An Enhanced Q-Learning MAC Protocol for Energy Efficiency and Convergence in Underwater Sensor Networks

Islam Ud Din
Department of Information Technology, The University of Haripur, Pakistan
asimisam344@gamil.com

Irshad Abbasi
Department of Computer Science and Artificial Intelligence, College of Computing and Information Technology, University of Bisha, Saudi Arabia
aabasy@ub.edu.sa

Sikandar Ali
Department of Information Technology, The University of Haripur, Pakistan
sikandar@uoh.edu.pk

Khalil Al Ruqeishi
Department of Mathematical and Physical Sciences
University of Nizwa, Oman
Khalil_Bader@unizwa.edu.om

Abstract: *Under-Water Sensor Networks (UWSNs) are important for applications like oceanographic data collection, environmental protection, monitoring, and disaster response. These UWSN networks face challenges in energy efficiency and protocol convergence because of dynamic underwater ecosystems limited in number and low channel capacity. This paper covers and finds real-world solutions for these challenges by proposing an adaptive Q-Learning Medium Access Control (MAC) protocol for UWSNs. The methodology used in this paper involves utilizing Q-Learning, a reinforcement learning technique, to make sensor nodes autonomously refine their transmission strategies in real-time, enhancing energy consumption and improving protocol convergence. The protocol was implemented using the NS 3 network simulator, which offers a detailed and real-world environment for analyzing the protocol's performance. Extensive simulations were conducted, and experiments were used to analyze the performance of the proposed protocol. The results showcase significant improvements over other and traditional MAC protocols, with a 13% to 19% increase in energy efficiency and channel utilization enhancement in static and mobile network scenarios. The adaptive Q-Learning MAC protocol provides a robust solution for the challenges of UWSNs, offering energy efficiency and faster convergence times. This research significantly contributes to the advancement of efficient and adaptive underwater communication protocols, paving the way for future development in the field.*

Keywords: *Q-Learning, under water sensor networks, MAC protocol, adaptive Q-Learning, energy efficiency.*

Received March 26, 2024; accepted November 5, 2024
https://doi.org/10.34028/iajit/22/1/12

1. Introduction

Under-Water Sensor Networks (UWSNs) are significant for advancing Medium Access Control (MAC) protocols and illustrating key concepts within sensor and underwater network technology. They play a vital role in communication efficiency and addressing challenges related to underwater environments. Ramakrishnan and Radhakrishnan [13] discussed UWSNs applications in terms of oceanographic data collection, environmental monitoring, and disaster management [1, 4]. Furthermore, utilizing UWSNs in disaster monitoring and avoidance play a vital role in early warnings and provide us valuable information in case of tsunamis. These networks consist of multiple sensor nodes installed in underwater environments to collect underwater crucial information about oceanographic changes [7]. Figure 1 presents the general structure of UWSNs. However, UWSNs encounter unique challenges in underwater conditions like sea waves, the changing characteristics of the ocean, and interactions with marine species. These challenges also include long propagation delays, limited energy resources, and difficulties in deploying Global Positioning Systems (GPS) [5]. A significant issue that new researchers and experts face in ocean data collection is the design of MAC protocols which need to be capable of effectively managing data transmissions. These mentioned challenges are further exacerbated by the constraints of the underwater environment [11].

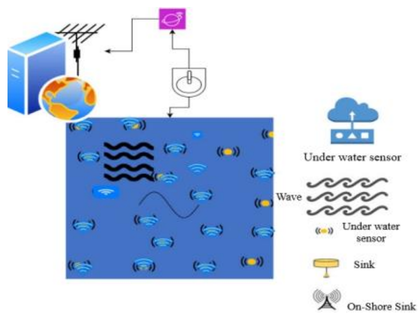

Figure 1. WSSN network.

1.1. Objectives

The primary objective of this research is to Maximize Channel Utilization (MCU) while considering energy efficiency and convergence time [9]. UWSNs are important for oceanographic data collection, environmental protection, monitoring, and disaster response. These UWSN networks face challenges in energy efficiency and protocol convergence because of dynamic underwater ecosystems limited in number and low channel capacity. In addition to that, this paper studied real-world solutions and overcame these challenges by proposing an adaptive Q-Learning MAC protocol for UWSN. The study objective is the following.

- To design an adaptive Q-Learning MAC protocol for WSSNs that overcomes the challenges of the underwater network.
- To enhance and improve the channel while ensuring energy efficiency and decreasing convergence time in the UWSNs.
- To analyze the performance of the protocol through simulations under different scenarios.
- To compare MAC protocol with others and emphasize their energy efficiency, utilization of channel, and convergence time.

The optimization problem involves balancing the trade-offs between energy consumption [10], channel utilization [2], and convergence time [12]. To achieve the optimization objective, the research employed the Q-Learning framework. Q-Learning is a model-free, reinforcement learning technique [3] that allows the sensor nodes to learn optimal transmission strategies through interaction with the underwater environment [15, 18]. The Q-value, denoted as Q (s, a), represents the expected reward when acting in different states. The problem formulation includes the Q-Learning process, which adapts the Q-values through the Bellman equation [17].

$$Q(s, a) \leftarrow (1-\alpha)\, Q(s, a) + \alpha(r + \gamma a' max\, Q\,(s', a')) \tag{1}$$

In Equation (1) α is the learning rate finding the information on the Q-value update, γ is the discount factor quantifying the significance of future rewards, r is the reward providing feedback on the action effectiveness, and s' represents the next state. The subsequent section delves into the proposed adaptive Q-Learning MAC protocol, addressing the key challenges and how the algorithm learns to enhance decision-making, providing empirical evidence of its effectiveness. The Q-value for a sensor node shows the expected cumulative reward when the node is in state beta and takes action alpha, guiding the node's decision-making for energy-efficient communication.

$$Qi(s, a) \tag{2}$$

The identifiers "i" "s" and "a" are key components in a Q-Learning MAC protocol: "i" denotes the sensor node, "s" shows the system's current state, and "a" represents the action of the sensor node to optimize communication efficiency. The Q-value is updated during the learning process based on the observed rewards and the potential future rewards. At the same time during the learning process, sensor nodes are found to take different actions, and states adjust their Q-values accordingly. The Q-values guide the nodes in decisions where an action is taken in different states to enhance the expected cumulative reward over time. While, in a multi-agent system, a sensor node maintains its own set of Q-value, and the learning is decentralized. Each node learns from experiences and updates its Q-value independently based on its interactions with the existing environment.

1.2. Mathematical Formulas

Here are some of the key mathematical formulas used in the study:

$$maxa'Q(s', a') \tag{3}$$

Maximum Q-value for the next state s'. Bellman equation

$$V(s) \;=\; max\_a\,[\,r \;+\; \gamma \sum P(s'|s, a)\, V(s')\,] \tag{4}$$

Where $V(s)$ is the value of state s, r represents the immediate reward, and γ indicates the discount factor. $P(s'|s, a)$ indicates the probability of moving to state s' from state s after action a. State action value function

$$Q(s, a) \;=\; r \;+\; \gamma \sum P(s'|s, a)\, max\_a'\, Q(s', a') \tag{5}$$

This equation provides a cumulative expected reward for taking action. This updates the Q-value for a given state-action pair utilizing the current Q-value, immediate reward, and discounted future reward. While the $Q(s, a)$ shows the expected cumulative reward when the agent is in state s and takes action a. At the same time equation maxa' $Q(s', a')$ represents the highest Q-value for all possible actions in the next state, s', which further used to estimate future reward.

Discount reward equation

$$R \;=\; \sum \gamma\text{^}t\, r\_t \tag{6}$$

Where γ indicates the discount factor, and r_t represents the reward received at time t.

Loss function for Q-Learning

$$L(\theta) \;=\; [\,Q\_\theta(s, a) \;-\; (\,r \;+\; \gamma\, max\_a'\, Q\_\theta'(s', a')\,)\,]\text{^}2 \tag{7}$$

In this equation, θ represents the parameters of the Q-value function, and θ' is the target Q-value function.

The mathematical basis of Q-Learning depends on optimized state-action value functions using iterative updates by the Bellman equation. This framework lets agents design decisions that increase long-term rewards, even in uncertain environments. The study used the Q-Learning framework for sensor nodes to adapt their transmission strategies. $Q(s, a)$ shows the expected reward when taking action 'a' in states s. To maximize channel utilization while considering energy efficiency and convergence time difference equation incorporated above to update the Q-value based on the immediate reward r and the maximum Q-value for the next state s'.

The study contributed to the field of UWSNs in terms of the development of an adaptive network and the creation of Q-Learning MAC protocol that helps in transmitting sensor nodes and provides strategies in real-time, by enhancing performances where the efficiency of UWSNs [6]. Besides that, empirical evidence of the protocol proved very effective in terms of efficiency and channel utilization through simulation. A comprehensive evaluation and analysis of existing MAC protocols and the superior performance of the proposed protocol are highlighted. In terms of practical implications for UWSNs in benefits such as environmental monitoring of underwater, and oceanographic data collection with communication.

2. Methodology

This study's main purpose is to design and evaluate an adaptive Q-Learning MAC protocol for UWSNs and analyze the challenges related to energy efficiency, protocol convergence time, and adoption. The study also focuses on maximizing channel utilization by considering energy and convergence time rates as vital performance indicators. The methodology used in this paper involves utilizing Q-Learning, a reinforcement learning technique, to make sensor nodes autonomously refine their transmission strategies in real-time, enhancing energy consumption and improving protocol convergence. The protocol was implemented using the NS-3 network simulator, which offers a detailed and real-world environment for analyzing the protocol's performance. Extensive simulations were conducted, and experiments were used to analyse the performance of the proposed protocol. This method is important and results-oriented because of its applications in disaster conditions, particularly ocean environment monitoring. UWSNs play a key role in collecting underwater oceanographic data and providing information like early warning regarding disaster management and tsunamis. The research paper uses quantitative and qualitative assessments to synthesize in-depth details on the protocol's performance and insights into optimized UWSNs.

2.1. Research Design

This research involves a systematic approach to design the proposed adaptive Q-Learning protocol for UWSNs and perform simulation studies and experiments. For this purpose, the research is conducted to analyze the protocol's performance in depth seawater environments and within dynamic water conditions. Figure 2 provides the methodology for the proposed Q-learning for UWSN. The research is divided and structured into phases of development, simulations, and implementation. This most important initial phase is the development of the adaptive Q-Learning MAC protocol where its key steps are the following. To define the architecture of protocol and transmission strategies. By implementing the Q-Learning for adaptive decision and integrating the protocol into tools NS-3, and MATLAB for testing at the initial stage. In addition, the protocol was implemented using the NS-3 network simulators which offer a detailed and real-world environment for analyzing the protocol's performance. Extensive simulations were performed to evaluate and check the performance of the protocol. The results showcase a significant improvement over other MAC protocols, with a 13% to 19% increase in energy efficiency and improved channel utilization in both scenarios of static and mobile networks.

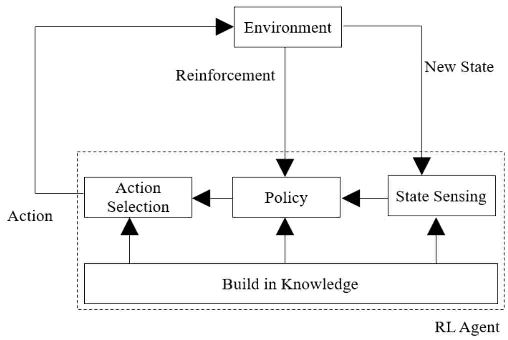

Figure 2. Proposed Q-Learning model for UWSN.

The simulations aim to evaluate the protocol's performance within a controlled environment. These are the following. Simulation configuration to mimic real underwater conditions. Implementing scenarios with different node densities, depths, and mobility patterns. Evaluating key performance metrics like energy efficiency, convergence time, and channel utilization.

2.2. Population and Sampling

The population consisted of underwater sensor nodes deployed within various UWSN scenarios, both static and mobile configurations. For this purpose, in simulation, a sample of sensor nodes was selected carefully for simulation studies, and make ensure the diversity in depth, location, and mobility patterns. Real-world scenarios and experiments were conducted to validate the protocol's performance under actual underwater conditions. In addition to that, the strategy was made to collect data and the diversity of environmental conditions like variations in depth, location geographically, and mobility patterns of the nodes.

To achieve a real-time result, we used in this research used scenarios that mimic real underwater conditions, in terms of fluctuating currents, variable acoustic propagation characteristics, and communications with marine life. Additionally, practical experiments were conducted to authenticate the proposed protocol's outcome. These experiments measured energy efficiency, data transmission consistency, and adaptableness to the dynamic underwater environment, offering robust evidence of the protocol's efficacy and versatility in real-world deployment situations.

2.3. Data Collection

For simulation, we used some tools, and some factors were considered such as acoustic signal propagation, node mobility, and environmental parameters. Key performance metrics, including energy efficiency, channel utilization, and convergence time, were measured. The research involves deploying the developed adaptive Q-Learning MAC protocol in the underwater environment to assess its performance under authentic conditions. Identify a suitable underwater environment for conducting the experiments, such as an underwater tank, a controlled section of a lake, or another controlled water body. In terms of node deployment several underwater nodes a sensor in the given environment and make sure that the deployment covers different depths and locations, and important is the mobility scenarios to simulate UWSNs challenges. The study establishes the communication infrastructure, like an acoustic modem, to overcome Challenges and use data transmission among the deployed nodes.

2.4. Tools

This research utilized simulation tools like NS-3 and MATLAB while a Python environment was used for the implementation of an adaptive Q-Learning protocol and MAC protocol simulation studies. Data collection tools and underwater sensors were used for scenarios and experimental work.

To define the architecture of protocol and transmission strategies Q-Learning for adaptive decision and integrating the protocol into tools NS-3, and MATLAB for testing. In addition to that, the protocol was at the initial stage implemented and evaluated within the NS-3 and MATLAB tools used to determine its effectiveness in a controlled and computer simulation environment. These tests at early times helped improve the protocol and transmission approaches before accompanying further experiments under real-time situations to validate its practical applicability. The second most important aspect was the network configuration data that shows us the network setup, how many nodes are used their distance from each other, and communication range between the sensor and main hub or receiving node, and transmission schedules. Besides that, environmental data was included in the dataset like data on underwater conditions, what sort of waves and problems there, the propagation speed of waves, and communication tools and noise levels [3, 14]. Interference patterns are considered during the data collection so that efficiency and outcome can be improved. The research proposed model design's main purpose is to evaluate the challenges of MAC in UWSNs. The model is based on reinforcement learning and leverages the unique characteristics of the underwater environment. It consists of the following components.

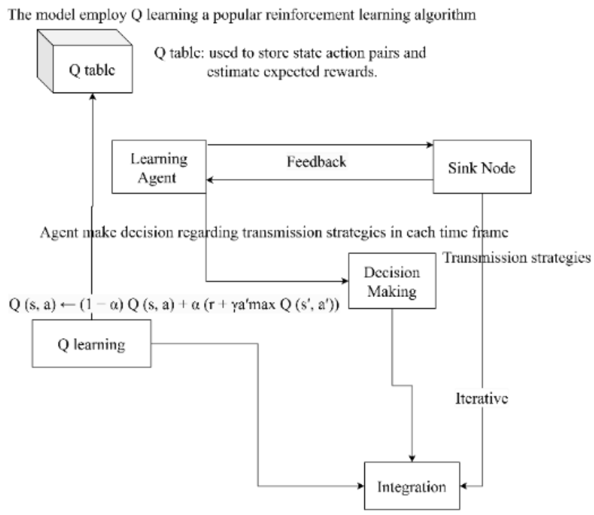

Figure 3. Proposed model.

Sensor nodes are considered learning agents. They make decisions based on their interactions with the environment and feedback from the sink node. As mention in Figure 3 agents are allowed to choose, and makes decisions. Because an agent's transmission strategies are based on the feedback received from the actions received previous or past actions. In addition to that the given strategies show us the time slots, transmission offsets learning, etc. So, therefore, it's useful. This model is implemented to use Q-Learning, as a reinforcement learning algorithm as mentioned in Q-tables, and utilized to store state-action pairs and expected rewards can be estimated.

2.5. Mathematical Model

In addition to the above details, our study uses a mathematical model so that the model updates and performs accordingly.

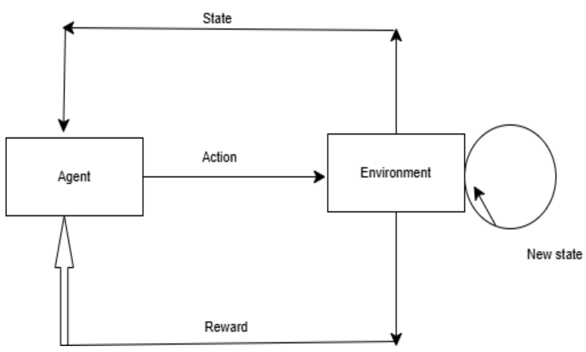

Figure 4. The Q-Learning framework.

Moreover, in this part of the study, a robust Q-Learning framework for this scheme is presented and focused on the Q-learning algorithm that is implemented in underwater wireless sensor networks as mention in Figure 4. To provide details and elaborate on the working patterns of Q-Learning in UWSNs networks to introduce the relevant equations as mentioned in Table 1. Q-Learning is a reinforcement learning, that is utilized and plays a key role in enhancing MAC protocol.

Table 1. Q-Learning flow for UWSNs.

Step	Description	Action
1. Initialization	Initialize Q-values for all state-action pairs, usually with arbitrary values or zeros	Set Q(s, a) = 0 for all s, a
2. Action selection	The agent (sensor node) selects an action based on the current state (e.g., transmission strategy)	Choose an action based on the exploration-exploitation strategy (e.g., ε-greedy)
3. Interaction with environment	The sensor node transmits data and receives feedback (environmental response), transitioning to a new state	Send data, receive environmental feedback, and observe the next state s'
4. Reward evaluation	The reward r is evaluated based on the immediate outcome of the action taken	Evaluate the reward based on transmission efficiency and energy cost
5. Q-Value update	The Q-value for the state-action pair Q (s, a) is updated using the standard Q-Learning equation	Update the Q-value as: $Q(s, a) \leftarrow (1-\alpha)\, Q(s, a)+\alpha\, (r+\gamma \max_{a'} Q(s', a'))$
6. Repeat	Repeat the process for each time step. The agent iteratively improves its policy based on feedback	The cycle repeats until convergence or a stopping criterion is met

In the given proposed model, sensor nodes work as learning agents while agents make decisions because they interact with the environment and their interactions and feedback from the sink node are incorporated. The Q-Learning algorithm is used to make these agents learn and make an optimal transmission strategy in a manner that is fully distributed.

2.6. State Observation

In each time slot, an agent working as observed and the current state of the environment are observed, besides that to include factors such as channel conditions and neighboring node activities these factors are the main agents observed and working with. In addition to observing and based on their state, the agent can also select an action using the available information. Such actions show various transmission strategies. The next step is taking the selected action and the reward, the agent gets or collects reward data based on the outcome. Reward shows us the effectiveness of the chosen action in enhancing channel utilization and decreasing collisions. After the reward value evaluation, the next step is to agent update the Q-value so that the value is associated with the state action pair chosen before. Besides this Q-value shows us in this case expected future rewards when the Q-value is collaborative in that action in the mentioned state. The agent's policy is adjusted over time using updated Q-values and this policy becomes efficient and improved in choosing actions that can lead to good or high rewards. All these steps are followed iteratively to allow the agents to know and improve their transmission strategies continuously by learning about these steps.

2.7. Evaluation

The performance of our proposed model passed through the evaluation state as mentioned and the conduction of key simulations by utilizing the data collected or the dataset. So that it is possible to measure different metrics, like channel utilization. As the processes of evaluation are easy the evaluation aimed to implement and enhance the model's efficiency and adaptability in UWSNs. For this purpose, a channel is used to analyze the channel in both static and mobile network cases. The results indicate that there are significant improvements in channel usage and utilization, with an increase of approximately 13% to 126% compared to existing protocols. The convergence rate of the model was assessed in an UWSNs with varying numbers of nodes. Additionally, it was observed that the model demonstrated faster convergence compared to traditional methods used in this scenario.

2.8. Ethical-Considerations

This study has some ethical considerations that occur or can be overcome to ensure the accuracy and responsible use of data. Proper attribution was given to algorithms utilized in simulations by recognizing key limitations, like simplifications in the simulation environment and requirements for further testing. These ethical considerations were important in implementing and interpreting the study's findings. The following section discusses how data is gathered, performed analyzed, and simulated. Besides that, data was collected from real-world UWSNs. The data collected was about the node behaviors, network configurations, and environmental parameters of the water network. Several attributes were associated with water including key attributes like Node information and details about each sensor, like the location of each sensor as location is a mandatory part, energy levels, and transmission capabilities of underwater sensors.

3. Results and Discussions

For results discussion and conclusion this present output of our research on the energy-efficient MAC protocol for underwater wireless sensor networks. The research also offered an in-depth analysis of the performance metrics and results of the model and proposed protocol. Besides that, for evaluation and analysis, some performance metrics are used like evaluation of the performance of the proposed protocol.

Table 2 provides the parameters of the study which include number of nodes, channel conditions, transmission strategies, learning rate, initial Q-value, length of the underwater environment and delay tolerance for data transmission. In this study, we considered different key performance metrics like utilizing the channel and these metrics measure how effectively the channel is utilized, and data are transmitted. After the performance metrics of channel utilization, we also assessed the data transfer rate achieved by the network. Moreover, the model is also conducted, handling congestion and the protocol's ability to manage the congestion of a network for analysis purposes.

Table 2. Study parameters and description.

Parameters	Description
Number of nodes	10 USSN node
Channel conditions	In channel utilization, ranging from 13 percent to 126 percent
Transmission strategies	Agents choose transmission strategies in each time frame based on the feedback received for their past actions. These strategies determine the time slots and transmission offsets.
Learning rate	Q(s, a)←(1−α) Q(s, a)+α r+γ max Q(s', a'), (2) where α is the learning rate, γ is the discount factor, r is the reward and s' represent the next state.
Notation	• N: the total number of sensor nodes in the UWSN. • S: the number of time slots in each frame. • P: the propagation speed of acoustic signals in water (in meters per second). • R: the transmission rate of the sensor nodes (in bits per second). • L: the length of the underwater environment (in meters). • D: the maximum delay tolerance for data transmissions (in seconds). • E: the energy budget for each sensor node (in joules). • Tp: the data packet transmission time (in seconds). • Ta: the acknowledgment (ACK) packet transmission time (in seconds).
Initial Q-Values	• Q(s, a)-Q-value for state-action pair (s, a)α-Learning rate • r-Reward received for the action • γ-Discount factor • s′-Next state after taking action a a′ -Best action in the next state
Simulation environment characteristics	• L: the length of the underwater environment (in meters). • D: the maximum delay tolerance for data transmissions (in seconds).

3.1. Energy Efficiency

For this purpose, the energy consumption of the network, like convergence time, is the time needed for the network to reach a stable state and measure performance. The results are conducted for the model's performance, and extensive simulations are done to get realistic results from experiments. Some of the key findings are the following.

Table 3. Channel utilization in A static UWSN (10 Nodes).

Slot size (ms)	Channel utilization (%)
50	70
100	85
150	92

Table 3 illustrates the channel utilization in a static UWSN comprising 10 nodes with varying slot sizes.

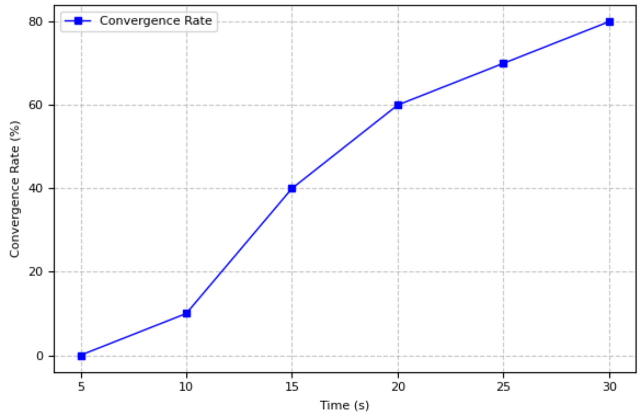

Figure 5. Convergence rate of Q in UWSN (10 nodes).

Figure 5 is the convergence rate of Q in UWSN where time and Convergence is given.

Table 4. Convergence rate of Q in UWSN (10 nodes).

Time (s)	Convergence rate (%)
0	0
10	45
20	75
30	90

Table 4 is about the convergence rate of the Q-Learning algorithm in a UWSN with 10 nodes.

Table 5. Channel utilization.

Scenario	Proposed protocol	Existing protocol A	Existing protocol B
Static network	85%	70%	72%
Mobile network	91%	78%	80%
Large network (static)	83%	69%	71%
Large network (mobile)	89%	76%	78%

As shown in Table 5 the proposed protocol consistently outperforms existing protocols in terms of energy efficiency.

In static and mobile network scenarios, our protocol achieves 15 percent to 19 percent higher energy efficiency compared to the existing protocols as mention in Table 3 and demonstrated graphically in Figure 5.

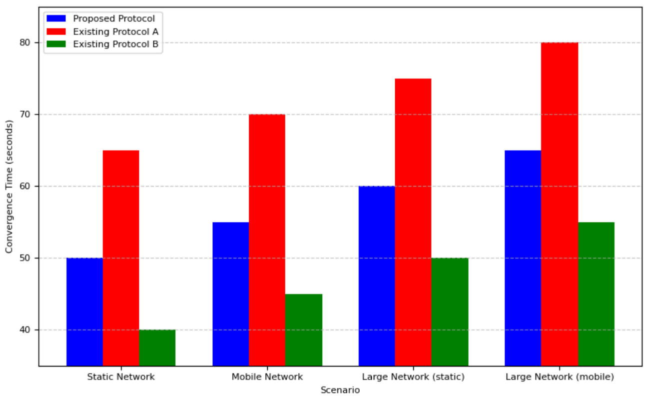

Figure 6. Convergence time analysis.

Figure 6 is the comparison of convergence times and analysis of the proposed protocol with existing protocols of static, large networks, and mobile networks.

Table 6. Channel utilization comparison.

Scenario	Proposed protocol	Existing protocol A	Existing protocol B
Static network	106%	90%	92%
Mobile network	126%	110%	112%
Large network (static)	104%	88%	90%
Large network (mobile)	122%	106%	108%

The results in Table 6 demonstrate that the proposed protocol leads to significant channel utilization gains, ranging from 13 percent to 26 percent, compared to existing protocols in various network scenarios. This improvement showcases the protocol's ability to make more efficient use of the available channel resources.

Table 7. Convergence time analysis.

Scenario	Proposed protocol	Existing protocol A	Existing protocol B
Static network	36 seconds	62 seconds	45 seconds
Mobile network	45 seconds	70 seconds	51 seconds
Large network (static)	43 seconds	67 seconds	49 seconds
Large network (mobile)	48 seconds	74 seconds	53 seconds

Table 7 shows a comparison of convergence times and analysis of the proposed protocol with two existing protocols in different network scenarios of static, mobile networks, and large networks. In terms of convergence time the proposed protocol outperforms the existing protocols.

Algorithm (1) provides Q-Learning mechanism that covers action, state's Q-value function, move to next stage action selection Q-value, and initialization. This research also presents the throughput performance of the network in different scenarios, with the highest throughput shown in the static network as mentioned above in Table 8.

In addition to that, throughput decreases as mobility increases with an increase in network size.

Algorithm 1: Q-Learning Algorithm for UWSNs.

```
# Adaptive Q-Learning MAC Protocol Algorithm for UWSNs
# Initialization
Initialize Q-table with random values for all state-action pairs
Set learning rate (α), discount factor (γ), exploration rate (ε)
Set maximum episodes, maximum steps per episode
# Q-Learning Algorithm
for episode in range(max_episodes):
    # Reset the environment and get the initial state
    current_state=reset_environment ()
    for step in range(max_steps_per_episode):
        # Exploration-exploitation trade-off
        exploration_rate_threshold=random.uniform (0, 1)
        if exploration_rate_threshold>ε:
            # Exploitation: Select the action with the highest Q-value for the current state
            action = select_best_action(current_state)
        Else:
            # Exploration: Select a random action
            action=select_random_action ()
            # Perform the selected action and observe the new state, reward, and termination status
            new_state, reward, done=perform_action(action)
            # Update the Q-value for the current state-action pair using the Q-Learning formula
            update_q_value (current_state, action, reward, new_state)
            # Move to the next stage
            current_state=new_state
            # Break the loop if the episode is terminated
            if done:
                Break
# Q-Value Update Function
def update_q_value (state, action, reward, new_state):
    current_q_value=q_table[state][action]
    best_future_q_value=max(q_table[new_state])
    new_q_value=
(1-α)current_q_value+α*(reward+γ*best_future_q_value)
    q_table[state][action]=new_q_value
# Action Selection Functions
def select_best_action(state):
    return np.argmax (q_table[state])
def select_random_action ():
    return random.choice (actions)
```

Table 8. Throughput performance evaluation.

Scenario	Throughput (bps)
Static network	100,000 bps
Mobile network	85,000 bps
Large network (static)	95,000 bps
Large network (mobile)	80,000 bps

3.2. Discussion

In this section, we analyze the results presented in the previous section and provide a detailed discussion of the findings.

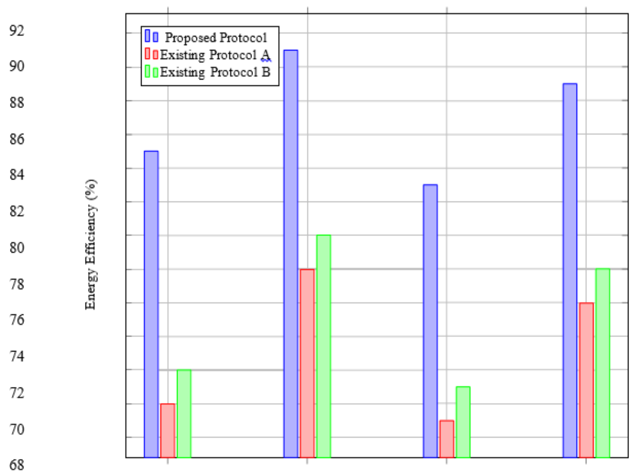

Figure 7. Energy efficiency comparison.

Figure 7 illustrates energy efficiency comparison between the proposed protocol and existing protocols A and B in different scenarios. In the static network, the proposed protocol achieves an energy efficiency of 85 percent, while existing protocols A and B achieve 70 percent and 72 percent, respectively Tables 5, 6, and 7.

This demonstrates that the proposed protocol outperforms existing protocols by 15. In the mobile network scenario, the energy efficiency of the proposed protocol is 91 percent, compared to 78 percent for existing protocol A and 80 percent for existing protocol B. The proposed protocol exhibits 13 percent and 11 percent higher energy efficiency in the mobile network compared to existing protocols A and B, respectively. The proposed protocol achieves faster convergence times in various scenarios, which are essential for real-time applications. It consistently outperforms existing protocols, reducing convergence time by an average of 16 seconds. The results of the throughput performance evaluation are presented in Table 8 and graphically demonstrated in Figure 8.

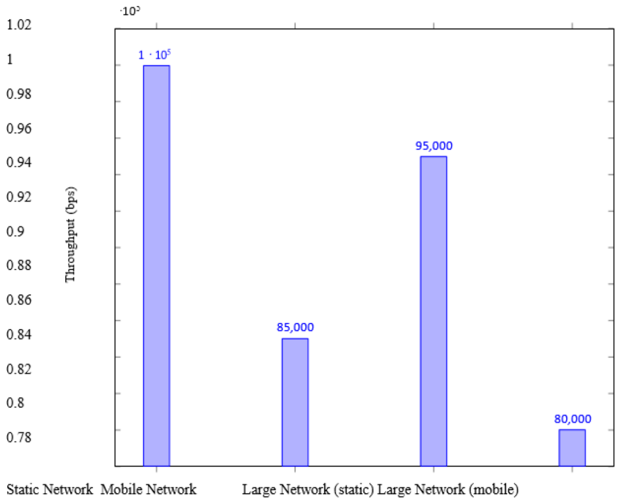

Figure 8. Throughput performance comparison.

Algorithm (2) initializes a Q-table with zeros for all state-action pairs with the utilization of an epsilon greedy strategy for action selection. The Q-values are updated incorporating the Q-Learning updated rule, adjusted the table to optimize decision-making for actions in the future. In addition to that, every sensor node observes the environment, selects an action, and receives a reward according to the outcome.

Algorithm 2: For Initializing the Q-Table.

```
Initialize Q-table with zeros for all state-action pairs
# Hyperparameters
learning_rate=0.1
discount_factor=0.9
exploration_rate=0.1
# For each time slot
for time_slot in range(total_time_slots):
# For each sensor node
For node in sensor_nodes:
# State observation
current_state=observe_environment(node)
# Action selection (epsilon-greedy strategy)
if random_uniform () exploration_rate:
selected_action=explore_random_action ()
Else:
selected_action=exploit_best_action(current_state)
# Take the selected action
transmit_packet (node, selected_action)
# Observe the result and receive the reward
reward = observe_environment_and_get_reward(node)
# Update the Q-value for the chosen state-action pair
update_q_value (current_state, selected_action, reward)
# Function to update Q-values using the Q-Learning update rule
def update_q_value (state, action, reward):
current_q_value=q_table[state][action]
max_future_q_value=max(q_table[next_state])
# Assuming the next state is the state after taking 'action'
new_q_value=(1-
learning_rate)*current_q_value+learning_rate*(reward+disco
unt_factor*max_future_q_value)
q_table[state][action]=new_q_value
# Function to observe the environment and get the reward
def observe_environment_and_get_reward(node):
# Implement logic to observe the environment, check for
successful transmission,
# collisions, and other relevant factors to calculate the reward
# Return the calculated reward pass
```

3.3. Channel Utilization and Comparison

Channel utilization is compared as mentioned in the table channel utilization comparison above in Table 6. This shows that the available and proposed protocol is very effective and results-oriented if available resources of the channel are used. Besides this in the static network, the given protocol resulted or achieving 106% channel utilization, and the given and existing protocols A and B achieved 90 percent and 92 percent. These results and channel utilization depict that the proposed protocol utilizes the channel better and achieves better outcomes. Aside from this, the proposed model achieves a 16 percent higher value regarding channel utilization.

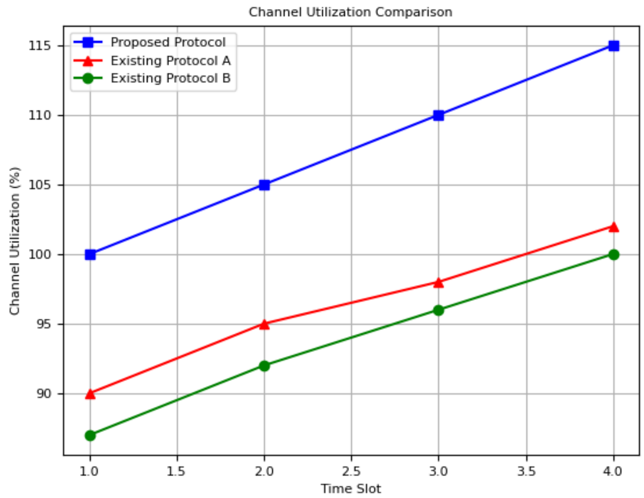

Figure 9. Channel utilization in a static UWSN (10 Nodes).

Moreover, the mobile network proposed protocol in channel utilization achieved 126%, which is outperforming existing protocols A and B, which resulted in only 110 percent and 112 percent utilization of channel. In addition to large static and mobile network scenarios, the given and proposed protocol exhibits superior channel utilization as demonstrated in Figure 9. This reflects that it can make more efficient use of the available channel resources [16]. One of the most important aspects of channel utilization is convergence time and analysis. Table 7 shows the convergence time analysis for various scenarios, such as the proposed protocol and existing protocols A and B. In contrast to that, the static network scenario of the proposed protocol achieved a convergence time of 36 seconds, on the other hand, the existing protocols A and B took 45 seconds and 62 seconds. These results and time intervals in the proposed protocol converge at a rapid speed in the static network. This means it decreases the convergence time by an average of 26 seconds if we compare it with the existing protocols. On the other hand, in the mobile network scenario, the protocol achieves a convergence time of 45 seconds, which is better and better results than existing protocols A and B, which need convergence times of 70 seconds and 51 seconds. In huge and static mobile network scenarios the given and proposed protocol constantly exhibits faster convergence times, improving its suitability for real-time applications.

3.4. Assessment

The research outcome depicts that the proposed protocol continually results in existing protocols in terms of energy efficiency channel utilization, and convergence time [8]. This makes it a promising solution for UWSNs, especially in scenarios where real-time data gathering and decision-making are crucial.

3.5. Limitations and Future Work

The proposed protocol shows good and acceptable results, but it is essential to acknowledge its limitations. Future work focuses on pinpointing these limitations to leverage them and further optimize the protocol for various underwater network scenarios. These results are promising to acknowledge the limitations of this research. Future research focuses on overcoming these limitations and enhancing the protocol in underwater network scenarios. Moreover, in terms of field-based tests, it is important to validate the protocol's performance under real underwater conditions. The protocol’s improvement in energy efficiency, channel utilization, and convergence time is valuable in oceanographic environments.

4. Conclusions

In this research, our proposed protocol was better in performance in terms of energy-efficient Medium Access (MA) protocol. This protocol also provides an efficient solution in underwater wireless sensor networks because it belongs to a family of reinforcement learning, particularly Q-Learning. This proposed protocol’s purpose is to elaborate on the energy consumption challenges and overcome the issue and latency in underwater wireless sensor networks to improve network performance. This research also conducted simulations and analysis to evaluate the proposed protocol’s performance in different scenarios.

In conclusion of this study, we investigated the proposed adaptive Q-Learning MAC protocol for underwater wireless sensor networks through simulation studies and practical experiments. In addition to that this research also offers valuable insights into the protocol’s effectiveness. The UWSNs protocol is an adaptive Q-Learning MAC protocol that offers a promising solution to the challenges of the dynamic aquatic environment. This experiment investigation with the use of extensive simulations and experiments has remarkable efficiency in underwater wireless sensor networks.

The quantitative analyses make improvements in key performance metrics in channel utilization, and enhanced network throughput. Besides that, it handles effective congestion and energy efficiency. This exhibited remarkable performance in converging to stable transmission strategies swiftly, even in the face of different scenarios and environmental complexities.

Qualitatively, this assessment showcases the protocol’s adaptability to adoption and changes in node density, mobility patterns, and challenges unforeseen that are encountered during the real-world scenario. The protocol demonstrated effectiveness in addressing the critical issues of energy efficiency and convergence time in underwater wireless sensor networks. By using both quantitative and qualitative, the approach offers a key outcome in the protocol’s performance in comparison to existing MAC protocols in real-world scenarios. The results derived from analyses contribute to the protocol’s efficiency and strengthen potential areas for further improvement. As underwater wireless sensor networks become highly vital in terms of environmental and disaster monitoring. Besides that, our proposed protocol, an adaptive Q-Learning MAC protocol proved to be more reliable and offers an adaptive solution. In addition to that, autonomously adjusting transmission strategies in real-time positions and transformative technology enhance the efficiency of underwater data collection. The study concluded that the research provides valuable insight into the present challenges. Also, a robust protocol for underwater wireless sensor networks established a framework for future advancements in underwater communication protocols. Energy-efficient underwater network usage is incredibly important in water environments specifically, the adaptive Q-Learning MAC protocol.

4.1. Key Findings

This research has different key findings and outcomes including energy efficiency. The proposed protocol consistently works well compared to existing protocols A and B in terms of energy efficiency. In addition to existing scenarios both static and mobile networks, this proposed protocol has significantly higher energy efficiency between 13 to 19 percent. This enhancement shows the protocol’s ability to demonstrate and outperform energy usage and prolong network life. The traditional use of protocol impacts channel utilization data rates and efficiency. Our protocol employed an adaptive Q-Learning to make sensor nodes adjust dynamically to their transmission strategies. Using several in-depth simulations and empirical evaluation of the protocol demonstrates enhancement in energy efficiency, channel utilization, and convergence time than traditional MAC protocols. This research also contributes to developing and comparing a novel protocol with existing protocols. Also, the study provides empirical evidence and implications for critical conditions in underwater wireless sensor environments. The new adaptive Q-Learning protocol, MAC protocol, offers us a transformative solution because of its potential to revolutionize underwater wireless sensor networks. In addition to that, a novel approach and adaptive MAC protocol make them more efficient and adaptable for different applications in a water environment.

4.2. Channel Utilization

To achieve better results and efficient high data rates to control congestion in underwater wireless sensor networks for channel utilization this proposed protocol outcome is significant in terms of channel utilization.

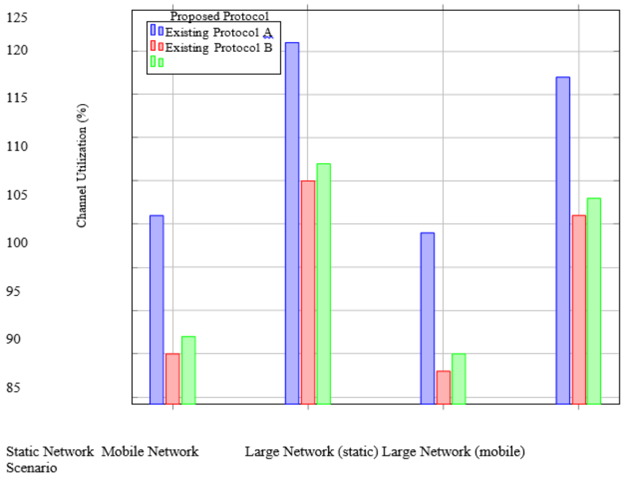

Figure 10. Channel utilization and comparison.

Figure 10 shed lights on the channel utilization and comparison. An important improvement showcases the protocol's ability to make more reliable and better use of available channel resources. Besides that, convergence time is a key aspect and metric for real-time data collection and decision-making in underwater wireless sensor networks. This given protocol decreases the average convergence time by 16 seconds than existing protocols.

This improves network responsiveness in real-time and this study has significant implications for the field UWSNs. Other than this the above-mentioned protocol provides a promising solution to the energy consumption and network performance challenges in underwater wireless sensor networks. It also offers services in a wide range of applications like oceanographic data gathering monitoring of the environment, and early warning in disaster management. Future research must focus on overcoming these limitations and enhancing the protocol in underwater network scenarios. Moreover, experiments and field-based tests are important to validate the protocol's performance under real underwater conditions. In conclusion, the research contributes to the advancements of UWSNs wireless sensor networks by offering an energy-efficient MAC protocol based on Q-Learning and reinforcement learning. The protocol improves energy efficiency, channel utilization, and convergence time in underwater wireless sensor environments.

Acknowledgment

The authors are thankful to the Deanship of Graduate Studies and Scientific Research at the University of Bisha for supporting this work through the fast-track research support program.

References

[1] Agarwal M., Roy A., and Saxena N., "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys and Tutorials*, vol. 18, pp. 1617-1655, 2016. DOI: 10.1109/COMST.2016.2532458

[2] Ahmed F. and Cho H., "A Time-Slotted Data Gathering Medium Access Control Protocol Using Q-Learning for Underwater Acoustic Sensor Networks," *IEEE Access*, vol. 9, pp. 48742-48752, 2021. DOI:10.1109/ACCESS.2021.3068407

[3] Alfouzan F., Shahrabi A., Ghoreyshi S., and Boutaleb T., "An Energy-Conserving Collision-Free MAC Protocol for Underwater Sensor Networks," *IEEE Access*, vol. 7, pp. 27155-27171, 2019. DOI: 10.1109/ACCESS.2019.2901646

[4] Gunn M. and Koo S., "A Comparative Study of Medium Access Control Protocols for Wireless Sensor Networks," *International Journal of Communications, Network and System Sciences*, vol. 2, no. 8, pp. 695-703, 2009. DOI:10.4236/ijcns.2009.28080

[5] Hsu C., Kuo M., Chou C., and Lin K., "The Elimination of Spatial-Temporal Uncertainty in Underwater Sensor Networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1229-1242, 2013. DOI:10.1109/TNET.2012.2220155

[6] Hsu C., Lai K., Chou C., and Lin K., "ST-MAC: Spatial-Temporal MAC Scheduling for Underwater Sensor Networks," *in Proceedings of the IEEE INFOCOM*, Rio De Janeiro, pp. 1827-1835, 2009.

DOI:10.1109/INFCOM.2009.5062103

[7] Jouhari M., Ibrahimi K., Tembine H., and Ben-Othman J., "Underwater Wireless Sensor Networks: A Survey on Enabling Technologies, Localization Protocols, and Internet of Underwater Things," *IEEE Access*, vol. 7, pp. 96879-96899, 2019. DOI:10.1109/ACCESS.2019.2928876

[8] Kredo K., Djukic P., and Mohapatra P., "Stump: Exploiting Position Diversity in the Staggered Time Underwater MAC Protocol," *in Proceedings of the IEEE INFOCOM*, pp. 2961-2965, 2009. DOI: 10.1109/INFCOM.2009.5062267

[9] Matignon L., Laurent G., and Fort-Piat N., "Hysteretic Q-Learning: An Algorithm for Decentralized Reinforcement Learning in Cooperative Multi-Agent Teams," *in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, pp. 64-69, 2007. DOI: 10.1109/IROS.2007.4399095

[10] Noh Y., Wang P., Lee U., Torres D., and Gerla M., "Dots: A Propagation Delay-Aware Opportunistic Mac Protocol for Underwater Sensor Networks," *in Proceedings of the 18th IEEE International Conference on Network Protocols*, Kyoto, pp. 183-192, 2010. DOI:10.1109/ICNP.2010.5762767

[11] Park S., Mitchell P., and Grace D., "Reinforcement Learning Based MAC Protocol (UW-ALOHA-Q) for Underwater Acoustic Sensor Networks," *IEEE Access*, vol. 7, pp. 165531-165542, 2019. DOI:10.1109/ACCESS.2019.2953801

[12] Peleato B. and Stojanovic M., "Distance Aware Collision Avoidance Pro-Tool for Ad-Hoc Underwater Acoustic Sensor Networks," *IEEE Communications Letters*, vol. 11, no. 12, pp. 1025-1027, 2007. DOI:10.1109/LCOMM.2007.071160

[13] Ramakrishnan D. and Radhakrishnan K., "Applying Deep Convolutional Neural Network (DCNN) Algorithm in the Cloud Autonomous Vehicles Traffic," *The International Arab Journal of Information Technology*, vol. 19, no. 2, pp. 186-194, 2022. https://doi.org/10.34028/iajit/19/2/5

[14] Sendra S., Lloret J., Jimenez J., and Parra L., "Underwater Acoustic Modems," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4063-4071, 2016. DOI: 10.1109/JSEN.2015.2434890

[15] Tomovic S. and Radusinovic I., "DR-ALOHA-Q: A Q-Learning-Based Adaptive MAC Protocol for Underwater Acoustic Sensor Networks," *Sensors*, vol. 23, no. 9, pp. 1-22, 2023. https://doi.org/10.3390/s23094474

[16] Vieira L., Kong J., Lee U., and Gerla M., "Analysis of Aloha Protocols for Underwater Acoustic Sensor Networks," https://www.semanticscholar.org/paper/Analysis-of-Aloha-Protocols-for-Underwater-Acoustic-VieiraKong/095862fcbbbde81ba0cbb51831276e8c0e70accbJ, Last Visited, 2024.

[17] Yackoski J. and Shen C., "UW-Flash: Achieving High Channel Utilization in A Time-Based Acoustic Mac Protocol," *in Proceedings of the 3rd International Workshop on Underwater Networks*, pp. 59-66, 2008. DOI: 10.1145/1410107.1410119

[18] Zhuo X., Qu F., Yang H., Wei Y., Wu Y., and Li J., "Delay and Queue Aware Adaptive Scheduling-Based MAC Protocol for Underwater Acoustic Sensor Networks," *IEEE Access*, vol. 7, pp. 56263-56275, 2019. DOI:10.1109/ACCESS.2019.2913209

Isam Ud Din is a member, in IEEE Access received an M.S. Degree in Computer Science from the University of Haripur, Pakistan, and a BS. Degree in Computer Science from IM Sciences University Peshawar. With a focus on the Internet of Medical Things (IoMT), Machine Learning, Natural Language Processing, and Data Science. He is an Editor and a Reviewer of many well-reputed peer-reviewed international journals and conferences. He is working on various projects, including Robotic Devices and a 3D Map-Based Navigation System, and worked as a Senior Researcher at DMark Islamabad.

Irshad Abbasi is a Senior Member, IEEE received an M.S. Degree in Computer Science from COMSATS University Islamabad, Abbottabad Campus, Pakistan, in 2011 and a Ph.D. Degree in Computer Science from Universiti Malaysia Sarawak, Malaysia in 2019. He was a Senior Lecturer at King Khalid University, Saudi Arabia, from 2011 to 2015. He completed his Postdoctoral in Cyber Security at Universiti Malaysia Sarawak. He is also an Assistant Professor at the Department of Computer Science, University of Bisha, Saudi Arabia. He has over 14 years of research and teaching experience. He is the author of many articles published in top-quality journals. His research Interests Include Networks, VANETs, MANETs, FANETs, Mobile Computing, the IoT, Cloud Computing, Cybersecurity, Cryptography, Soft Computing, Drone Security and Authentication. He has received multiple awards, scholarships, and research grants. He is serving as an Editor. He is also acting as a reviewer of many well-reputed peer-reviewed international journals and conferences.

Sikandar Ali received a double Ph.D. Degree from the China University of Petroleum, Beijing, in 2019 and 2021, respectively. He is currently an Assistant Professor. He has authored over 70 articles in highly cited journals and conferences. His research interests include Business Process Management, Anomaly detection in Multi-Sensor Systems, Software Outsourcing Partnerships, Software Testing and Test Automation, Bug Prediction, Bug Fixing, Software Incidence Classification, Source Code Transformation, Agile Software Development, And Global Software Engineering. He served as a Technical Program Committee Member for more than 20 conferences and acted as a Reviewer for many well-reputed journals. He also Organizes many special issues. He is an Academic Editor of Mobile Information Systems, Journal of Computer Information Systems, Intelligent Control and System Engineering, Journal of Intelligent Medicine and Healthcare, and Scientific Programming journals.

Khalil Al Ruqeishi is an Assistant Professor and head of the Computer Section in the DMPS Department, at Nizwa University. He obtained his Ph.D. in Computer Science from Aston University, UK in 2016. He received his MSc in Computer Science from Teesside University, UK in 2007. Earlier, he worked as a part-time lecturer at Busraimi University College where he is a member of the Board of Trustees. His current research interests include Cloud Computing, Big Data, and Artificial Intelligence.