

Article

Cross-Layer Routing Protocol Based on Channel Quality for Underwater Acoustic Communication Networks

Jinghua He ^{1,2,3} , Jie Tian ^{1,3,*}, Zhanqing Pu ^{1,3}, Wei Wang ^{1,3} and Haining Huang ^{1,3,*}

¹ Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Key Laboratory of Science and Technology on Advanced Underwater Acoustic Signal Processing, Chinese Academy of Sciences, Beijing 100190, China

* Correspondence: tianjie@mail.ioa.ac.cn (J.T.); hhn@mail.ioa.ac.cn (H.H.)

Abstract: Due to the physical characteristics of acoustic channels, the performance of underwater acoustic communication networks (UACNs) is more susceptible to the impacts of multipath and Doppler effects. Channel quality can serve as a measure of the reliability of underwater communication links. A cross-layer routing protocol based on channel quality (CLCQ) is proposed to improve the overall network performance and resource utilization. First, the BELLHOP ray model is used to calculate the channel impulse response combined with the winter sound speed profile data of a specific sea area. Then, the channel impulse response is integrated into the communication system to evaluate the channel quality between nodes based on the bit error rate (BER). Finally, during the selection of the next hop node, a reinforcement learning algorithm is employed to facilitate cross-layer interaction within the protocol stack. The optimal relay node is determined by the channel quality index (BER) from the physical layer, the buffer state from the data link layer, and the node residual energy. To enhance the algorithm's convergence speed, a forwarding candidate set selection method is proposed which takes into account node depth, residual energy, and buffer state. Simulation results show that the packet delivery rate (PDR) of the CLCQ is significantly higher than that of Q-Learning-Based Energy-Efficient and Lifetime-Extended Adaptive Routing (QELAR) and Geographic and Opportunistic Routing (GEDAR).

Keywords: cross-layer routing protocol; channel quality; Q-learning; underwater acoustic communication networks



Citation: He, J.; Tian, J.; Pu, Z.; Wang, W.; Huang, H. Cross-Layer Routing Protocol Based on Channel Quality for Underwater Acoustic Communication Networks. *Appl. Sci.* **2024**, *14*, 9778. <https://doi.org/10.3390/app14219778>

Academic Editor: Alessandro Lo Schiavo

Received: 20 September 2024

Revised: 22 October 2024

Accepted: 23 October 2024

Published: 25 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

UACNs employ sound waves to transmit information and facilitate data exchange between underwater devices via underwater acoustic channels. UACNs play a pivotal role in numerous fields, including ocean observation, underwater operations, military applications, and business [1,2]. UACNs typically comprise three main components: fixed underwater acoustic sensor nodes, mobile autonomous underwater vehicles (AUVs), and water gateway nodes. A distributed, multi-hop, three-dimensional network is formed by the aggregation of multiple nodes at different locations. The nodes exchange data via underwater acoustic communication and relay information to the central node through the underwater relay nodes.

Although there are conceptual similarities between UACNs and terrestrial wireless electromagnetic wave networks, underwater acoustic channels are considerably more complex than terrestrial wireless channels. This complexity arises from several factors, including multipath effects, Doppler effects, time-varying characteristics, and limited available bandwidth. These factors result in a high BER, high susceptibility to link interruptions, and increased packet loss in communication transmissions [3–6]. Consequently, UACNs cannot directly adopt the protocols and techniques of terrestrial wireless electromagnetic wave protocols and technologies. When designing routing protocols for UACNs, it is

essential to fully consider the characteristics of the underwater acoustic channels and the underwater environment in order to establish a reliable transmission mechanism that achieves a high PDR in the multi-hop network. In UACNs, factors such as data surges and node failures can lead to network congestion, resulting in packet loss and reception interference, as well as increased queuing delay. This further increases the end-to-end delay and reduces the PDR [7]. Furthermore, sensor nodes in UACNs typically operate on independently sourced power, which is limited and challenging to replenish. Consequently, the energy consumption must be taken into account when designing routing protocols for UACNs.

To address the aforementioned challenges, researchers have proposed cross-layer design approaches, as evidenced by the following references: [8–10]. In contrast to traditional network protocols, which are designed independently within their respective layers, cross-layer protocols facilitate information sharing between different layers, enabling the network to dynamically adjust routing decisions and resource allocation based on real-time channel conditions, node states, and application requirements. This optimization enhances overall network performance [11]. In [12–14], the authors emphasize the importance of considering node energy consumption, transmission path selection, and routing strategy optimization in routing protocol design to improve energy utilization efficiency. However, the impact of node congestion on energy consumption, end-to-end delay, and PDR is not investigated. References [15,16] address the dual issues of energy efficiency and congestion in protocol design; however, insufficient attention is paid to the critical impact of channel quality on PDR. In [17], the receiver signal-to-noise ratio (SNR) is employed to forecast the channel quality, which incorporates the propagation characteristics of signals. However, the physical layer effects during orthogonal frequency-division multiplexing (OFDM) modulation and demodulation are not fully considered.

To increase the PDR in complex marine environments, a cross-layer routing protocol, the CLCQ, for UACNs is proposed that combines channel quality and node congestion state. The design of this protocol fully accounts for the specific characteristics of the marine environment. The winter sound velocity profile data from a specific sea area are employed to calculate the channel impulse response and propagation loss using the BELLHOP ray model. These real marine channel characteristics are then introduced into the OFDM communication system. Subsequently, the signal traversing the OFDM communication system is used to compute the BER, which serves to assess the quality of the channel. In the context of a variable underwater environment, the CLCQ designs a reward function that considers channel quality, node buffer state, and residual energy. The reinforcement learning method enables the CLCQ to adjust the next hop node selection dynamically. Simulation results indicate that the CLCQ exhibits a higher PDR in the context of complex underwater acoustic channels. The main contributions made by this paper are the following:

1. The cross-layer routing protocol is designed by combining the physical layer and data link layer parameters to realize the information exchange between different layers. Simulation results show that the protocol can effectively improve the network performance.
2. The channel impulse response is calculated using winter sound speed profile data from a specific sea area, and the results are applied to the OFDM communication system to obtain the BER of underwater acoustic channels under specific marine environments, different transceiver positions, and fixed modulation modes. The BER serves as a channel quality evaluation index, which provides an important basis for designing cross-layer routing protocols.
3. The reward function for reinforcement learning was designed by considering channel quality, node buffer state, and remaining energy to select reliable links and avoid congestion, thereby enhancing the PDR and reducing end-to-end delay. Additionally, a forwarding candidate selection method based on node depth, remaining energy, and buffer state was proposed to accelerate algorithm convergence.

The rest of the paper is organized as follows: Section 2 reviews related research. The system model, including the network model, energy model, and communication model,

is presented in Section 3. The proposed CLCQ protocol is described in detail in Section 4. Section 5 evaluates the performance of the CLCQ protocol through simulation. Section 6 summarizes the research results and concludes.

2. Related Work

In UACNs, underwater routing protocols can select appropriate paths for data transmission between underwater nodes by considering the characteristics of the underwater acoustic channel and the network topology [18]. Routing protocols are primarily classified into two categories, non-cross-layer routing and cross-layer routing which depends on the involvement of parameters from other layers of the network stack [19]. These parameters include node buffer states, receiver SNR, transmission power, and transmission rates.

Traditional non-cross-layer routing implements specific standards and rules only within network layers, with no interaction between layers. One classic example is the Vector-Based Forwarding (VBF) routing protocol, which relies on three-dimensional coordinate information for data transmission [20]. VBF uses source–destination vectors as axes to directionally forward packets within a limited-radius communication pipe, in contrast to the flooding approach. However, in a network with sparsely distributed nodes, the forwarding node may not be able to find the next hop node within the communication range of the pipe. To solve this problem, researchers have also proposed two improved routing protocols based on VBF: the HHVBF (Hop-by-Hop VBF) routing protocol [21] and VBVA (Vector-Based Void Avoidance) routing protocol [22]. HHVBF employs a hop-by-hop forwarding strategy that allows each node to maintain a virtual routing pipeline pointing to the destination node. VBVA combines vector shifting and backpressure mechanisms to avoid convex and concave voids in the network. QELAR [23] employs Q-learning for the first time in distributed routing protocols in underwater sensor networks. It integrates the residual energy of each node and the average residual energy of neighboring nodes into the design of the reward function to select an optimally energy-efficient transmission path. Experiments show that the lifetime of QELAR is on average 20% longer than that of VBF. However, the design of the protocol focuses only on the energy of the nodes and does not consider other factors that may affect the routing.

The objective of cross-layer routing protocols is to enhance the overall performance of routing protocols through the implementation of effective information-sharing mechanisms while maintaining the hierarchical structure of the network stack [24]. GEDAR [25] is an opportunistic routing protocol that employs the concept of clustering, which focuses on determining the next hop forwarding node by calculating the priority of the derived cluster. The priority calculation incorporates both the distance between nodes and the packet error rate. The sending node then forwards the packet to the cluster of nodes with the highest priority. However, the protocol does not fully consider the congestion state of the nodes during packet forwarding, which may result in higher end-to-end delay and energy consumption. In [15], a congestion avoidance routing protocol, RCAR, is proposed. The reward function takes into account the congestion level and the residual energy of the nodes, allowing for the avoidance of congested regions during packet forwarding. Furthermore, RCAR introduces a MAC layer-based handshake mechanism for updating information, which ensures that the optimal transmission path is selected. Although RCAR is a cross-layer routing protocol, it primarily concentrates on the design of the network layer and data link layer, with minimal consideration of how certain factors in the physical layer impact the protocol's performance.

MOR [26] is a cross-layer routing protocol proposed by Yuan et al. The protocol considers four key performance indicators, energy consumption, end-to-end delay, link quality, and link congestion, and uses the non-dominated sorting genetic algorithm II (NSGA-II) to find the optimal routing path. The effectiveness of MOR in reducing energy consumption, decreasing end-to-end delay, and improving packet transmission rate is verified by experiments. PCAQR [13] is a new network topology that realizes cross-layer application by optimizing the transmission power and routing path of nodes, reducing

energy consumption and delay. CLORP [27] is a cross-layer opportunistic routing protocol that uses multi-agent reinforcement learning with two reward functions for successful and failed transmissions, combined with cross-layer information to optimize routing. It employs an adaptive learning rate and a Q-value initialization strategy based on location and neighbor count to accelerate convergence and adapt to dynamic network topology changes. CLIC utilizes an integrated routing MAC to adaptively avoid nodes likely to experience high conflict or congestion during routing [28]. This approach provides a high PDR and low latency with low overhead. Nevertheless, the protocol does not fully account for the distinctive characteristics of underwater acoustic channels, such as signal propagation loss and multipath effects, which can influence the efficacy of data transmission. GO-MAC [17] employs a novel approach to determining the next hop node, integrating geographic routing protocols and OFDM techniques. This integration enables the simultaneous optimization of communication resources and the selection of next hop nodes through a handshaking mechanism. The degree of node congestion and channel quality are taken into account to optimize the adaptive backoff algorithm in the geographic routing protocol. Nevertheless, the assessment of the channel quality in this protocol is predominantly based on the ray propagation model to predict the receiver SNR, which primarily focuses on the propagation characteristics of the signals and does not fully consider the physical layer effects in the OFDM modulation and demodulation process.

In conclusion, when designing cross-layer routing protocols, several factors must be taken into account in order to ensure the efficiency and reliability of the protocols. Firstly, the protocol design should fully consider the acoustic propagation characteristics, particularly in special environments such as underwater acoustic channels, where signal propagation loss and multipath effects significantly impact data transmission performance. Secondly, cross-layer design factors are also critical, including but not limited to congestion state, residual energy, channel quality, and hop count, which collectively determine the merits of routing and the overall performance of the network. Furthermore, it is essential to consider the physical layer characteristics, particularly when employing modulation techniques such as OFDM. The unique physical layer characteristics of OFDM must be fully incorporated into the protocol architecture.

3. System Model

3.1. Network Model

In this paper, the information is generated by the underwater source node and transmitted upward to the destination node using relay forwarding. Subsequently, the destination node transmits the data packets to the ground base station for data analysis via electromagnetic waves. The network model is depicted in Figure 1. We make the following assumptions:

1. Numerous underwater sensor nodes are randomly scattered throughout a three-dimensional underwater network;
2. The destination nodes are energy unconstrained and can obtain their position information through GPS;
3. All underwater nodes can obtain their location information through positioning algorithms;
4. All nodes have access to their buffer status.

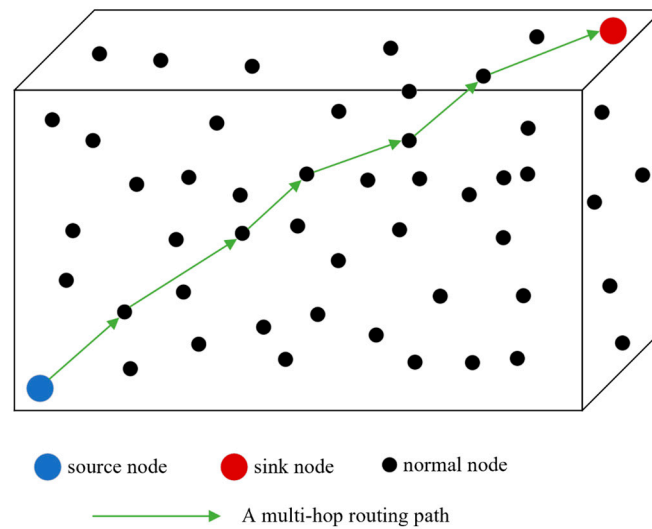


Figure 1. The schematic diagram of the network.

3.2. Energy Consumption Model

In underwater acoustic communication, the attenuation function within a channel, as defined by the Thorp propagation model, is given by Equation (1), where f represents the frequency in kHz, and d is the distance in km.

$$A(d, f) = d^k \alpha(f)^d \quad (1)$$

Here k and $\alpha(f)$ denote the energy scattering factor and the absorption coefficient, respectively. The energy scattering factor is assigned values based on the scattering type: 1 for columnar scattering, 1.5 for real scattering, and 2 for spherical scattering. The absorption coefficient $\alpha(f)$ is determined by Equation (2) and is measured in dB/km.

$$\alpha(f) = 10^{\alpha/10} \quad (2)$$

$$\alpha = \frac{0.11f^2}{1+f^2} + \frac{44f^2}{4100+f^2} + 2.75 \times 10^{-4}f^2 + 0.003 \quad (3)$$

When a node transmits m bits of data over a distance d , the model for node transmission energy consumption [29] can be expressed as follows:

$$E_S = P_R A(d, f) T_S \quad (4)$$

The energy consumption by the receiver to receive k bits of data can be calculated by the following:

$$E_R = P_R \times T_R \quad (5)$$

where P_R is the received power, and T_R is the time taken to receive the m bits of data.

3.3. Communication Model

In order to consider the impact of channel quality on network performance, the real propagation loss and channel impulse response are calculated using the BELLHOP ray model in conjunction with winter sound speed profile data from a specific sea area. The obtained channel impulse response is then applied to a physical-layer OFDM communication system to evaluate the BER of the underwater acoustic channel in conjunction with specific modulation modes and underwater environmental conditions.

The winter sound speed profile, as depicted in Figure 2, is utilized to describe the sound speed distribution of the deep-sea acoustic channel in a specific sea area. We obtain the number of multipaths N_{pa} for the known transmitting depth and the known receiving

position, as well as the amplitude and delay information in the $Npa \times 1$ dimension. The impulse response $h(t)$ is defined as follows:

$$h(t) = A_0\delta(t - \tau_0) + \sum_{u=1}^{Npa-1} A_u\delta(t - \tau_u) \quad (6)$$

where A_u and τ_u are the amplitude and delay of the u -th acoustic line at the receiver, respectively. An example of the channel is shown in Figure 3. If the signal at the transmitter is denoted by $x_m(t)$, the received signal of the multipath channel is denoted by the following:

$$y_m(t) = A_0x_m(t - \tau_0) + \sum_{u=1}^{Npa-1} A_u x_m(t - \tau_u) + n_c(t) \quad (7)$$

where $y_m(t)$ is the superposition of the direct acoustic signal and multiple reflections, and $n_c(t)$ is the channel noise.

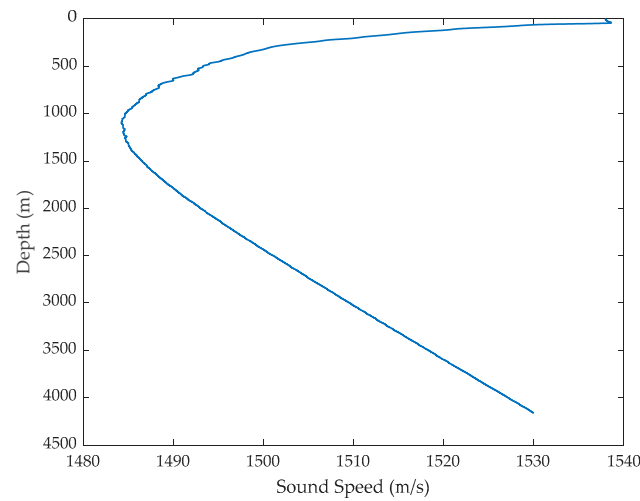


Figure 2. The sound speed profile.

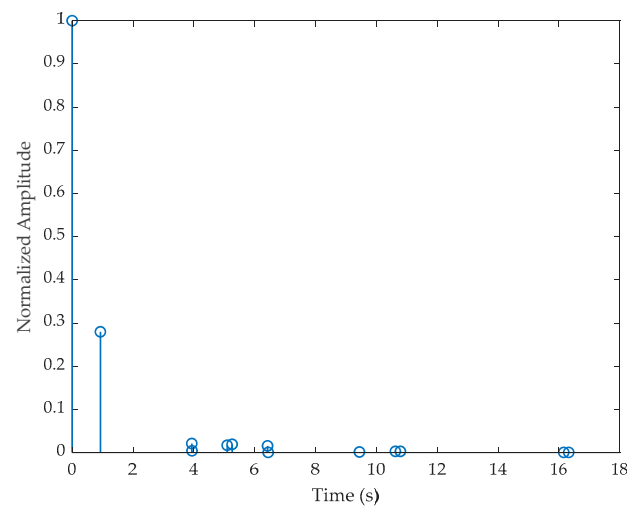


Figure 3. Example of channel (transmit depth 1000 m, receive depth 878.1 m, distance 536.8 m).

The delay and amplitude of the multipath channel, as calculated by BELLHOP, are incorporated into the physical layer of the OFDM system depicted in Figure 4. OFDM is a modulation technique that divides a high-speed data stream into multiple low-speed subcarriers for transmission. This process has the characteristics of resistance to multipath

fading, high-frequency utilization, and strong resistance to jamming. It is therefore suitable for the complex channel environment of underwater acoustic communication.

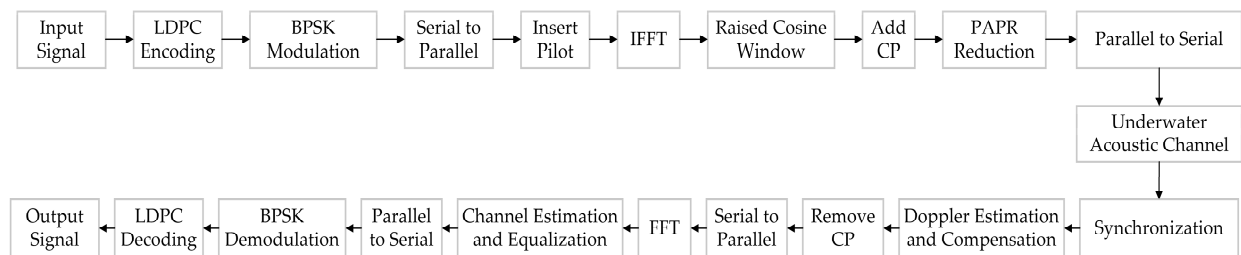


Figure 4. OFDM underwater acoustic communication system implementation process.

The OFDM system described in this paper employs LDPC code for channel coding, BPSK for modulation and demodulation, and a frequency-based channel estimation and equalization method. In order to eliminate the Doppler effect in the underwater acoustic channel, a time-domain correlation method is employed for Doppler estimation, while an interpolation method is used for Doppler compensation. Based on the positional information of the transmitting and receiving nodes, the BER for known transmitting and receiving depths and distance can be calculated. The bit error rate is inversely proportional to the channel quality. A lower bit error rate indicates a higher channel quality. Figure 5 depicts the 2D BER for a transmitting depth of 800 m, varying receiving depths, and varying receiving distances. As stated in [25], the PDR can be calculated from the BER.

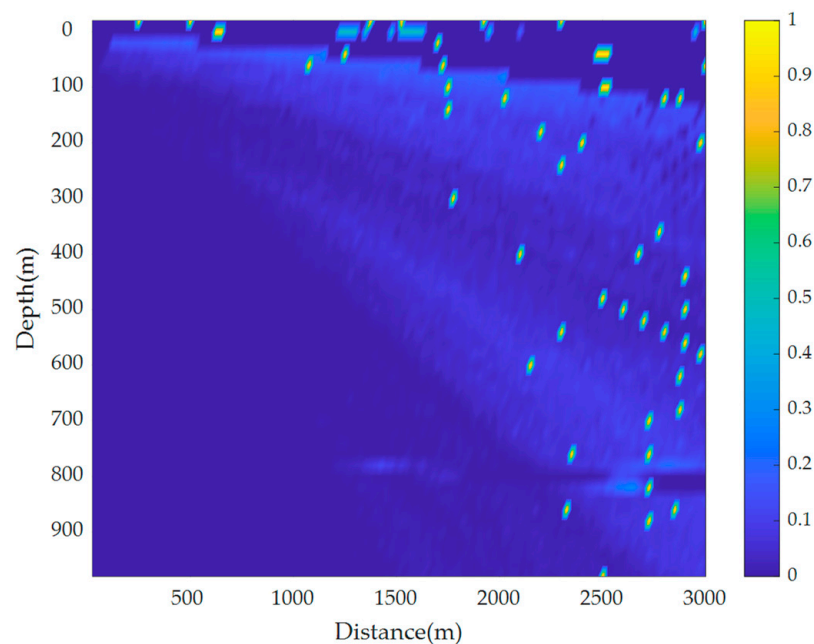


Figure 5. BER at different receiving depths and distances (transmitting depth 800 m).

4. Proposed CLCQ Protocol

4.1. Q-Learning-Based Routing Protocol

Q-learning is a model-free reinforcement learning algorithm which learns the optimal actions by interacting with the environment and estimating the long-term value of actions in specific states [30]. The objective of Q-learning is to identify the optimal policy that maximizes the expected future rewards. The complexity of the underwater acoustic channel and the uncertainty of the environment make it challenging to obtain a large number of real environmental data. Q-learning is capable of modifying the node behavior strategy in real time based on the current network status and data characteristics. This makes it a

highly versatile solution for UACNs that frequently change. The fundamental principle underlying Q-learning is the utilization of the Bellman equation to iteratively update the Q-function until it converges to the optimal Q-value. The updated formula for the Q-function can be expressed as follows:

$$Q^\pi(s_i, a_i) = r_i + \gamma \sum_{s_j \in S} P_{s_i s_j}^{a_j} Q^\pi(s_j, a) \quad (8)$$

where s_i and s_j represent the current state and the next state, respectively. a_i represents the action taken, and r_i represents the immediate reward for the current action. $\gamma \in (0, 1)$ denotes the discount factor, and $P_{s_i s_j}^{a_i}$ is the transition probability from state s_i to state s_j through action a_i . Ref. [31] has proven that there exists at least one optimal policy π^* that can achieve the optimal value, which can be described as follows:

$$V^*(s) = \max_a (Q^*(s, a)) \quad (9)$$

$$Q^*(s_i, a_i) = r_i + \gamma \sum_{s_j \in S} P_{s_i s_j}^{a_i} V^*(s_j) \quad (10)$$

Among these, $Q^*(s_i, a_i)$ represents the expected return of taking action a_i under the optimal policy in the state s_i . The protocol framework that is the subject of this study is depicted in Figure 6. In order to calculate the value of the reward function, nodes must obtain information from the environment, such as the remaining energy of neighboring nodes. The agent then evaluates the actions for candidate nodes and sends the data packet to the next hop probabilistically. This method fully exploits the information transfer between the network layer, data link layer, and physical layer within the network protocol stack to achieve efficient routing in practical applications.

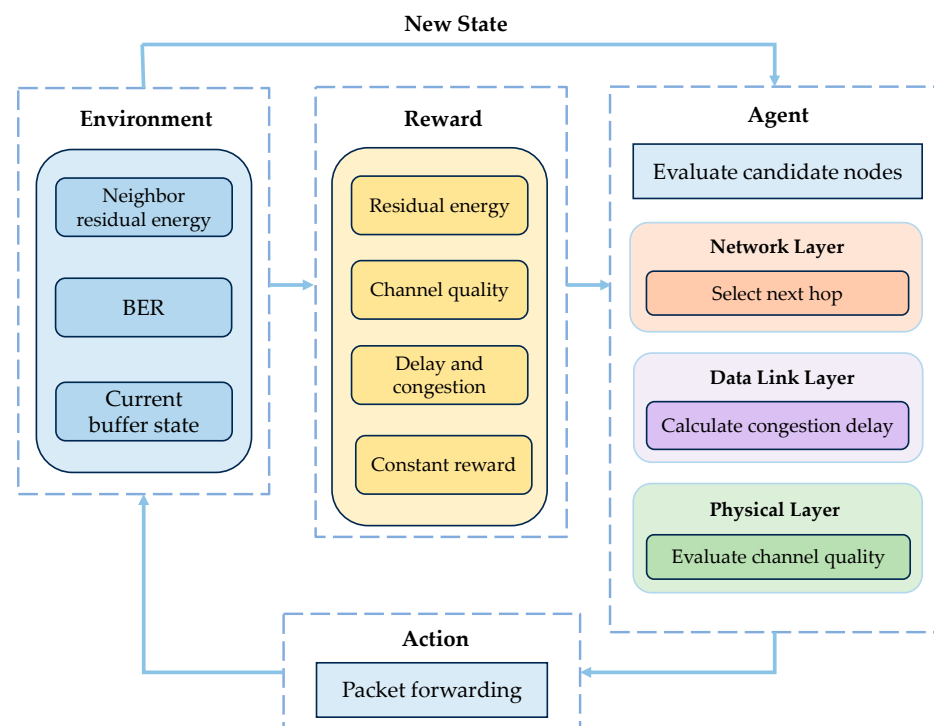


Figure 6. Protocol framework.

4.2. Selection of Forwarding Candidate Set

This section will examine the process of selecting the forwarding candidate set in detail. By comprehensively considering factors such as node depth, energy, and buffer

state, we will further filter out eligible forwarding nodes from the neighboring nodes. Only nodes that meet the specified criteria can be selected as the next hop.

In UACNs, the neighbor set of node n_i is defined as

$$Neighbor_i = \{n_{i1}, n_{i2}, n_{i3}, \dots, n_{im}\} \quad (11)$$

where m represents the number of neighbor nodes of node n_i . If node n_j satisfies Equation (12), then n_j belongs to the forwarding candidate set of n_i .

$$candidateSet = Subset_1 \cap Subset_2 \cap Subset_3 \quad (12)$$

$$Subset_1 = \{n_j \subseteq Neighbor_i \mid depth(n_j) \leq depth(n_i)\} \quad (13)$$

$$Subset_2 = \{n_j \subseteq Neighbor_i \mid E_{res}(n_j) \geq 0.2E_{init}(n_j)\} \quad (14)$$

$$Subset_3 = \{n_j \subseteq Neighbor_i \mid Buffer_{now}^j \leq 0.8Buffer_{max}^j\} \quad (15)$$

To ensure that packets are transmitted towards a decreasing depth, nodes with a depth less than or equal to that of node n_i are first selected from the neighbor nodes according to Equation (13), forming a subset $Subset_1$. Furthermore, in order to prevent certain nodes from depleting their energy prematurely and creating routing holes, neighbor nodes with less than 80% of the initial energy are filtered out according to Equation (14), which helps extend the lifespan of the network. Concurrently, to alleviate network congestion and enhance packet transmission efficacy, only nodes whose current buffer length is not more than 80% of the maximum buffer length are selected according to Equation (15). Finally, the intersection of the aforementioned subsets is selected as the final forwarding candidate set. However, if the forwarding candidate set is empty, this article proposes that the forwarding candidate set be assigned all neighbor nodes of the current node n_i , denoted as $Neighbor_i$. This approach ensures uninterrupted packet transmission.

4.3. Selection of Next Hop

The design of this protocol incorporates considerations of the channel quality of the actual underwater environment, thereby improving the PDR. As a cross-layer routing protocol based on Q-learning, the CLCQ can select the next hop forwarding node by calculating function values through the current buffer state, the residual energy, and the BER of point-to-point communication. Note that competitive MAC protocols can potentially lead to collisions, which may result in packet loss. In order to mitigate this risk, the CLCQ employs Slotted FAMA [32] as the MAC protocol and makes further improvements. To prevent packet collisions, each packet must be transmitted within the corresponding time slot. The length of a time slot is defined as follows:

$$slot = T_{prop} + T_{trans} + T_{guard} \quad (16)$$

$$T_{prop} = \frac{range}{V_{sound}} \quad (17)$$

$$T_{trans} = \frac{packetSize}{Rb} \quad (18)$$

where T_{prop} is the maximum propagation delay, T_{trans} is the transmission delay, and T_{guard} is the guard interval, set to 0.001 s.

The goal of Q-learning is to learn a policy that can achieve the maximum cumulative reward by continuously updating the Q-value. The design of the reward function is crucial for the effectiveness of Q-learning as it determines the behavioral goals and learning motivations of the agent. Assuming that node n_i holds a packet under state s_i , the action of node n_i sending the packet to node n_j in its forwarding candidate set is denoted as a_{ij} , with the obtained reward denoted as R_{ij} , and the Q-function defined as $Q(s_i, a_{ij})$. Some symbols are shown in Table 1.

Table 1. The list of symbols.

Parameters	Symbols
Constant reward	g_0
Channel quality sensitivity	φ_q
Channel-quality-related reward	$c(q)$
Delay and congestion sensitivity	φ_t
Delay-and-congestion-related reward	$c(t)$
Residual energy sensitivity	φ_{en1}
Residual energy reward	$b(n_i), b(n_j)$
Energy distribution sensitivity	φ_{en2}
Energy distribution reward	$d(n_i), d(n_j)$
Residual energy of node n_i	$E_{res}(n_i)$
The initial energy of node n_i	$E_{init}(n_i)$
Average residual energy of n_i neighbor nodes	$\bar{E}(n_i)$
Current buffer length	$Buf_{f_{now}}$
Maximum buffer length	$Buf_{f_{max}}$
A handshake time	t_{MAC}
Proportionality factor	α_1, α_2
BER from n_i to n_j	BER_{ij}
Average BER to its neighbor nodes	\overline{BER}
Maximum number of retransmissions	N_{max}
Discount factor	γ

If the packet is successfully transmitted from node n_i to node n_j , the reward function can be obtained by the following:

$$R_{s_{ij}}^{a_{ij}} = -g_0 - \varphi_q \times c(q) - \varphi_t \times c(t) - \varphi_{e1} \times [b(n_i) + b(n_j)] + \varphi_{e2} \times [d(n_i) + d(n_j)] \quad (19)$$

The reward function mainly consists of four parts: a fixed reward, a channel-quality-related reward, a delay-and-congestion-related reward, and an energy-consumption-related reward. Among them, g_0 is the fixed reward with a weight set to 1, while the other weights (φ_q , φ_t , φ_{en1} , and φ_{en2}) are all less than or equal to 1.

Energy-consumption-related reward: $b(n_i)$ is the reward related to the remaining energy, defined as follows [23]:

$$b(n_i) = 1 - \frac{E_{res}(n_i)}{E_{init}(n_i)} \quad (20)$$

wherein $E_{res}(n_i)$ and $E_{init}(n_i)$, respectively, represent the remaining energy and initial energy of the node. When all nodes have the same initial energy, the lower the remaining energy of node n_i , the higher the reward $b(n_i)$, and the more difficult it is for the two nodes to communicate. $d(n_i)$ is a function measuring the energy distribution of the node, defined as follows:

$$d(n_i) = \frac{2}{\pi} \arctan(E_{res}(n_i) - \bar{E}(n_i)) \quad (21)$$

where $\bar{E}(n_i)$ is the average remaining energy of all neighbor nodes of node n_i . The greater the discrepancy between a node's residual energy and the mean residual energy of its neighboring nodes, the more likely it is that the node will be selected as the next hop node.

Delay-and-congestion-related reward: Congestion occurs when there are a large number of packets in the buffer of a node. The handshake-based Slotted FAMA protocol transmits a limited number of packets per handshake. For nodes with a high load, a certain number of handshakes is required to release the load. Consequently, the congestion issue in the network can be viewed as a delay, and the reward can be expressed as follows [15]:

$$c(t) = 1 - \frac{1}{t_{total} + 1} \quad (22)$$

$$t_{total} = \alpha_1 \times \left(\left\lceil \frac{Buff_{now}}{Buff_{max}} \right\rceil + 1 \right) \times t_{MAC} + \alpha_2 \times t_{MAC} \quad (23)$$

where t_{MAC} is the total delay of data transmission in the Slotted FAMA protocol, and $\alpha_1 \in [0, 1]$ and $\alpha_2 \in [0, 1]$ are two delay coefficients, respectively. The candidate node with a larger buffer length is less likely to be selected as the next hop node.

Channel-quality-related reward: The channel quality and connection reliability between network nodes are factors that cannot be ignored in the system networking. Here, the BER is used to measure the reliability of the communication link. We propose an improved reward related to channel quality as follows:

$$c(q) = \sigma^{BER_{ij}} - 1 \quad (24)$$

where σ is a predetermined base, and, in order to ensure that the range of $c(q)$ is $[0, 1)$, σ is set to 2. A greater BER results in a greater channel-quality-related cost, which makes it less likely that the node will be chosen as the next hop.

If the number of retransmissions of the data packet reaches the maximum limit and the next hop has not received the data packet, the data packet forwarding process is unsuccessful. The reward function for forwarding failure is presented by the following:

$$R_{Fij}^{a_{ij}} = -g_0 - \varphi_q \times c'(q) - \varphi_t \times c'(t) - \varphi_{e1} \times b(n_i) + \varphi_{e2} \times d(n_i) \quad (25)$$

$$c'(q) = \sigma^{\overline{BER}} - 1 \quad (26)$$

$$c'(t) = 1 - \frac{1}{t'_{total} + 1} \quad (27)$$

$$t'_{total} = \alpha_1 \times \left(\left\lceil \frac{Buff_{now}}{Buff_{max}} \right\rceil + 1 \right) \times t_{MAC} + \alpha_2 \times t_{MAC} \times N_{max} \quad (28)$$

According to [15], N_{max} is the maximum number of retransmissions, and $c'(t)$ is the delay-and-congestion-related overhead of transmission failure. \overline{BER} is the average bit error rate from node n_i to all neighbor nodes, and $c'(q)$ is the channel-quality-related reward of transmission failure.

The dynamic nature of the underwater acoustic channel makes it challenging to ascertain the likelihood of successful data packet transmission. Consequently, this paper employs the bit error rate of point-to-point communication to delineate the channel state. The state transition success probability and packet loss rate are expressed as follows:

$$P_{s_i s_j}^{a_{ij}} = 1 - BER_{ij} \quad (29)$$

$$P_{s_i s_i}^{a_{ij}} = BER_{ij} \quad (30)$$

Accordingly, we define the direct reward function for packet transmission:

$$R_{ij} = P_{s_i s_j}^{a_{ij}} \times R_{s_{ij}}^{a_{ij}} + P_{s_i s_i}^{a_{ij}} \times R_{Fij}^{a_{ij}} = (1 - BER_{ij}) \times R_{s_{ij}}^{a_{ij}} + BER_{ij} \times R_{Fij}^{a_{ij}} \quad (31)$$

The action utility function is defined as follows:

$$Q^*(s_i, a_i) = R_{ij} + \gamma \times ((1 - BER_{ij}) \times V^*(s_j) + BER_{ij} \times V^*(s_i)) \quad (32)$$

The optimal value is updated as follows:

$$V^*(s_i) = \max_a (Q^*(s_i, a)) \quad (33)$$

Initially, the Q-values and V-values of all nodes are set to 0, the constant reward g_0 is set to 1, and the discount factor γ is set to 0.5.

4.4. Packet Structure Design

The packet structure design in this protocol is shown in Figure 7.

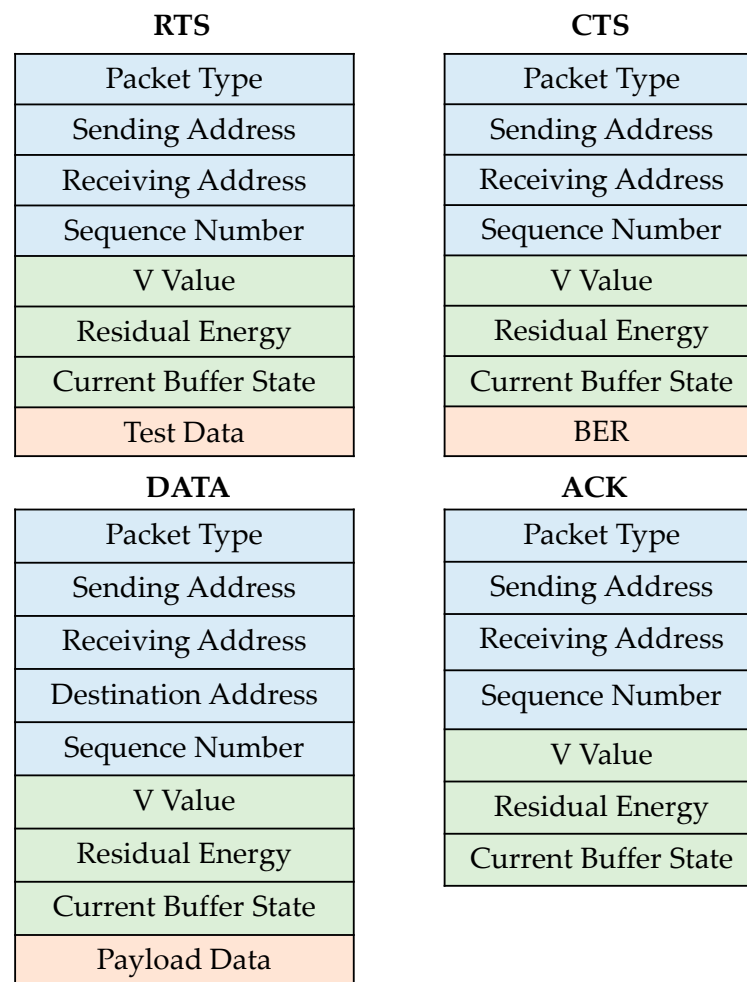


Figure 7. Packet structure.

RTS: This packet is used by sending nodes to request channel access from their neighbors and carries a test data segment for calculating the BER. After receiving the RTS packet, the neighbor node calculates the BER through the test data segment and writes the result to the corresponding field of the CTS packet.

CTS: Neighbors send CTS packets to the sending node simultaneously at the beginning of the slot to confirm that they have received the RTS packet. To avoid multiple CTS packets colliding at the sending node, each neighbor node uses different subcarrier frequency bands to transmit CTS packets.

DATA: In addition to the payload to be transmitted, it also contains routing information such as the destination address and receiving address, as well as additional information such as the V-value and residual energy.

ACK: After the receiving node successfully receives the DATA packet, it sends an ACK packet to confirm receipt of the data and provides feedback about its own status information to the sending node.

4.5. Overview of CLCQ

The process of the CLCQ protocol is shown in Algorithm 1. The main steps of the CLCQ algorithm are outlined below:

- (1) Initialize the network: Set parameters such as node coordinates, Q-values, V-values, communication range, maximum buffer length, and initial energy.
- (2) Broadcast beacon: Obtain the status information of neighbor nodes, such as the residual energy, current buffer state, and locations.
- (3) Send RTS: Broadcast RTS packet with a test data segment. After receiving the RTS packet, the neighbor node uses the test data segment to calculate the BER and writes the result to the CTS packet.
- (4) Receive CTS packets from neighbor nodes.
- (5) Determine the forwarding candidate set: Select based on node depth, remaining energy, and buffer state.
- (6) Calculate the Q-values of all nodes in the forwarding candidate set.
- (7) Select the node with the highest Q-value as the next hop, update the V-value, and then proceed with the data packet transmission.
- (8) Determine if the sink node has received the data packet: If the sink node has received the data packet, the process ends; if not, repeat steps (3) to (7) until the sink node receives it.

Algorithm 1 CLCQ Algorithm

```

1: Initialize network;
2: Broadcast beacon;
3: Get  $E_{res}$ ,  $Buff_{now}$ ,  $Buff_{max}$ , and the location of neighbors;
4: Begin
5:   If (node  $n_i \neq$  sink node) then
6:     Send RTS;
7:     Receive CTS;
8:     Select forwarding candidate set;
9:     While the next hop is not found do
10:      For  $n_j \in$  candidateSet do
11:        Calculate the reward function  $R_{ij}$ ;
12:        Calculate the action-utility function  $Q^*(s_i, a_i)$ ;
13:      End for
14:      Select  $n_j$  with the max  $Q^*(s_i, a_i)$  as the next hop;
15:      Update the  $V(s_i)$  of  $n_i$  with max  $Q^*(s_i, a_i)$ ;
16:    End while
17:    Packet transmission;
18:    Receive ACK;
19:  End if
20: End
  
```

5. Simulation Results and Performance Analyses

This section simulates the CLCQ in three aspects. Firstly, the convergence speed is compared with that of QELAR. Subsequently, the impact of varying parameters on the CLCQ is assessed. Finally, the four performance indicators of energy consumption, residual energy variance, average end-to-end delay, and PDR are compared with QELAR and GEDAR.

5.1. Simulation Setting

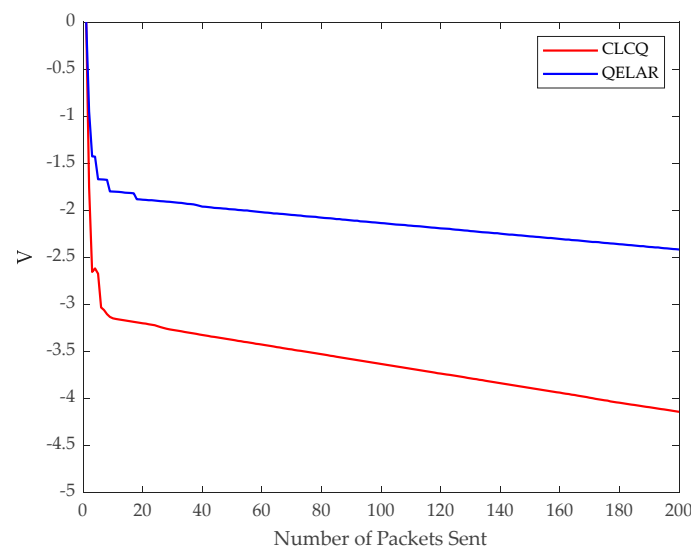
The simulation model is a network structure including one source node, one destination node, and multiple relay nodes. The three-dimensional coverage water area is 3000 m * 3000 m * 1000 m, and the relay nodes are randomly deployed. The packet is generated by the source node and sends it at a time. Other parameters are shown in Table 2.

Table 2. The simulation parameters.

Simulation Parameters	Values
Transmission range	1000 m
Frequency	9.75 kHz
Receiving power	0.6 W
Idle power	1 mW
Transmission rate	2 kb/s
Data packet size	512 bit
Energy initialization of nodes	1000 J
The number of nodes	[50,60,70,80,90,100]

5.2. Performance Evaluation of CLCQ

The V-value of the source node is employed to assess the depth of its comprehension of the network topology and the accuracy of its prediction of the optimal global path. When the V-value of the source node is observed to be stable or to exhibit a convergence trend, it can be inferred that the node has learned a relatively accurate optimal path selection strategy, resulting in a stable state of transmission efficiency and performance for the routing protocol within the network. Figure 8 illustrates the evolution of the V-value of the source node following the transmission of each packet in a network comprising 60 nodes. A comparison of the V-value change observed in this protocol with that observed in the QELAR protocol reveals several significant differences.

**Figure 8.** V-value of the source node.

With the continuous transmission of data packets, the V-value experiences a sharp decline and then gradually slows down, which suggests that the protocol progressively identifies the optimal route based on the current environmental conditions. In contrast, the V-value of QELAR slows down when the 20th packet is sent, while the speed of this protocol begins to slow down when the 6th packet is sent. This suggests that the protocol converges more quickly than QELAR. This rapid convergence is attributed to the optimization strategy employed by this protocol before data packet transmission. The forwarding candidate set is determined based on node depth, residual energy, and buffer state. By filtering the set of node neighbors, unnecessary exploration is reduced, thereby accelerating the convergence speed of the algorithm. This strategy not only enhances the efficacy of routing selection but also optimizes the performance of the entire network, thereby conferring upon the protocol enhanced adaptability and stability in a dynamic network environment.

Figures 9 and 10 illustrate the impact of varying φ_{en1} and φ_{en2} on the total energy consumption and residual energy variance when 60 nodes are deployed in the network.

Firstly, it can be observed from Figure 9 that, as the value of φ_{en2} increases, the total energy consumption displays an upward trend. Furthermore, it can be seen that, as the value of φ_{en1} increases, the energy consumption also increases, although the overall change is not significant. As the value of φ_{en2} increases, the network is more likely to select a node with a significant discrepancy between its residual energy and the average residual energy of its neighbors as the next hop node. This choice increases the transmission path length, thereby raising the overall energy consumption. This heightened energy consumption indicates that the network may be compromising some energy efficiency in pursuit of a balanced energy distribution.

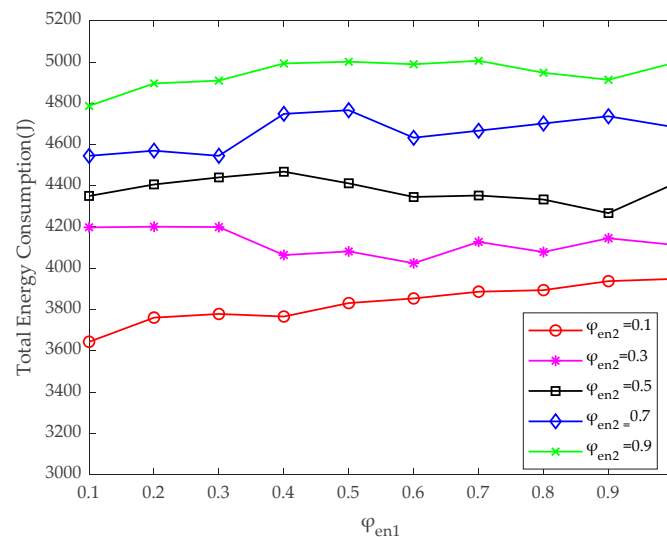


Figure 9. Total energy consumption under different φ_{en1} and φ_{en2} values.

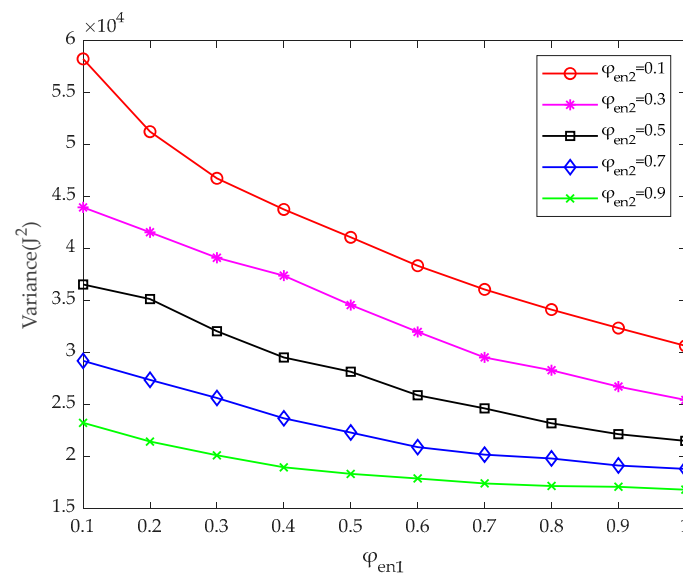


Figure 10. Residual energy variance under different φ_{en1} and φ_{en2} values.

Figure 10 illustrates the trend of residual energy variance. As the values of φ_{en1} and φ_{en2} increase, the residual energy variance of the network gradually decreases. For instance, when $\varphi_{en1} = 1$ and $\varphi_{en2} = 0.9$, the residual energy variance is only 28.87% of that when $\varphi_{en1} = 0.1$ and $\varphi_{en2} = 0.1$. This is because the proportion of energy consumption in the reward function increases, rendering the residual energy of the node a crucial factor in routing selection. When φ_{en2} increases, the network will select a node with a significant discrepancy between the residual energy of its neighbors and its residual energy. This

approach helps to achieve a balanced distribution of energy, thereby reducing the residual energy variance. On the other hand, when φ_{en1} increases, the network is inclined to select nodes with a considerable residual energy, which also serves to diminish the residual energy variance. Consequently, the residual energy variance of nodes gradually tends to become balanced over time following multiple packet transmissions.

In conclusion, by modifying the values of φ_{en1} and φ_{en2} , a trade-off between network energy consumption and residual energy balance can be achieved. While increasing these two parameters can reduce the variance of residual energy and promote the balanced distribution of energy, it will also increase the total energy consumption of the network. Consequently, in practical applications, it is necessary to select reasonable values for these two parameters based on the specific needs and objectives of the network to achieve optimal network performance.

Figure 11 illustrates the impact of varying φ_t , α_1 , and α_2 on the average end-to-end delay when 100 nodes are deployed in the network. Under the configuration of $\varphi_t = 1$, $\alpha_1 = 0.9$, $\alpha_2 = 0.1$, the average end-to-end delay is 76.78 s. In contrast, when $\varphi_t = 0.1$, $\alpha_1 = 0.1$, and $\alpha_2 = 0.9$, the average end-to-end delay is 85.78 s, which is approximately 9 s shorter than that of the latter. The results illustrate that, by adjusting the values of φ_t , α_1 , and α_2 , the network delay can be effectively reduced and the network performance can be optimized. Firstly, an increase in the value of φ_t makes the protocol more inclined to select those nodes with a lower latency when choosing the next hop node, and consciously avoid the congestion area. This is because the time required for multiple retransmissions caused by congestion is considerably longer than the time required for a successful single transmission, indicating that congestion has a more significant impact on the network. Furthermore, while maintaining φ_t constantly, increasing α_1 can more effectively avoid severely congested neighboring nodes. This strategy facilitates a reduction in the number of packets that must traverse the congested area, thereby reducing the average end-to-end delay of the entire network.

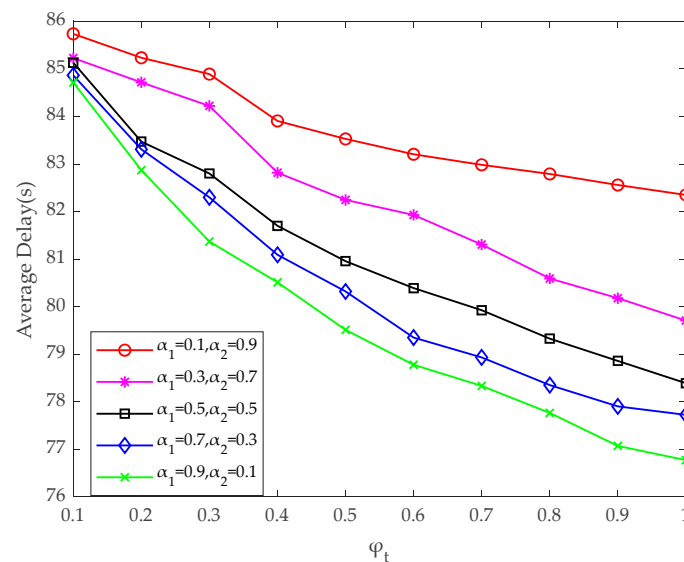


Figure 11. Average end-to-end delay under different φ_t , α_1 , and α_2 values.

Figure 12 illustrates the impact of varying φ_q and φ_t values on PDR when 60 nodes are deployed in the network. Firstly, an increase in φ_q means that the protocol is more inclined to select nodes with a lower bit error rate as the next hop. A reduction in the bit error rate will lead to an improvement in link quality and reliability, resulting in fewer errors and retransmissions during data transmission. This significantly enhances the PDR. On the other hand, increasing φ_t tends to cause the protocol to avoid nodes that lead to multiple retransmissions due to congestion, which also increases the PDR. Although congestion management is equally important for improving the PDR, the impact of channel

quality on the PDR seems to be more significant, as illustrated in Figure 12. When $\varphi_t = 0.1$, as φ_q increases, the PDR increases more rapidly than when $\varphi_t = 0.9$. This indicates that channel quality is a more significant factor in this protocol, and its impact on PDR is more pronounced than that of congestion. Therefore, it is necessary to consider both the channel quality of physical layer and the congestion degree of the link layer in cross-layer design. Such a design can optimize data transmission in different network environments, significantly improving communication reliability and packet delivery rates.

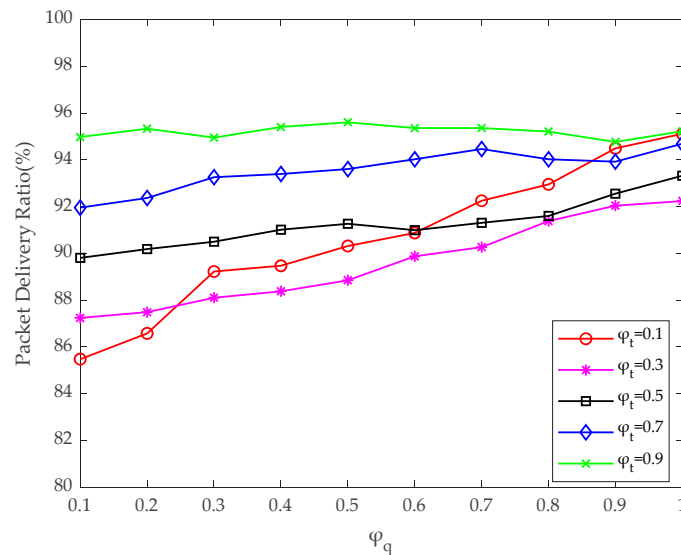


Figure 12. PDR under different φ_q and φ_t values.

5.3. Performance Comparisons

Figure 13 illustrates the total energy consumption of the proposed CLCQ, as well as QELAR and GEDAR, under varying numbers of nodes. It is evident from the data that, as the number of network nodes increases, the total energy consumption of this protocol and QELAR nodes exhibits a downward trend, while the energy consumption of GEDAR shows a slight increase but remains relatively consistent. Once the number of nodes exceeds 70, the energy consumption advantage of this protocol becomes noticeable. This phenomenon reveals that, as the network size increases, the advantages of this protocol in energy consumption control become more pronounced. In particular, when there are 90 nodes in the network and both φ_{en1} and φ_{en2} are set at 0.1, the energy consumption of this protocol is approximately 16.04% and 17.93% lower than that of QELAR and GEDAR, respectively. This result not only corroborates the efficacy of this protocol in terms of energy conservation but also highlights its performance advantages, especially in large-scale networks. This is because the protocol is designed to consider how to reduce energy consumption by optimizing routing as the network scale expands.

As illustrated in Figure 14, the residual energy variance of the CLCQ, QELAR, and GEDAR varies with the number of nodes. The selection mechanism of relay nodes in the GEDAR protocol is primarily dependent on the location information. Consequently, when node locations do not change significantly over a short period, the network tends to follow a fixed transmission path. This results in GEDAR having significantly higher residual energy variance compared to the other two protocols. In contrast, the proposed protocol and QELAR show superior performance in addressing the residual energy variance. As the node count grows, the number of next hop nodes available for selection also increases, thereby providing more routing choices for the protocol. This, in turn, helps to achieve more uniform energy consumption. Consequently, the residual energy variance of the two protocols gradually decreases as the number of nodes increases. Furthermore, as the number of nodes increases, the CLCQ presents a more pronounced advantage in energy balance and achieves superior energy utilization efficiency. This indicates that

the protocol proposed in this paper is capable of more intelligent energy consumption allocation within the network, preventing excessive energy depletion on specific nodes and thereby enhancing the overall energy utilization efficiency of the network.

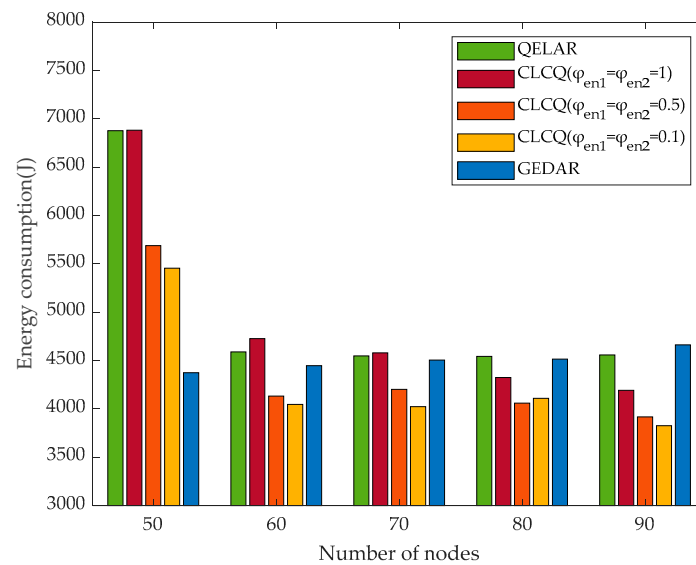


Figure 13. Total energy consumption of different protocols.

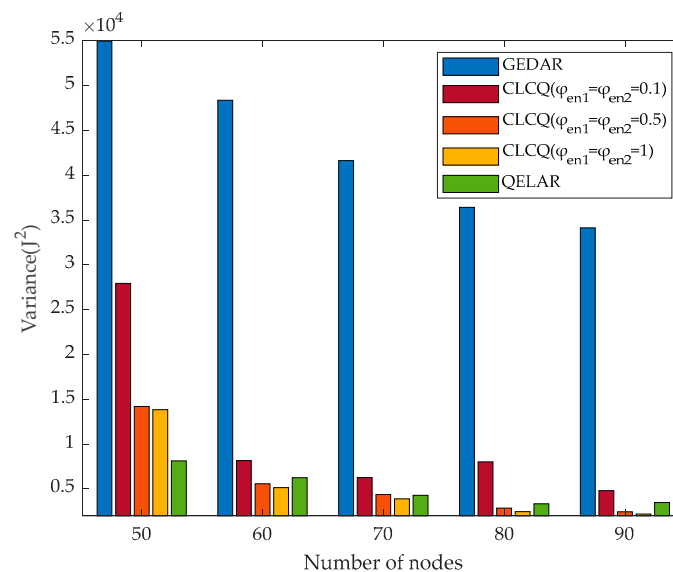


Figure 14. Residual energy variance of different protocols.

Figure 15 depicts the average end-to-end delay of the CLCQ under varying numbers of nodes, with comparisons made to QELAR and GEDAR. It can be observed that the average end-to-end delay of all protocols decreases as the number of nodes increases. This is because an increase in node density allows the network to identify a path with fewer hops between the source and destination nodes, thereby reducing delay. Furthermore, GEDAR does not consider the issue of congestion, resulting in a significantly longer end-to-end delay than the other two protocols. In the case of a small number of nodes, the average end-to-end delay of QELAR is observed to be shorter than that of the CLCQ. Nevertheless, as the number of nodes increases, the delay gap between the CLCQ and the QELAR protocol begins to narrow. Once the network reaches a certain scale, the CLCQ exhibits a lower average end-to-end delay than QELAR. This indicates that, as the network scale expands, the CLCQ exhibits enhanced adaptability and superiority in addressing congestion, selecting routing, and optimizing data transmission paths. The advantage of the

CLCQ is that it fully considers the channel quality of the physical layer and the congestion of the data link layer in the protocol design. By dynamically evaluating the BER of links and the buffer state of nodes, the CLCQ can effectively avoid congested areas and select more reliable links, thereby reducing the number of retransmissions in data transmission and the delay caused by congestion. Therefore, the design of the CLCQ improves packet transmission efficiency while reducing end-to-end delay, which validates its contribution to improving communication performance in dynamic network environments.

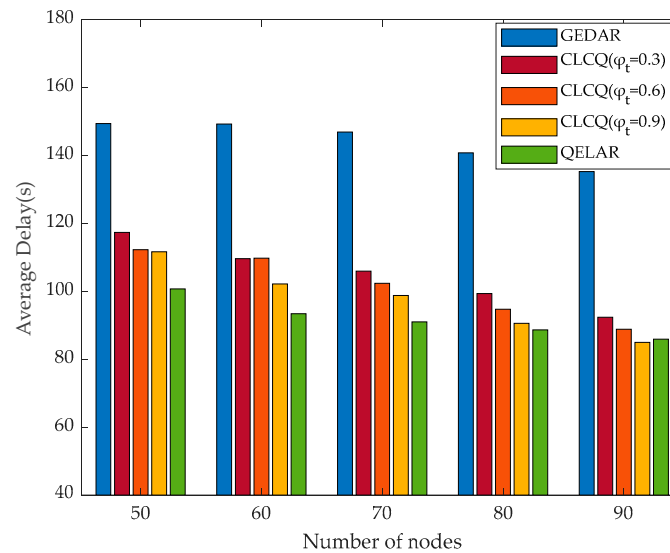


Figure 15. Average end-to-end delay of different protocols.

Figure 16 illustrates the PDR of each routing protocol under varying node counts. As the number of nodes increases, the PDR of all protocols improves. For GEDAR and QELAR, an increase in node count means that more nodes participate in the routing process, which, in turn, improves the PDR. Nevertheless, the protocol proposed in this paper exhibits a more pronounced advantage in terms of PDR. This advantage is not solely attributable to the increased route selection opportunities afforded by the increase in the number of nodes. It also benefits from the cross-layer mechanism of protocol design. In the event of a high-quality channel between nodes, there is a greater probability of the nodes being selected as the next hop, thereby reducing the occurrence of transmission failures due to unreliable channels. At the same time, the CLCQ effectively avoids congested links and further improves the PDR. For instance, at a node count of 90, the PDR of QELAR is 78.71%, while that of GEDAR is 86.93%. When $\phi_q = 0.9$, the PDR of the CLCQ is 97.02%. Because GEDAR has shortcomings in network congestion management, QELAR's assessment of channel quality is not comprehensive enough, and the CLCQ is better at solving these problems.

Figures 13–16 provide a comprehensive evaluation of the four key indicators of the CLCQ, with a comparison to QELAR and GEDAR. In terms of energy consumption, once the network node count surpasses 60, the CLCQ's energy consumption advantage becomes apparent, especially in large-scale networks. Although the CLCQ initially shows a marginally higher average end-to-end delay compared to QELAR, this gap narrows as the number of nodes increases, as illustrated in Figure 15. At a certain point, the CLCQ reaches the lowest average end-to-end delay. Additionally, the CLCQ presents outstanding results with respect to the crucial metric of the PDR, with a significantly higher success rate than that of QELAR and GEDAR. Overall, the protocol proposed in this paper is the most balanced of the three protocols.

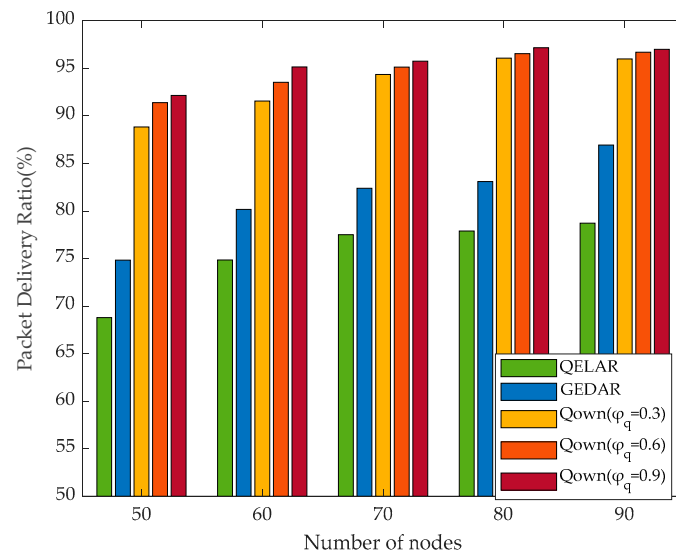


Figure 16. PDR of different protocols.

5.4. Summary and Discussion

The simulation results in Sections 5.2 and 5.3 show the effectiveness of the CLCQ design. First, the forwarding candidate set selection method proposed in Section 4.2, which comprehensively considers node depth, residual energy, and buffer state, reduces the exploration of invalid nodes and significantly speeds up the convergence rate of the algorithm. As shown in Figure 8, the convergence rate of the CLCQ is notably faster than that of QELAR, proving the contribution of this method to improving the overall efficiency of the protocol.

Second, Figures 9–12 highlight the necessity of the reward function design. The reward function in Section 4.3 accounts for not only channel quality (BER) but also node residual energy and buffer state. As shown in Figures 9 and 10, despite the increase in energy consumption, the residual energy distribution is more balanced, indicating the effectiveness of energy-related rewards. Figures 11 and 12 further illustrate that the reward function based on channel quality and congestion effectively reduces end-to-end latency and improves the PDR.

Finally, the comparison of the CLCQ with QELAR and GEDAR (Figures 13–16) shows that the CLCQ has clear advantages in energy consumption control, delay reduction, and delivery rate improvement in large-scale networks. This is primarily due to the cross-layer design presented in Section 4.1, which enhances transmission reliability and efficiency by dynamically evaluating channel quality and buffer state.

6. Conclusions

The objective of this study was to propose a novel cross-layer routing protocol, designated as the CLCQ. The actual channel impulse response is calculated by using the winter sound velocity profile data from a specific sea area. Subsequently, the underwater acoustic channel is incorporated into the OFDM communication system, and the channel quality is evaluated by calculating the bit error rate, which is then applied to the protocol design. Furthermore, this paper proposes an optimized forwarding candidate set selection method that accelerates the convergence speed of the reinforcement learning algorithm, thereby further enhancing the overall performance and efficiency of the routing protocol.

The CLCQ effectively integrates the information from the physical layer and data link layer to intelligently select the next hop node. To guarantee the reliability of data transmission, this study devised a reward function based on a reinforcement learning algorithm. This function considers various factors, including channel quality, buffer state, and residual energy, in order to select the optimal route for forwarding data packets. Consequently, the CLCQ circumvents congestion areas while selecting high-reliability links.

The results of the simulation present that the PDR of the CLCQ is markedly enhanced compared to the traditional underwater acoustic routing protocols.

However, some unresolved issues remain in this study. For instance, the BER utilized in the protocol design is calculated under the deterministic communication system, without considering different communication environments.

Author Contributions: J.T. provided academic guidance and offered practical suggestions for the research content. J.H. designed the core method proposed in this paper, developed the program, conducted the relevant simulation verification, and drafted the manuscript. Z.P. revised the manuscript, enhancing both its content and structure. W.W. provided theoretical guidance for the paper. H.H. guided the overall direction of the research and provided a research framework for the study. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ali, M.F.; Jayakody, D.N.K.; Chursin, Y.A.; Affes, S.; Dmitry, S. Recent Advances and Future Directions on Underwater Wireless Communications. *Arch. Comput. Methods Eng.* **2020**, *27*, 1379–1412. [\[CrossRef\]](#)
2. Xiao, X.; Huang, H.; Wang, W. Underwater Wireless Sensor Networks: An Energy-Efficient Clustering Routing Protocol Based on Data Fusion and Genetic Algorithms. *Appl. Sci.* **2021**, *11*, 312. [\[CrossRef\]](#)
3. Akyildiz, I.F.; Pompili, D.; Melodia, T. Underwater Acoustic Sensor Networks: Research Challenges. *Ad Hoc Netw.* **2005**, *3*, 257–279. [\[CrossRef\]](#)
4. Akkaya, K.; Younis, M. A Survey on Routing Protocols for Wireless Sensor Networks. *Ad Hoc Netw.* **2005**, *3*, 325–349. [\[CrossRef\]](#)
5. Luo, Y.; Pu, L.; Zuba, M.; Peng, Z.; Cui, J.-H. Challenges and Opportunities of Underwater Cognitive Acoustic Networks. *IEEE Trans. Emerg. Top. Comput.* **2014**, *2*, 198–211. [\[CrossRef\]](#)
6. Awan, K.M.; Shah, P.A.; Iqbal, K.; Gillani, S.; Ahmad, W.; Nam, Y. Underwater Wireless Sensor Networks: A Review of Recent Issues and Challenges. *Wirel. Commun. Mob. Comput.* **2019**, *2019*, e6470359. [\[CrossRef\]](#)
7. Ghaffari, A. Congestion Control Mechanisms in Wireless Sensor Networks: A Survey. *J. Netw. Comput. Appl.* **2015**, *52*, 101–115. [\[CrossRef\]](#)
8. Emokpae, L.E.; Liu, Z.; Edelmann, G.F.; Younis, M. A Cross-Stack QoS Routing Approach For Underwater Acoustic Sensor Networks. In Proceedings of the 2018 Fourth Underwater Communications and Networking Conference (UComms), Lercici, Italy, 28–30 August 2018; pp. 1–5.
9. Liu, J.; Yu, M.; Wang, X.; Liu, Y.; Wei, X.; Cui, J. RECRP: An Underwater Reliable Energy-Efficient Cross-Layer Routing Protocol. *Sensors* **2018**, *18*, 4148. [\[CrossRef\]](#)
10. Su, Y.; Xu, Y.; Pang, Z.; Kang, Y.; Fan, R. HCAR: A Hybrid-Coding-Aware Routing Protocol for Underwater Acoustic Sensor Networks. *IEEE Internet Things J.* **2023**, *10*, 10790–10801. [\[CrossRef\]](#)
11. Pompili, D.; Akyildiz, I.F. A Multimedia Cross-Layer Protocol for Underwater Acoustic Sensor Networks. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 2924–2933. [\[CrossRef\]](#)
12. Wang, B.; Zhang, H.; Zhu, Y.; Cai, B.; Guo, X. Adaptive Power-Controlled Depth-Based Routing Protocol for Underwater Wireless Sensor Networks. *J. Mar. Sci. Eng.* **2023**, *11*, 1567. [\[CrossRef\]](#)
13. Shen, Z.; Yin, H.; Jing, L.; Ji, X.; Liang, Y.; Wang, J. A Power Control-Aided Q-Learning-Based Routing Protocol for Optical-Acoustic Hybrid Underwater Sensor Networks. *IEEE Trans. Green Commun. Netw.* **2023**, *7*, 2117–2129. [\[CrossRef\]](#)
14. Xu, H.; Yuan, X. Cross-Layer Design for Energy-Efficient Reliable Multi-Path Transmission in Event-Driven Wireless Sensor Networks. *Sensors* **2023**, *23*, 6520. [\[CrossRef\]](#)
15. Jin, Z.; Zhao, Q.; Su, Y. RCAR: A Reinforcement-Learning-Based Routing Protocol for Congestion-Avoided Underwater Acoustic Sensor Networks. *IEEE Sens. J.* **2019**, *19*, 10881–10891. [\[CrossRef\]](#)
16. Ali, R.; Sohail, M.; Almagrabi, A.O.; Musaddiq, A.; Kim, B.-S. greenMAC Protocol: A Q-Learning-Based Mechanism to Enhance Channel Reliability for WLAN Energy Savings. *Electronics* **2020**, *9*, 1720. [\[CrossRef\]](#)
17. Guo, J.; Song, S.; Liu, J.; Chen, H.; Lin, B.; Cui, J.-H. An Efficient Geo-Routing-Aware MAC Protocol Based on OFDM for Underwater Acoustic Networks. *IEEE Internet Things J.* **2023**, *10*, 9809–9822. [\[CrossRef\]](#)
18. Li, N.; Martinez, J.-F.; Meneses Chaus, J.M.; Eckert, M. A Survey on Underwater Acoustic Sensor Network Routing Protocols. *Sensors* **2016**, *16*, 414. [\[CrossRef\]](#)

19. Shovon, I.I.; Shin, S. Survey on Multi-Path Routing Protocols of Underwater Wireless Sensor Networks: Advancement and Applications. *Electronics* **2022**, *11*, 3467. [\[CrossRef\]](#)
20. Xie, P.; Cui, J.-H.; Lao, L. VBF: Vector-Based Forwarding Protocol for Underwater Sensor Networks. In *Proceedings of the NETWORKING 2006. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*; Boavida, F., Plagemann, T., Stiller, B., Westphal, C., Monteiro, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1216–1221.
21. Nicolaou, N.; See, A.; Xie, P.; Cui, J.-H.; Maggiorini, D. Improving the Robustness of Location-Based Routing for Underwater Sensor Networks. In *Proceedings of the OCEANS 2007–Europe, Aberdeen, Scotland, 18–21 June 2007*; pp. 1–6.
22. Xie, P.; Zhou, Z.; Peng, Z.; Cui, J.-H.; Shi, Z. Void Avoidance in Three-Dimensional Mobile Underwater Sensor Networks. In *Wireless Algorithms, Systems, and Applications*; Liu, B., Bestavros, A., Du, D.-Z., Wang, J., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5682, pp. 305–314; ISBN 978-3-642-03416-9.
23. Hu, T.; Fei, Y. QELAR: A Machine-Learning-Based Adaptive Routing Protocol for Energy-Efficient and Lifetime-Extended Underwater Sensor Networks. *IEEE Trans. Mob. Comput.* **2010**, *9*, 796–809. [\[CrossRef\]](#)
24. Luo, J.; Chen, Y.; Wu, M.; Yang, Y. A Survey of Routing Protocols for Underwater Wireless Sensor Networks. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 137–160. [\[CrossRef\]](#)
25. Coutinho, R.W.L.; Boukerche, A.; Vieira, L.F.M.; Loureiro, A.A.F. GEDAR: Geographic and Opportunistic Routing for Underwater Sensor Networks. *IEEE Trans. Comput.* **2016**, *65*, 548–561. [\[CrossRef\]](#)
26. Yuan, Y.; Zhuo, X.; Wei, Y.; Qu, F. Multi-Objective Routing Optimization to Support IoUT Applications in UWSNs. *IEEE Internet Things J.* **2024**, *11*, 27664–27675. [\[CrossRef\]](#)
27. Liu, S.; Wang, J.; Shi, W.; Han, G.; Yan, S.; Li, J. CLORP: Cross-Layer Opportunistic Routing Protocol for Underwater Sensor Networks Based on Multiagent Reinforcement Learning. *IEEE Sens. J.* **2024**, *24*, 17243–17258. [\[CrossRef\]](#)
28. Sun, Y.; Ge, W.; Li, Y.; Yin, J. Cross-Layer Protocol Based on Directional Reception in Underwater Acoustic Wireless Sensor Networks. *J. Mar. Sci. Eng.* **2023**, *11*, 666. [\[CrossRef\]](#)
29. Sozer, E.M.; Stojanovic, M.; Proakis, J.G. Underwater Acoustic Networks. *IEEE J. Ocean. Eng.* **2000**, *25*, 72–83. [\[CrossRef\]](#)
30. Watkins, C.J.C.H.; Dayan, P. Q-Learning. *Mach Learn* **1992**, *8*, 279–292. [\[CrossRef\]](#)
31. Naem, M.; Rizvi, S.T.H.; Coronato, A. A Gentle Introduction to Reinforcement Learning and Its Application in Different Fields. *IEEE Access* **2020**, *8*, 209320–209344. [\[CrossRef\]](#)
32. Molins, M.; Stojanovic, M. Slotted FAMA: A MAC Protocol for Underwater Acoustic Networks. In *Proceedings of the OCEANS 2006–Asia Pacific, Singapore, 16–19 May 2006*; pp. 1–7.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.