et al.*, "IoT and agriculture data analysis for smart farm," *Computers and Electronics in Agriculture*, vol. 156, no. 9, pp. 467–474, 2019.
- [4] K. Jaiswal and V. Anand, "A Grey-Wolf based optimized clustering approach to improve QoS in wireless sensor networks for IoT applications," *Peer-to-Peer Networking and Applications*, vol. 14, no. 4, pp. 1943–1962, 2021.
- [5] R. K. Yadav and R. P. Mahapatra, "Hybrid metaheuristic algorithm for optimal cluster head selection in wireless sensor network," *Pervasive and Mobile Computing*, vol. 79, no. 3, pp. 1–16, 2022.
- [6] N. Ahmed, D. De and I. Hussain, "Internet of Things (IoT) for smart precision agriculture and farming in rural areas," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4890–4899, 2018.
- [7] H. Dutta and S. Biswas, "Medium access using distributed reinforcement learning for IoTs with low-complexity wireless transceivers," in *IEEE 7th World Forum on Internet of Things*, New Orleans, LA, USA, pp. 356–361, 2021.
- [8] C. Savaglio, P. Pace, G. Aloï, A. Liotta and G. Fortino, "Lightweight reinforcement learning for energy-efficient communications in wireless sensor networks," *IEEE Access*, vol. 7, pp. 29355–29364, 2019.
- [9] V. L. Narayanan and P. Selvan, "Energy-efficient Q learning based Kullback sparse encoder for traffic and congestion control data delivery in WSN," *International Journal of Nonlinear Analysis and Applications*, vol. 12, no. 2, pp. 2601–2618, 2021.
- [10] M. Y. Alnaham and I. A. M. Abdelrahman, "Improving QoS in WSN based on an optimal path from source to sink node routing algorithm," in *Int. Conf. on Computer, Control, Electrical, and Electronics Engineering*, Khartoum, Sudan, pp. 1–6, 2018.
- [11] M. Maalej, S. Cherif and H. Besbes, "QoS and energy-aware cooperative routing protocol for wildfire monitoring wireless sensor networks," *The Scientific World Journal*, vol. 2013, no. 12, pp. 1–11, 2013.
- [12] G. Merline, "Retrieval of dropped and modified packets by dynamic rerouting in wireless sensor networks," *International Journal of Engineering Research & Technology*, vol. 4, no. 2, pp. 594–599, 2015.
- [13] N. Javaid, A. Sher, W. Abdul, I. A. Niaz, A. Almogren *et al.*, "Cooperative opportunistic pressure based routing for underwater wireless sensor networks," *Sensors*, vol. 17, no. 3, pp. 1–26, 2017.
- [14] M. E. Migabo, T. O. Olwal, K. Djouani and A. M. Kurien, "Cooperative and adaptive network coding for gradient based routing in wireless sensor networks with multiple sinks," *Journal of Computer Networks and Communications*, vol. 2017, pp. 1–10, 2017.
- [15] A. Yahya, S. U. Islam, M. Zahid, G. Ahmed, M. Raza *et al.*, "Cooperative routing for energy efficient underwater wireless sensor networks," *IEEE Access*, vol. 7, pp. 141888–141899, 2019.

- [16] H. J. Abdul Nasir, K. R. Ku-Mahamud and E. Kamioka, "Enhanced ant colony system for reducing packet loss in wireless sensor network," *International Journal of Grid and Distributed Computing*, vol. 11, no. 1, pp. 81–88, 2018.
- [17] S. Abkari, J. Mhamdi and A. H. Abkari, "Real-time RSS-based positioning system using neural network algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 3, pp. 1601–1610, 2021.
- [18] T. Kaur and D. Kumar, "MACO-QCR: Multi-objective ACO-based QoS-aware cross-layer routing protocols in WSN," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6775–6783, 2020.
- [19] N. Li, S. McLaughlin and D. Laurenson, "Traffic-aware routing for wireless sensor networks in built environment," in *Fourth UKSim European Symp. on Computer Modeling and Simulation*, Pisa, Italy, pp. 397–401, 2010.
- [20] C. Park, H. Kim and I. Jung, "Traffic-aware routing protocol for wireless sensor networks," *Cluster Computing*, vol. 15, no. 1, pp. 27–36, 2012.
- [21] M. Gholipour, A. T. Haghghat and M. R. Meybodi, "Hop-by-hop traffic-aware routing to congestion control in wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–13, 2015.
- [22] S. Mohan Kumar and T. Kumanan, "Machine learning based driver distraction detection," *International Journal of MC Square Scientific Research*, vol. 12, no. 4, pp. 16–24, 2020.
- [23] S. Kumarapandian, "Melanoma classification using multiwavelet transform and support vector machine," *International Journal of MC Square Scientific Research*, vol. 10, no. 3, pp. 01–07, 2018.
- [24] Y. Tripathi, A. Prakash and R. Tripathi, "A sleep scheduling based cooperative data transmission for wireless sensor network," *International Journal of Electronics*, vol. 109, no. 4, pp. 596–616, 2022.