A Time-Slotted Data Gathering Medium Access Control Protocol Using Q-Learning for Underwater Acoustic Sensor Networks

FAISAL AHMED, (Student Member, IEEE), AND HO-SHIN CHO[ID], (Senior Member, IEEE)
School of Electronic and Electrical Engineering, Kyungpook National University, Daegu 41556, South Korea
Corresponding author: Ho-Shin Cho (hscho@ee.knu.ac.kr)

ABSTRACT Contention-based medium access control (MAC) protocols for underwater acoustic sensor networks are designed to handle packet collisions that are caused by long propagation delays. However, existing protocols are known to suffer from relatively high collisions, which decrease system performance. To enhance system performance, we propose a contention-based MAC protocol that employs a widely-popular machine learning technique, namely, Q-learning. Using Q-learning, the proposed protocol allows the sensor nodes to intelligently select the back-off slots and accordingly schedule the transmission of data packets such that collisions are minimized at the receiver. Unlike in existing protocols, the sensor nodes are not required to exchange scheduling information, which implies that the proposed protocol has low complexity and overhead. Under varying traffic loads and node numbers, the proposed protocol is compared with the state-of-the-art ALOHA-Q for underwater environment (UW-ALOHA-Q), multiple access collision avoidance for underwater (MACA-U) and exponential increase exponential decrease (EIED) protocols. Results demonstrate the effectiveness of the proposed protocol in terms of energy efficiency, channel utilization, and latency.

INDEX TERMS Back-off, collisions, medium access control, machine learning, Q-learning, slot selection, underwater acoustic sensor networks.

I. INTRODUCTION

With recent technological advances, it is now possible to explore and monitor the ocean through the application of underwater acoustic sensor networks (UWASNs) [1]. Generally, UWASNs contain a large number of battery powered acoustic sensors [2] that execute collaborative tasks such as oceanographic environment monitoring, data collection, disaster prevention, and tactical surveillance [3]. Unlike terrestrial wireless sensors, underwater acoustic sensors use acoustic waves to communicate with each other; these are approximately five-fold slower than the radio frequency waves used by terrestrial wireless sensors [4]. Despite being slower, acoustic waves are preferable to radio waves as they are relatively robust to attenuation while radio waves suffer from high attenuation in underwater environments [5]. However, due to the lower propagation speed of the underwater acoustic channel, some distinct features such as limited

bandwidth, low channel capacity, high bit error rate, and high dynamics of channel quality are observed in underwater acoustic sensor networks [3].

Medium access control (MAC) protocols are pertinent for UWASNs as they provide high system performance in an energy efficient manner by coordinating access to a shared medium [6]. The aforementioned characteristics of the underwater channel make direct application of terrestrial MAC protocols in UWASNs an inefficient process; thus, the process must be rethought [7].

Generally, the major task of both underwater and terrestrial MAC protocols is resolving collisions of data packets at the receiver. However, for underwater MAC protocols, resolving data packet collision at the receiver must include consideration of transmission time along with the distance between the senders and the receiver, while terrestrial MAC protocols require consideration of transmission time only as the propagation delay is negligible in the terrestrial domain [8]. Consider a small network with one receiver and two senders in the underwater environment. In this network,

The associate editor coordinating the review of this manuscript and approving it for publication was Miguel López-Benítez[ID].
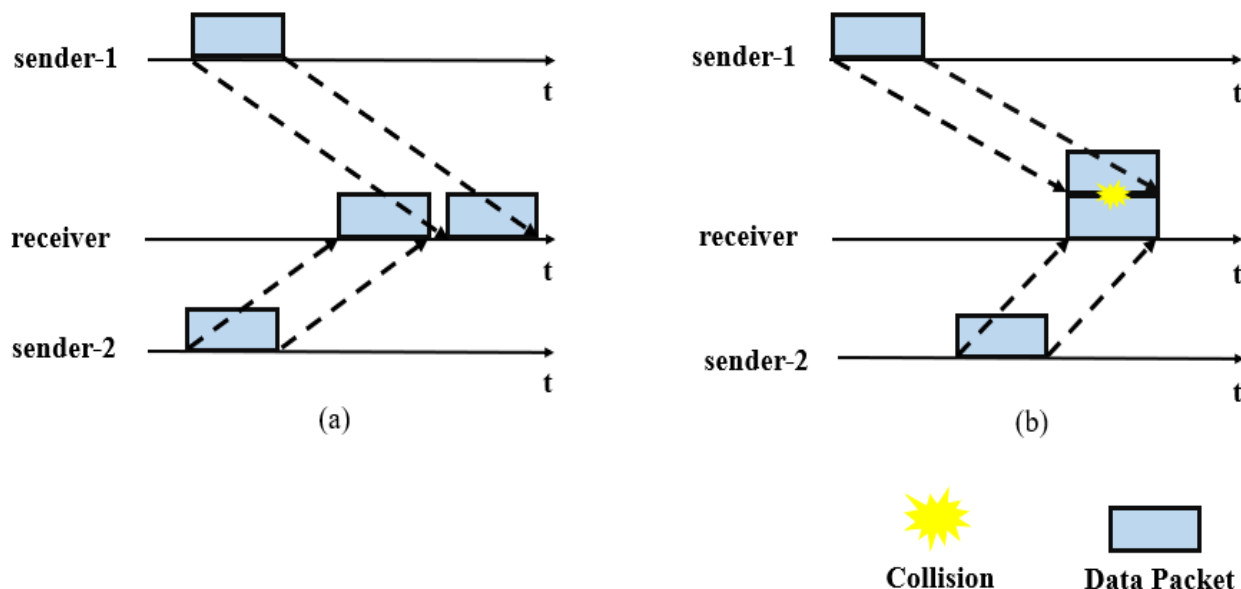
FIGURE 1. Collisions caused by space-time uncertainty.

sender-1 and sender-2 have different distances with reference to the receiver. Suppose that sender-1 and sender-2 want to transmit one data packet to the receiver, and both senders select the same sending time (Fig. 1(a)) for data packet transmission. In this case, even with the same transmission time collision is possibly avoided as both data packets arrive successfully at the receiver at different receiving times. Moreover, assuming, that sender-1 and sender-2 transmit data packets at different sending times and, despite these differences, the reception time for these data packets is possibly the same at the receiver; a collision would then occur (Fig. 1(b)). This phenomenon is a two-dimensional uncertainty, which is better known as space–time uncertainty [9].

Considering the dynamicity and the complexity of underwater acoustic channels, applying artificial intelligence (AI) techniques found to be an effective solution. Hence, various works on UWASNs consider AI techniques to solve a range of problems. For example, a combination of three AI techniques, i.e., ant colony optimization, artificial fish swarm, and dynamic coded operation, is designed to reduce energy consumption and increase robustness [10]. Boosted regression tree technique is used to classify the modulation and coding scheme levels by investigating the underwater channel characteristics [11]. A modulation selection method, which is based on a conventional neural network and random forest technique to ensure a reliable underwater acoustic communication in the time-varying underwater acoustic channel, is proposed in [12]. In [13], [14], and [16], Q-learning is used to determine the optimal routing path, the selection of adaptive underwater channels, and schedule the transmission of data packets, respectively.

In the present study, we focused on reducing the collision probability at the receiver by designing a novel Q-learning-based MAC protocol for UWASNs. Q-learning is a popular model-free reinforcement learning technique that is

based on the *state–action* pair value denoted Q(s,a), where s denotes a *state* and a denotes an *action*. Briefly, Q-learning predicts the Q(s,a) value without any prior knowledge on the parameters of the system of a given environment. The Q(s,a) value for taking an *action* in a particular state is defined by a *policy* π, which determines how an agent behaves at a given time [15]. The agent is considered an entity that learns and makes decisions independently while everything outside the agent is considered the environment [11]. For learning, the agent explores its environment by selecting a particular *action* in each *state* based on the Q(s,a) values. Depending on the *action* performed, the agent receives a positive or negative *reward*. The ultimate goal is to determine the *action* that generates the maximum *reward*.

Q-learning, if applied to underwater acoustic sensor networks, may increase the performance of these networks by reducing collisions at the receiver. Additionally, as Q-learning works on a trial-and-error basis, it does not require signaling information exchange to schedule the transmission of data packets and its algorithm has a relatively low computational complexity.

The main contributions of the current study are as follows.

- An underwater Q-learning-based MAC protocol, which has the major advantage of low-signaling overhead and low complexity, is proposed to optimize the selection of back-off slots while offering low packet collisions.
- A newly designed reward function that is based on the outcome of transmitted data packets, i.e., success or collision. Moreover, for the first time, the normalized received power level, which is not considered in the existing contention-based MAC protocols, is considered to determine the number of collisions.
- Through a system-level simulator, the energy efficiency, latency, and channel utilization of the proposed protocol

under varying traffic loads and the number of nodes is investigated.

The remainder of this paper is organized as follows. Section II provides a discussion on the related works. In Section III, the assumptions and conditions are presented, whereas in Section IV the proposed scheme is introduced. Numerical results are presented in Section V and conclusions are drawn in Section VI.

II. RELATED WORKS

Previous studies have shown that the two general classes of MAC protocol, contention-based and contention-free MAC protocols, attempt to complete a common task, namely, collision minimization at the receiver [16]. Contention-based underwater MAC protocols can use full underwater channel bandwidth [17]; thus, most efforts to design underwater MAC protocols have focused on the contention-based class.

Handshaking and random-access MAC are the two classifications of contention-based MAC protocols. In handshaking-based MAC protocols, sender and receiver nodes exchange small control packets before data packet transmission for the reservation of channels and the avoidance of collisions. In random access-based MAC protocols, collision avoidance is a probabilistic approximation [6] because the senders' data transmission has no prior coordination. Hence, handshaking-based MAC protocols have dominance over random access-based protocols in terms of collision avoidance but they show less efficiency in terms of energy and latency.

Fang *et al.* proposed the carrier sense multiple access with collision avoidance for underwater (UW-CSMA/CA) [18], an asynchronous underwater MAC protocol in which a combination of carrier sensing and handshaking is applied to reduce collision probability. The main principle of the UW-CSMA/CA algorithm is similar to that of the original CSMA/CA algorithm but with enhancements in the state transition rules, transmission deferment rules, and waiting slot time. These three improvements have been introduced into UW-CSMA/CA to harmonize the underwater environment features.

Similar to the UW-CSMA/CA algorithm, the handshaking-based MAC protocol slotted floor acquisition multiple access (slotted FAMA) was proposed by Molins and Stojanovic [19]; this was adapted from the FAMA protocol with the addition of an ARQ technique. Although similar, slotted FAMA differs from UW-CSMA/CA because it is synchronized and channel access is divided into fixed time slots where all the packets (RTS, CTS, DATA, or ACK) must be delivered at the beginning of the time slot. The length of the time slot is the combination of maximum propagation delay, length of RTS/CTS packet, and guard time. The introduction of the time-slotting technique in slotted FAMA minimizes the chance of collisions and eliminates the requirement for long control packets.

Peleato and Stojanovic [20] proposed an improvement to the slotted FAMA protocol that minimizes overheads and

increases energy efficiency. The improved protocol also utilizes the handshaking technique, but it includes the addition of a waiting time period between the time when a CTS packet is received and the time the sender node starts sending its data packet, which avoids potential packet collisions.

Another improvement to the slotted FAMA proposed by Dong *et al.* [21] considered an adaptive slot duration strategy that avoids packet collisions by using the handshaking technique while trying to offer good throughput of MAC in the UWASNs.

Multiple access collision avoidance for underwater MACA-U [22], a collision avoidance underwater MAC protocol, proposed by Ng *et al.* is an adaption of the terrestrial MACA protocol but with changes to the control rules and packet forwarding approach. Although MACA-U is a handshaking-based MAC protocol like the other abovementioned MAC protocols, it differs by not using carrier sensing before sending an RTS, which improves energy efficiency, and not including an ARQ technique, which eliminates extra overheads.

These handshaking-based underwater MAC protocols guarantee collision avoidance but degrade the channel utilization as significant amount of idle time is introduced in the network. Additionally, their exchange of control packets can introduce high energy consumption, latency, and extra overheads into the network.

In [8], [23], and, [24] for UWASNs, random access MAC protocols i.e., different variants of Aloha-based protocols have been proposed, where one of the protocol use slot guard time for the mitigation of packet collisions [8], while others, such as, protocol proposed in [17] add additional control packets for collision avoidance. A receiver synchronized approach is used in the protocol proposed in [18] to minimize packet collisions at the receiver. However, for high traffic conditions or large number of sensor nodes, the random-access protocols suffer from low channel utilization following high latency and energy consumption.

Apart from the above-mentioned approaches, several other procedures, such as back-off algorithms, that can reduce the collision probability at the receiver are applied to contention-based MAC protocols. One of the examples named exponential increase and exponential decrease (EIED) is applied to many existing works [25]–[28]. In EIED, the back-off value is exponentially incremented or decremented based on the status of a packet transmission, i.e., success or collision. The fixed tuning of EIED introduces a considerable amount of idle time at the network when the traffic load is high and for large number of sensor nodes, which affects the network's channel utilization.

Contention-free MAC protocols, that is, TDMA based MAC protocols are well-known for scheduled data packet transmission that can mitigate data packet collisions at the receiver. In [29], [30], and [31], a TDMA based MAC protocol is proposed where collision reduction at the receiver comes at the cost of a large scheduling delay. The reason is that first the central node has to collect the information to

determine the transmission/reception schedule, then it has to send out the schedule information to the sensor node, and finally the sensor nodes, on receiving the schedules, transmit their data packets accordingly.

Owing to low-computational requirements, Q-learning is widely-studied in terrestrial MAC protocols to solve issues such as time-slot scheduling [32], data transmission scheduling [33], and active and sleep time adjustment for duty cycling [34]. At present, very few studies have been conducted on using Q-learning-based protocols in underwater. Most existing research focuses on routing, while only two protocols focus on the MAC layer. One extends the lifetime of underwater acoustic sensor networks [35] and the other focuses on improving the channel utilization of underwater acoustic sensor networks [16].

III. ASSUMPTIONS & CONDITIONS

A data gathering star network topology is considered wherein the sink is located at the center and sensor nodes are uniformly distributed within the transmission range of the sink [24], [36], [38]. In this study, the sensor nodes are considered to be the last-hop nodes of a multi-hop data-gathering network. These last-hop relay nodes collect data either from the environment or from sensor nodes that are more than two-hop away from the sink and send to the sink. Such networks can be useful in real-world applications, such as military surveillance, oceanographic data collection, underwater

environment monitoring, etc. [1], [3], [5]. Time is slotted with the duration of the packet length plus guard time.

We assume that sensor nodes are synchronized with the sink and that the propagation delays are known. Synchronization in this study means that each sensor node knows the exact transmission time that corresponds to the data packet arrival time at the sink [24], [37]. Based on synchronization, the data packet is transmitted in a way that the data packet arrives at the beginning of the slot at the sink regardless of sensor–sink distance. The associated events are illustrated in Fig. 2, where the n-th data packet of $Node_i$ denoted by $D_i(n)$ and m-th data packet of $Node_j$ denoted by $D_j(m)$ are generated at any time and backed-off before being transmitted to alleviate possible collision at the sink. The transmission time was calculated by considering the targeted slot at the sink and the distance. The transmitted data packets are then received at the beginning of the targeted slot at the sink, e.g., in Fig. 2, the k-th and $(k+2)$-th slots for $D_i(n)$ and $D_j(m)$, respectively. Sensor nodes were considered static; hence, the movement of sensor nodes caused by water current and waves, for example, is ignored.

Based on Throp's underwater channel model, we assume that by using pre-acquired information on distance and frequency, each sensor node is capable of controlling the transmission power in such a way that the received power level, P_R, at the sink must be constant regardless of the locations of the sensor nodes [38].
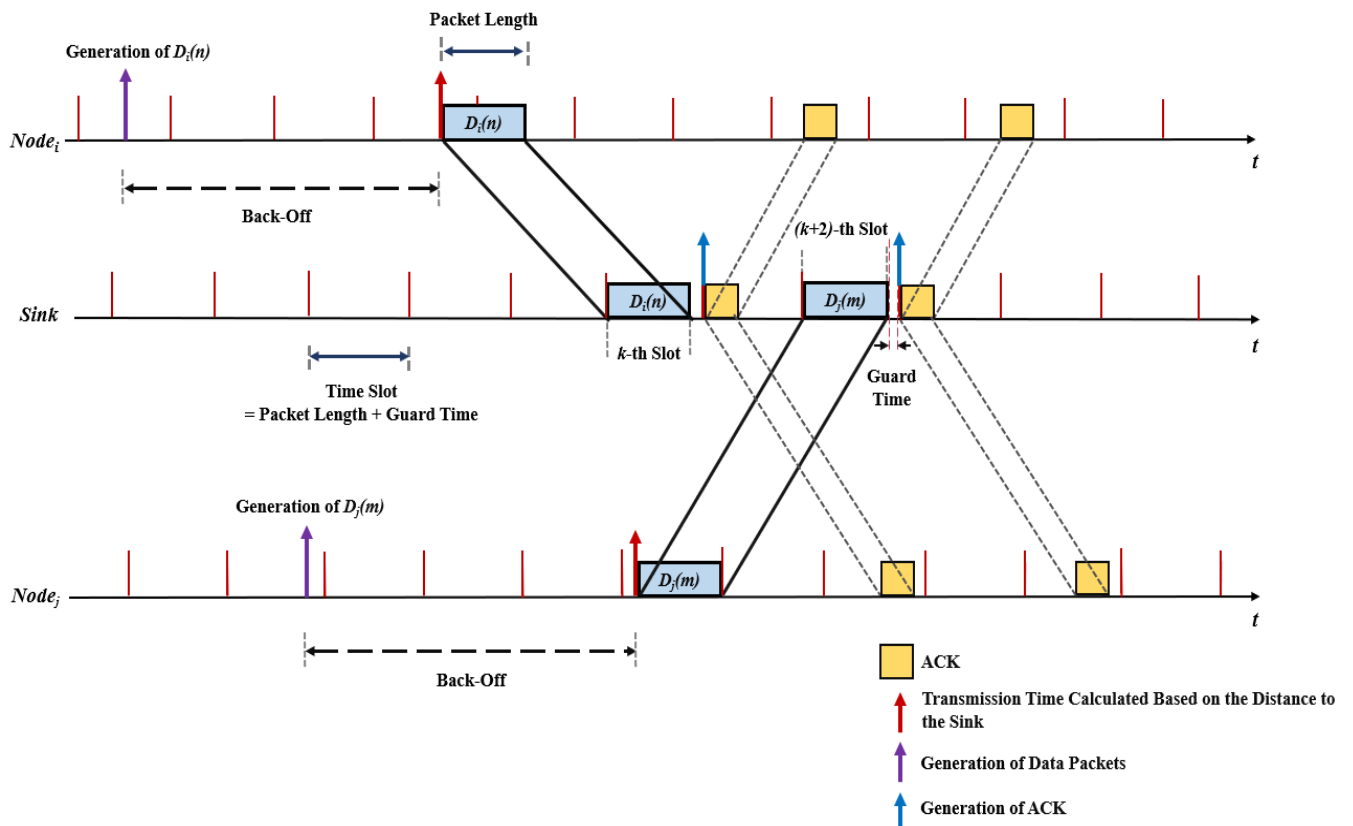

FIGURE 2. Illustration of events that happen in a time domain such as packet generation, back-off, packet transmission, and packet arrival.
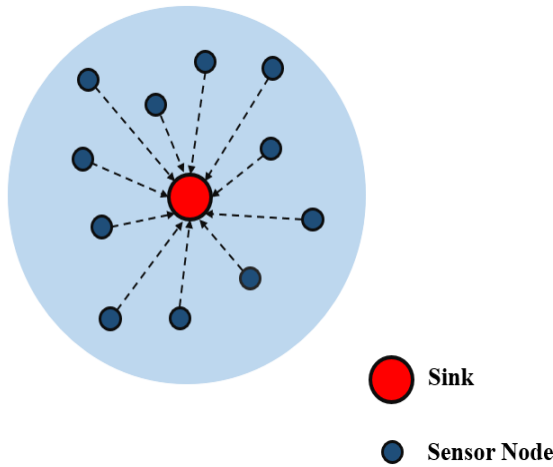
FIGURE 3. A data gathering star network topology.

IV. PROPOSED SCHEME

Where only one data packet arrives at the beginning of a given slot of the sink the transmission is considered successful. Otherwise, reception of more than one data packet at the beginning of same time slot is considered a collision.

After reception of data packets, the sink responds by broadcasting ACK or NACK at the beginning of the next slot depending on whether the data packet transmission was successful or not; thus, the corresponding sender node decides to move forward to the next data packet transmission or to retransmit the failed data packet. Unlike ACK, NACK

contains the received power level, denoted by nP_R, where n is the number of collided packets.

Fig. 4 illustrates an associated event in which the $(n+1)$-th packet of $Node_i$ denoted by $D_i(n+1)$ and $(m+1)$-th packet of $Node_j$ denoted by $D_j(m+1)$ collide at the $(k+25)$-th slot of the sink. Since the number of packets received by the sink at the $(k+25)$-th slot is 2, at the beginning of the $(k+26)$-th slot, a NACK is broadcast that contains the twice received power level, $2P_R$. This $2P_R$ value informs the $Node_i$ and $Node_j$ that two data packets collided at the sink; based on this information, both nodes will perform back-off and retransmit the collided data packets at a later time. As the proposed system works in a full duplex mode, a separate downlink channel is used for broadcasting the ACK or NACK.

The decision to select a back-off value before the transmission of generated data packets or to retransmit the collided data packets is made by sensor nodes using the Q-learning technique. Each sensor node operates as an agent of Q-learning and manages a Q table. Initially, all the Q values are set to 0 but are updated according to the obtained *reward* after sensor nodes take an *action*. In general, the following equation is considered for updating Q values:

$$Q(s_{t+1}, a_t) = Q(s_t, a_t) + \alpha[r_t + \theta \max_a Q(s_{t+1}, a) - Q(s_t, a_t)], \quad (1)$$

where $Q(s_t, a_t)$ is the current Q value obtained when *action* a_t is performed in *state* s_t. The maximum possible Q value is found by the agent in the next *state* s_{t+1}, given that a_t is taken, *reward* r_t is obtained, and the current Q value get updated.
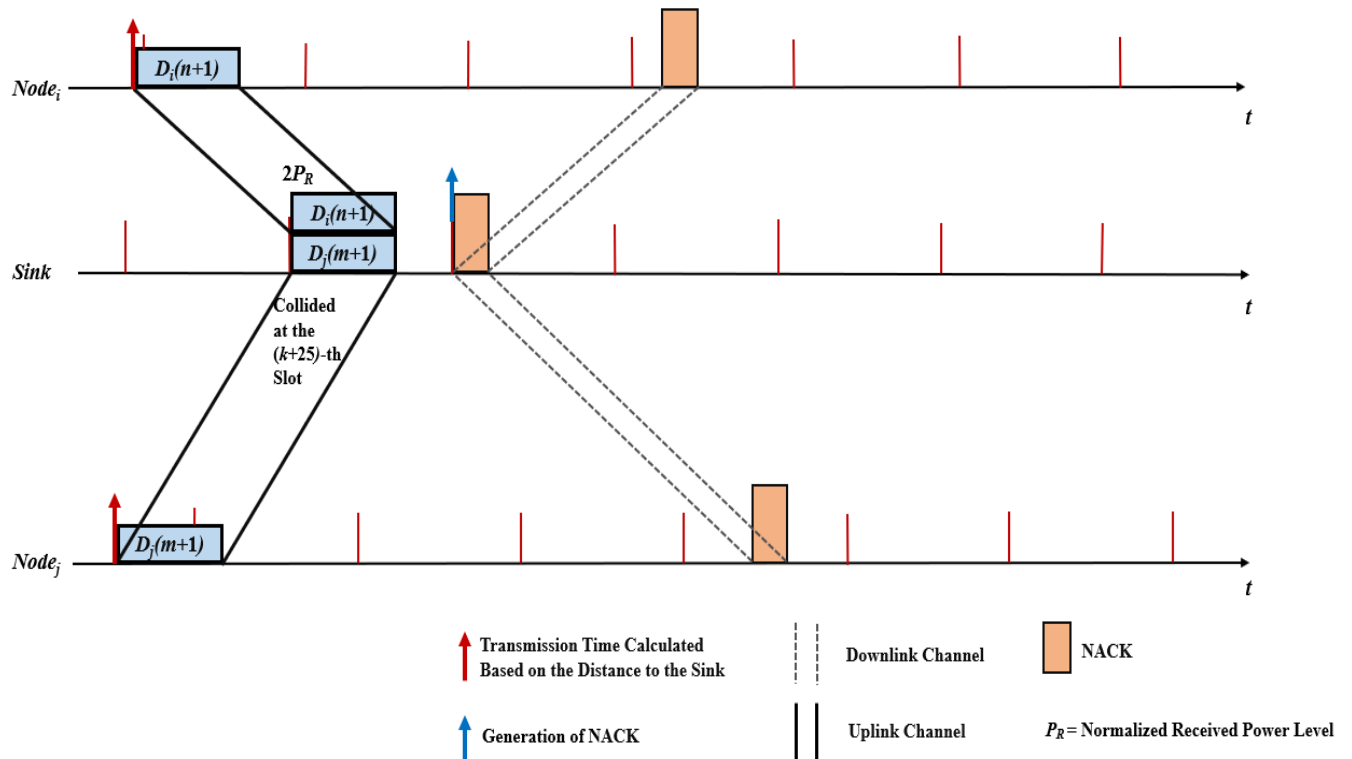

FIGURE 4. Packet collision and NACK broadcasting with received power level.

The discount factor, θ, which ranges between 0 and 1, gives more weight to either immediate *rewards* (if $\theta < 1$) or future *rewards* (if $\theta > 0$), whereas, the *learning rate*, α, which also ranges between 0 and 1, is used to tune the learning speed.

In the proposed protocol, *action* is defined as the decision made by each sensor node for selecting the back-off value denoted as B. The value of B will be in the range $[B_{\min}, B_{\max}]$. where $B_{\min}$ and $B_{\max}$ represent the minimum and maximum back-off values, respectively. *State* is defined as the receiver of data packets.

In the deployed single-hop network, only one receiver exists, i.e., the sink. Since the sink is the sole data packet receiver, the applied Q-learning technique is basically a single-*state* Q-learning technique.

In single-*state* Q-learning, only the immediate *reward* is considered because there is no new or future *state* and, accordingly, no future *reward*. Differing from the update rule shown in (1), the following expression is used in the proposed protocol for the update of Q values [11]:

$$Q_{t+1}(a) = Q_t(a) + \alpha(r_t - Q_t(a)), \qquad (2)$$

where $Q_t(a)$ and $Q_{t+1}(a)$ denote the current and new Q values at time steps t and $t + 1$, respectively.

The *reward* is the main element that controls the decrement or increment of the Q values and determine not only the behavior but also the performance of the system. After an *action* is performed, a sensor node transmits a data packet to the sink and receives a certain amount of *reward* consequent to the performed *action*. The *reward* settings differ depending on whether a collision occurred or not at the sink. If data packet reception is acknowledged by the sink, then the positive earned *reward* denoted by r_{suc} is defined as follows:

$$r_{suc} = \beta_1 \times r_1 + \beta_2 \times r_2 + \beta_3 \times r_3, \qquad (3)$$

where r_1, r_2, and r_3 are the fixed *reward* factor, back-off-related *reward* factor, and success-related *reward* factor, respectively. β_1, β_2, and β_3 are the respective related weights of the *reward* factors. The *reward* factors are given as follows:

$$r_1 = -1, \qquad (4)$$

$$r_2 = 1 - \frac{B}{B_{\max}}, \qquad (5)$$

$$r_3 = 10. \qquad (6)$$

The fixed *reward* factor r_1 represents the giving of punishment to the sensor nodes as energy is consumed when packet transmission occurs. The back-off-related *reward* factor r_2 depends on the performed *action* B. A higher value of B introduces high idle time and delay in the network, which is responsible for the degradation of channel utilization. Therefore, a lower value of B is desirable that provides higher rewards for successful data packet transmission. r_3, the success-related *reward* factor, is given when ACK broadcasting is received by sensor nodes.

On the other hand, if a collision occurs at the sink, a negative earned *reward* denoted by r_{col} is defined as follows:

$$r_{col} = \gamma_1 \times r_1 - \gamma_2 \times r_2 + \gamma_3 \times r_3^* - \gamma_4 \times r_4, \qquad (7)$$

where γ_1, γ_2, γ_3, and γ_4 are the respective *reward* weight terms. r_1 and r_2 *reward* factors are the same in both the r_{suc} and r_{col}. The other two *reward* factors, r_3^* and r_4, are the collision related *reward* factor r_3^*, which is awarded when NACK broadcasting is received by sensor nodes, and the related *reward* factor r_4 is the received power level. They can be defined as

$$r_3^* = -10, \qquad (8)$$

$$r_4 = nP_R. \qquad (9)$$

Based on the obtained *reward*, the Q values of sensor nodes are updated. Sensor nodes use these updated Q values to make the decision of selecting B values. Suppose that, at time step $t + 1$, the *reward* is r_{suc}, the following equation would then be used for the calculation of the B_{t+1} value:

$$\lfloor B_{t+1} \rceil = \begin{cases} \lceil B_t \rceil - ((Q_{t+1}(a) - Q_t(a)) \times \lfloor B_t \rceil), \\ \qquad \text{if } Q_t(a) \geq 0, Q_{t+1}(a) > 0 \\ \lfloor B_t \rceil + ((Q_{t+1}(a) - Q_t(a)) \times \lfloor B_t \rceil), \\ \qquad \text{if } Q_t(a) < 0, Q_{t+1}(a) < 0 \\ \lfloor B_t \rceil - \left(\left(\frac{Q_{t+1}(a) - Q_t(a)}{2} \right) \times \lfloor B_t \rceil \right), \\ \qquad \text{otherwise} \end{cases} \qquad (10)$$

The following equation is applied for r_{col} when calculating the B_{t+1} value:

$$\lfloor B_{t+1} \rceil = \begin{cases} \lfloor B_t \rceil + ((Q_t(a) - Q_{t+1}(a)) \times \lfloor B_t \rceil), \\ \qquad \text{if } Q_t(a) > 0, Q_{t+1}(a) > 0 \\ \lceil B_t \rceil - ((Q_{t+1}(a) - Q_t(a)) \times \lfloor B_t \rceil), \\ \qquad \text{if } Q_t(a) \leq 0, Q_{t+1}(a) < 0 \\ \lceil B_t \rceil + \left(\left(\frac{Q_t(a) - Q_{t+1}(a)}{2} \right) \times \lfloor B_t \rceil \right), \\ \qquad \text{otherwise} \end{cases} \qquad (11)$$

where $\lfloor \cdot \rceil$ denotes rounding. $Q_t(a)$ and $Q_{t+1}(a)$ represents Q value at time step t and $t + 1$, respectively. B_t is the back-off value at time step t.

Initially, when Q values are set to 0, the sensor nodes select B values randomly from the range of $[B_{\min}, B_{\max}]$. Later on, the B values are calculated using (10) or (11) based on the success or collision of a packet. A simple example of how the Q values of sensor nodes help to select B values after receiving a broadcast from the sink is given below and illustrated in Fig. 5. Suppose that, at the beginning of the process when all Q values are 0, $Node_i$ randomly chooses a B value of 25 and transmits the $D_i(n)$-th data packet after back-off at the known transmission time. The transmission is successful and the earned *reward* is positive for $Node_i$. Now, according to equation (2), the Q value of $Node_i$ is updated and the new Q value is 0.253. To simplify the calculation, it is assumed that $\beta_1 = 0.1$, $\beta_1 = 0.7$, $\beta_3 = 0.2$, $\alpha = 0.1$, and $B_{\max} = 256$. The new B value for $Node_i$ is now 19 according to (10). Based on this obtained B value, $Node_i$ performs back-off and transmits the $D_i(n + 1)$-th data packet, but this time collision occurs at the sink with the $D_j(m + 1)$-th data packet of $Node_j$. Again, the Q value of $Node_i$ is updated. This time the earned *reward* is negative for $Node_i$ and the new Q value according
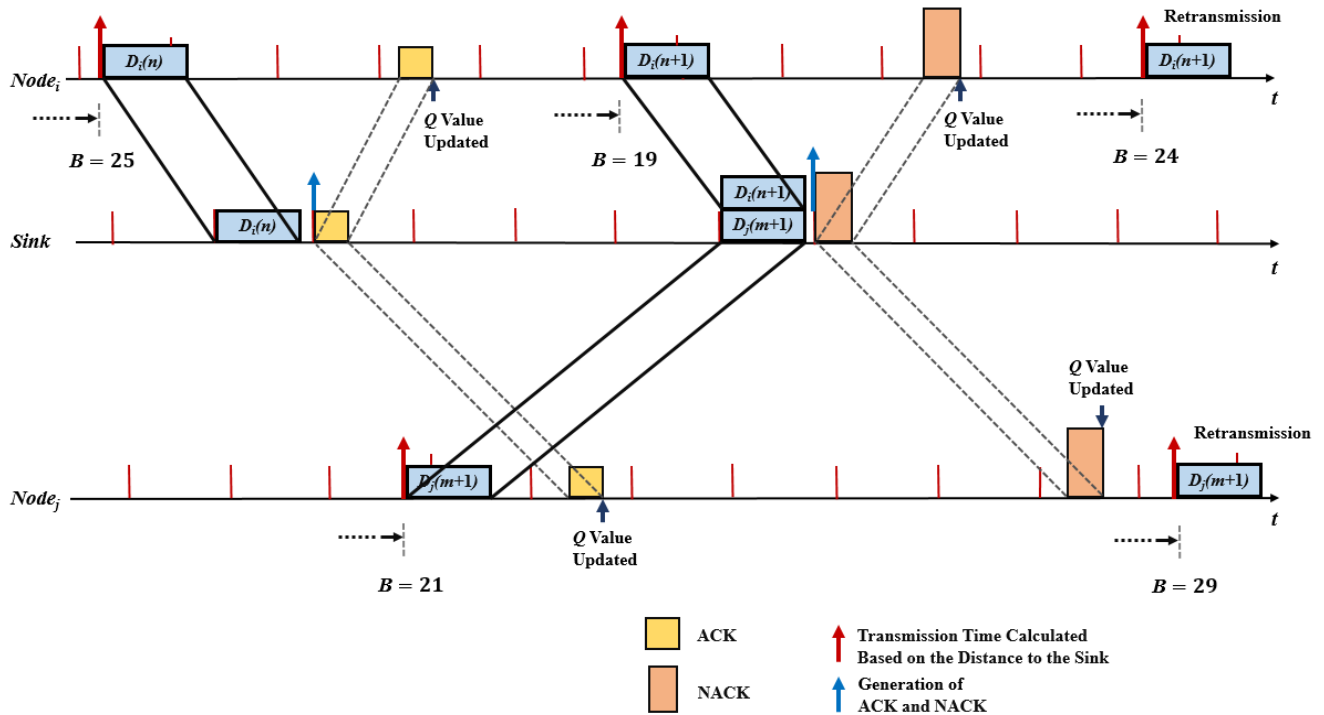
FIGURE 5. Operation of the proposed scheme.

the (2) becomes -0.30. Here, γ_1, γ_2, γ_3, and γ_4 are assumed to be 0.05, 0.05, 0.45, and 0.45, respectively, with the same values of α and $B_{\max}$ as previously mentioned. Based on (11), the new B value for $Node_i$ is now 24. $Node_i$ performs back-off according to this obtained B value and retransmits the collided $D_i(n+1)$-th data packet. If the same packet of $Node_i$ exceeds the limit for the maximum number of packet retransmissions, it will be discarded. As with $Node_i$, $Node_j$ also performs back-off according to its updated Q value and retransmits the collided $D_j(m+1)$-th data packet after selecting B with the aim of achieving successful packet retransmission.

V. SIMULATION MODEL AND RESULTS

To evaluate the proposed protocol, we provide simulation results obtained using MATLAB-(R2020a). In the simulation, 20 sensor nodes were randomly deployed within the transmission range of the sink. The traffic generated by each sensor node followed a Poisson distribution with the rate of 0.05–0.15 [packets/s]. To show the effectiveness of the proposed protocol, more simulations were performed by setting the generated traffic load as 0.1 [packets/s] where the number of sensor nodes are varied from 10 to 30. As the proposed protocol is a learning-based protocol, therefore, it is beneficial to perform the simulation to show how the performance of the proposed protocol varies over simulation time. In this regard, the proposed protocol is simulated by fixing the traffic generation rate and the number of nodes as 0.125[packets/s] and 18, respectively. Throp's underwater acoustic channel model was considered to characterize the underwater model [39]. For RX/TX power and data rate, a Teledyne Benthos ATM-903 commercial underwater modem was considered [40].

TABLE 1. Simulation parameters.

Parameter	Values
Transmission range	6 km
Propagation speed	1500 m/s
Data packet size	2376 bits
Transmission rate	2400 bps
Learning rate	0.1
Guard time	0.01sec
B_{max}	256
B_{min}	10
Max. number of retransmissions	3

The proposed protocol was compared with MACA-U [22], EIED and UW-ALOHA-Q [16]. MACA-U is a collision avoidance underwater MAC protocol in which RTS-CTS control packets are exchanged before the transmission of data packets to mitigate data packet collision at the receiver. EIED is a random back-off algorithm. To lower collision probability at the receiver, EIED increases the back-off value of a sensor node by a back-off factor κ_i if a data packet transmitted from that sensor node is involved in a collision; otherwise, the back-off value is reduced by back-off factor κ_d after a successful data packet transmission, where $\kappa_i = \kappa_d = 2$. UW-ALOHA-Q is a learning-based underwater MAC protocol adopted from terrestrial ALOHA-Q protocol where Q-learning technique is implemented in repeating frame structure to achieve collision free scheduling by finding an optimal data packet transmission time slot in each frame. Three improvements, namely, asynchronous operation,

refinement of the frame length and uniform random back-off is incorporated in UW-ALOHA-Q to address the key limitations of underwater environment.

Three performance metrics, i.e., energy efficiency, latency, and channel utilization, were considered for performance evaluation.

First, energy efficiency is defined as,

$$\varepsilon = \frac{N \times L_d}{\varepsilon_c} [bits/j] \qquad (12)$$

where L_d is the data packet size, N is the number of data packets received successfully by the sink, and ε_c is the total energy consumption for the duration over which the network was active.

Second, latency, τ, is defined as the average time interval between the generation of the data packet and successful delivery of that generated data packet at the sink. It is obtained as follows:

$$\tau = \frac{\sum\limits_{K=1}^{N} (T_{a,k} - T_{g,k})}{N} [s] \qquad (13)$$

where $T_{g,k}$ is the generation time of the k-th generated data packet, $T_{a,k}$ is the successful arrival time of the k-th generated data packet at the sink.

Finally, channel utilization, η, is defined as the ratio of the duration of successfully received data packets at the sink to the total simulation time and is defined as follows:

$$\eta = \frac{N \times T_{data}}{T_{slot} \times N_{slot}} \qquad (14)$$

where, T_{data} is the data packet duration, T_{slot} is the slot duration, N_{slot} is the number of slots during the total simulation time. Discarded data packets are not considered in this case.

Fig. 6 shows the energy efficiency for various average traffic loads. A downward performance was exhibited when average traffic load for the proposed, EIED and MACA-U protocols increased while the performance of UW-ALOHA-Q
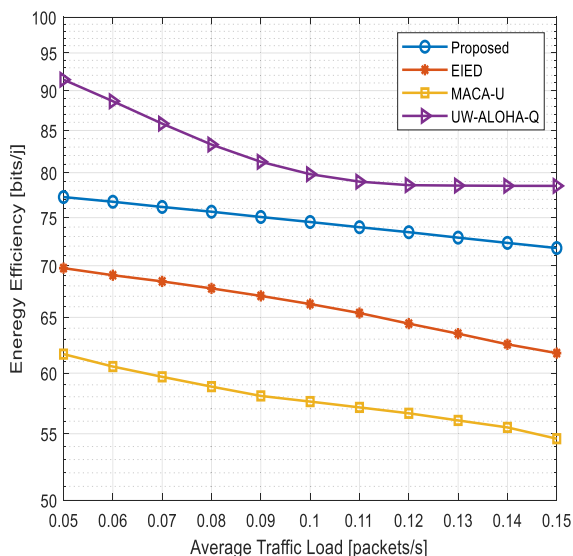
being saturated. UW-ALOHA-Q shows the highest energy efficiency as, in part, the number of collisions is greatly reduced because each node learns to determine the time-slot that has a low collision probability and, in part, because the time-slot duration is set to the maximum round trip time. However, it causes unavoidable long delay and subsequently poor channel utilization. On the other hand, for the proposed protocol and the remaining two comparison protocols the downward performance is observed with the increment of traffic load as collision occurs at the sink more frequently when the number of data packet transmission increases. The proposed protocol shows a higher energy efficiency compared with the MACA-U and EIED because more rewards are obtained as a result of increasing data packet transmission, which helps the sensor nodes to learn the back-off value for reducing data packet collision at the sink. In EIED, more collisions are introduced at the sink compared with the proposed protocol because, based on the received ACK or NACK, sensor nodes simply increase or decrease back-off values exponentially while ignoring how many packets have collided. MACA-U shows the lowest energy efficiency among the four protocols, which may be attributed to sensor nodes exchanging control packets before data packet transmission or to the propagation delay. Because of the propagation delay, there remains a possibility that control packets will collide in the process of handshaking. Both of these reasons can account for high energy consumption in the network.

Fig. 7 shows the energy efficiency for various numbers of nodes, which decreased with increasing node number. This occurs because, with an increasing number of sensor nodes, the sink faces a greater number of collisions. Consequently, data packet retransmission occurs more frequently in the network and energy consumption therefore increases. Greater energy consumption is an indication of lower energy efficiency. The energy efficiency of the proposed protocol is close to that of UW-ALOHA-Q's for the reasons mentioned
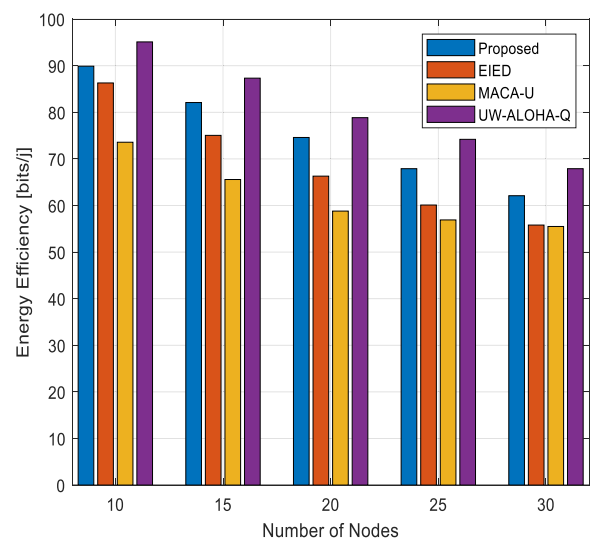

FIGURE 6. Energy efficiency versus average traffic load.

FIGURE 7. Energy efficiency versus number of nodes.

in Fig.6. For EIED and MACA-U, the trend illustrated in Fig. 6 is responsible for the poor performance compared with the proposed protocol and UW-ALOHA-Q.
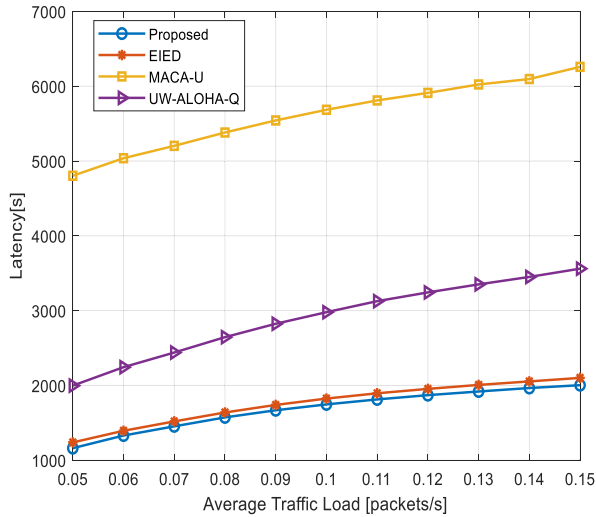

FIGURE 8. Latency versus average traffic load.

Fig. 8 shows the latency for various average traffic loads, which increased for all four protocols with increments of average traffic load. An increase in average traffic load indicates a greater number of data packet retransmissions due to a higher number of collisions at the sink; consequently, high latency is introduced into the network. Among the four schemes, MACA-U shows the highest latency. This is due to the handshaking procedure between the sender and receiver that takes place before data packet transmission, which causes a significant amount of waiting time. It is also due to the additional collisions induced by control packets during the handshaking. Where control packet collision occurs, the whole handshaking process starts again from the beginning, which introduces an indicative delay in the network. UW-ALOHA-Q shows higher latency compared to the proposed protocol and EIED due to two reasons. The first one is the large frame duration that introduces high idle time. The second one is the increase in the number of buffered data packets that are queued for transmission as the traffic load increases. Additionally, as total elimination of collisions is impossible, a certain number of retransmissions exist that increase network latency. The proposed protocol and EIED perform data packet transmission without applying the handshaking technique; thus, the latency of each protocol is much lower than that of MACA-U. As shown in Fig. 8, the lowest latency was achieved by the proposed protocol. This can be explained by the sensor nodes using Q values for the adjustment of back-off values, which results in a lower number of data packet retransmissions.

Fig. 9 shows the latency for various numbers of nodes, which increased with the increment of node numbers in the network. As shown, this latency trend was apparent for all four protocols. Generally, more nodes lead to more collisions at the sink, which increases the average delay. The proposed protocol has the lowest average latency, for reasons illustrated
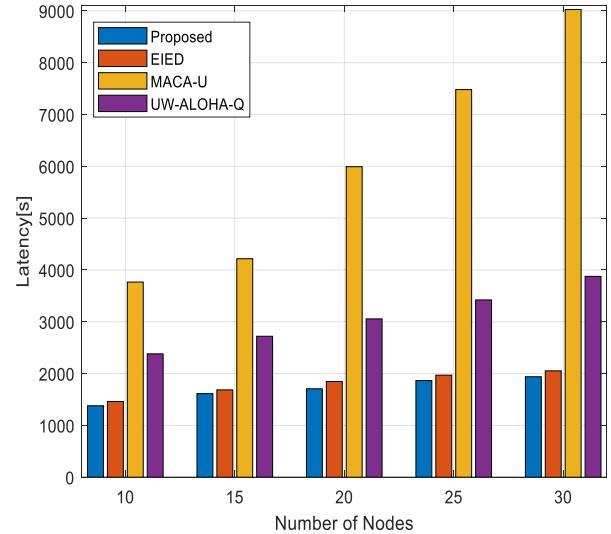

FIGURE 9. Latency versus number of nodes.

in Fig. 8. For the other three protocols, UW-ALOHA-Q, MACA-U and EIED, Fig. 8 also shows why their performance is poor compared with that of the proposed protocol.

Fig. 10 shows channel utilization, which increased as traffic load increased up to a certain point and then subsequently decreased for the proposed protocol, EIED and MACA-U. This is a general feature of channel utilization versus traffic load for contention-based protocols because, beyond a certain threshold, packet collisions become the dominant factor that determines channel utilization rather than traffic load, which proportionally increases channel utilization where traffic load is lower. In case of UW-ALOHA-Q, channel utilization increased and then saturated. Fig. 10 shows that the proposed protocol has the highest channel utilization, which reflects the ability of the Q-learning technique to select back-off values that reduce collision at the sink and thereby mitigate the average delay of the network compared with UW-ALOHA-Q, EIED and MACA-U. MACA-U had the lowest channel utilization because it performs handshaking
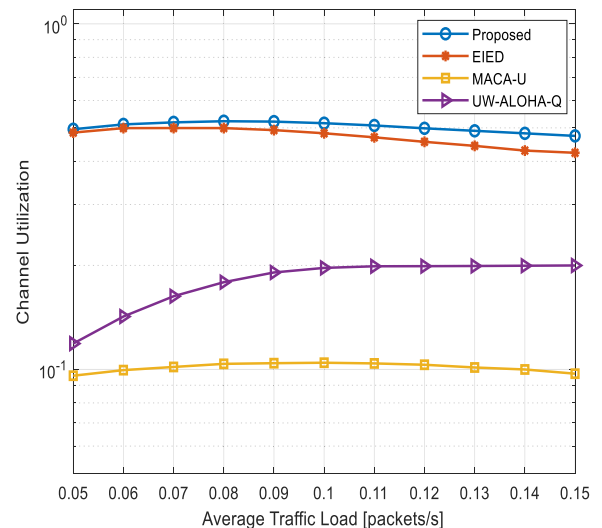

FIGURE 10. Channel utilization versus average traffic load.

to avoid possible collisions; handshaking introduces a significant amount of latency into the network. UW-ALOHA-Q has lower channel utilization compared to the proposed protocol and EIED because it has higher transmission delay due to the large frame duration. However, with incremental increases in average traffic load, the channel utilization of EIED decreases more rapidly than that of the proposed protocol because more retransmissions are performed in EIED. The advantage of fewer retransmissions in the proposed protocol arises from the adjustment of back-off based Q-learning.
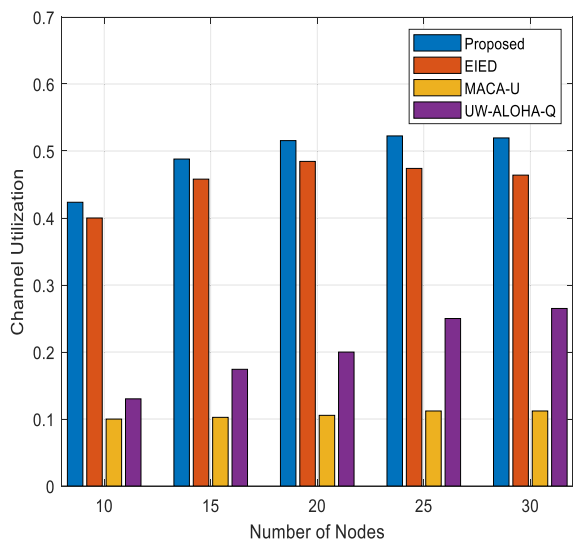

FIGURE 11. Channel utilization versus number of nodes.

Fig. 11 shows the channel utilization for various numbers of nodes in the network. Among the four protocols, the proposed protocol had the highest channel utilization for the same reasons described in relation to Fig. 10.

Fig.12 demonstrates the performances of the protocols over the simulation period. It can be seen that the proposed protocol and UW-ALOHA-Q show some variations while EIED is almost steady. This is because both the proposed protocol and UW-ALOHA-Q have a learning process, while
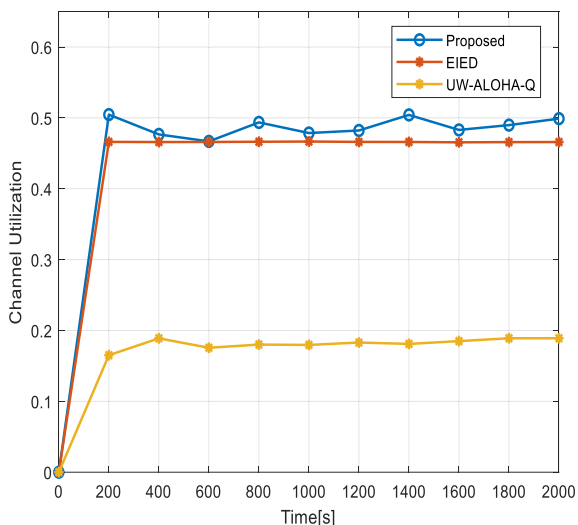

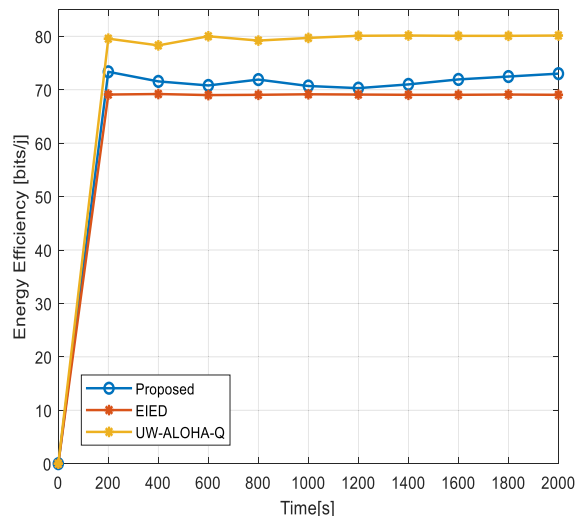**FIGURE 12.** Channel utilization versus simulation time.

FIGURE 13. Energy efficiency versus simulation time.

EIED does not, resulting in a relatively steady-state performance. Moreover, the reasons behind the proposed protocol's better performance compared to the other two protocols are already described in Fig.10.

Fig.13 follows the same description as given for Fig.12 except for the fact that in case of energy efficiency, UW-ALOHA-Q shows better performance for the same reasons described in Fig.6.

VI. CONCLUSION

Herein, a contention-based underwater MAC protocol was proposed in which a single-*state* Q-learning technique was integrated to optimize the back-off slots to reduce collision probability at the sink in underwater acoustic sensor networks. Through subsequent trial-and-error learning, the proposed protocol permitted sensor nodes to intelligently predict the number off backed-off slots that were needed to mitigate collision probability.

Through extensive simulations, the performance of the proposed protocol was compared with the UW-ALOHA-Q, MACA-U and EIED protocols in terms of energy efficiency, latency, and channel utilization. Moreover, the proposed protocol outperformed the UW-ALOHA-Q protocol in single-hop networking in terms of latency and channel utilization where for MACA-U and EIED, the proposed protocol outperformed in terms of energy efficiency, latency and channel utilization.

In future works, a multi-*state* Q-learning will be investigated to improve underwater contention-based MAC protocols in multi-hop networks.

REFERENCES

[1] Y. Noh and S. Shin, "Survey on MAC protocols in underwater acoustic sensor networks," in *Proc. 14th Int. Symp. Commun. Inf. Technol. (ISCIT)*, Sep. 2014, pp. 80–84.

[2] C.-C. Hsu, K.-F. Lai, C.-F. Chou, and K. C.-J. Lin, "ST-MAC: Spatial-temporal MAC scheduling for underwater sensor networks," in *Proc. IEEE INFOCOM 28th Conf. Comput. Commun.*, Apr. 2009, pp. 1827–1835.

[3] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Netw.*, vol. 3, no. 3, pp. 257–279, May 2005.

[4] J. Partan, J. Kurose, and B. N. Levine, "A survey of practical issues in underwater networks," in *Proc. 1st ACM Int. Workshop Underwater Netw. (WUWNet)*, 2006, pp. 17–24.

[5] S. Jiang, "State-of-the-art medium access control (MAC) protocols for underwater acoustic networks: A survey based on a MAC reference model," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 96–131, 1st Quart., 2018.

[6] G. A. Shah, "A survey on medium access control in underwater acoustic sensor networks," in *Proc. Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, May 2009, pp. 1178–1183.

[7] J. Preisig, "Acoustic propagation considerations for underwater acoustic communications network development," in *Proc. 1st ACM Int. Workshop Underwater Netw. (WUWNet)*, 2006, pp. 1–5.

[8] A. A. Syed, W. Ye, J. Heidemann, and B. Krishnamachari, "Understanding spatio-temporal uncertainty in medium access with ALOHA protocols," in *Proc. 2nd Workshop Underwater Netw.*, 2007, p. 41.

[9] J.-W. Lee and H.-S. Cho, "Cascading multi-hop reservation and transmission in underwater acoustic sensor networks," *Sensors*, vol. 14, no. 10, pp. 18390–18409, Oct. 2014.

[10] Y. Chen, J. Zhu, L. Wan, S. Huang, X. Zhang, and X. Xu, "ACOA-AFSA fusion dynamic coded cooperation routing for different scale multi-hop underwater acoustic sensor networks," *IEEE Access*, vol. 8, pp. 186773–186788, 2020.

[11] M. S. M. Alamgir, M. N. Sultana, and K. Chang, "Link adaptation on an underwater communications network using machine learning algorithms: Boosted regression tree approach," *IEEE Access*, vol. 8, pp. 73957–73971, 2020.

[12] Y. Kim, H. Lee, J. Ahn, and J. Chung, "Selection of CDMA and OFDM using machine learning in underwater wireless networks," *ICT Exp.*, vol. 5, no. 4, pp. 215–218, Dec. 2019.

[13] Z. Jin, Q. Zhao, and Y. Su, "RCAR: A reinforcement-learning-based routing protocol for congestion-avoided underwater acoustic sensor networks," *IEEE Sensors J.*, vol. 19, no. 22, pp. 10881–10891, Nov. 2019.

[14] A. Pottier, P. D. Mitchell, F.-X. Socheleau, and C. Laot, "Q-learning based adaptive channel selection for underwater sensor networks," in *Proc. 4th Underwater Commun. Netw. Conf. (UComms)*, Aug. 2018.

[15] S. Galzarano, A. Liotta, and G. Fortino, "QL-MAC: A Q-learning based MAC for wireless sensor networks," in *Algorithms and Architectures for Parallel Processing* (Lecture Notes in Computer Science), vol. 8286. Cham, Switzerland: Springer, 2013, pp. 267–275.

[16] S. H. Park, P. D. Mitchell, and D. Grace, "Reinforcement learning based MAC protocol (UW-ALOHA-Q) for underwater acoustic sensor networks," *IEEE Access*, vol. 7, pp. 165531–165542, 2019.

[17] P. Casari and M. Zorzi, "Protocol design issues in underwater acoustic networks," *Comput. Commun.*, vol. 34, no. 17, pp. 2013–2025, Nov. 2011.

[18] D. Fang, Y. Li, H. Huang, and L. Yin, "A CSMA/CA-based MAC protocol for underwater acoustic networks," in *Proc. Int. Conf. Comput. Intell. Softw. Eng.*, Sep. 2010, pp. 10–13.

[19] M. Molins and M. Stojanovic, "Slotted FAMA: A MAC protocol for underwater acoustic networks," in *Proc. OCEANS Asia–Pacific*, May 2006, pp. 1–7.

[20] B. Peleato and M. Stojanovic, "A MAC protocol for ad-hoc underwater acoustic sensor networks," in *Proc. 1st ACM Int. Workshop Underwater Netw. (WUWNet)*, Apr. 2006, pp. 113–115.

[21] C. Dong, Y. Chen, Q. Guan, F. Ji, H. Yu, and F. Chen, "Design and analysis of handshake-based MAC with delay variations in underwater acoustic networks," *Sensors*, vol. 19, no. 19, p. 4159, Sep. 2019.

[22] H.-H. Ng, W.-S. Soh, and M. Motani, "MACA-U: A media access protocol for underwater acoustic networks," in *Proc. GLOBECOM IEEE Global Telecommun. Conf.*, Nov. 2008, pp. 833–837.

[23] N. Chirdchoo, W. S. Soh, and K. C. Chua, "Aloha-based MAC protocols with collision avoidance for underwater acoustic networks," in *Proc. IEEE INFOCOM*, May 2007, pp. 2271–2275.

[24] P. Mandal, S. De, and S. S. Chakraborty, "A receiver synchronized slotted aloha for underwater wireless networks with imprecise propagation delay information," *Ad Hoc Netw.*, vol. 11, no. 4, pp. 1443–1455, Jun. 2013.

[25] C. Wang, "A dynamic back-off algorithm in ad hoc networks," *Appl. Mech. Mater.*, vols. 519–520, pp. 117–123, Feb. 2014.

[26] M. M. Rahaman, K. Ashrafuzzaman, M. S. Chowdhury, and M. O. Rahman, "Performance measurement of different backoff algorithms in IEEE 802.15.4," in *Proc. Int. Conf. Innov. Sci., Eng. Technol. (ICISET)*, Dhaka, Bangladesh, Oct. 2016, pp. 1–4.

[27] H.-M. Liang, S. Zeadally, N. K. Chilamkurti, and C.-K. Shieh, "A novel pause count backoff algorithm for channel access in IEEE 802.11 based wireless LANs," in *Proc. Int. Symp. Comput. Sci. Appl. (CSA)*, Oct. 2008, pp. 163–168.

[28] H.-H. Chin, C.-C. Lin, and D.-J. Deng, "Contention resolution algorithm for MAC protocol in wireless ad-hoc networks," in *Proc. 8th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Aug. 2012, pp. 856–859.

[29] N. Morozs, P. D. Mitchell, and Y. Zakharov, "Dual-hop TDA-MAC and routing for underwater acoustic sensor networks," *IEEE J. Ocean. Eng.*, vol. 44, no. 4, pp. 865–880, Oct. 2019.

[30] J. Zhang, M. Cai, G. Han, Y. Qian, and L. Shu, "Cellular clustering-based interference-aware data transmission protocol for underwater acoustic sensor networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3217–3230, Mar. 2020.

[31] S. Ergashev and M. Kim, "TDMA based colision free MAC protocol in underwater acoustic networks," *J. Eng. App. Sci.*, vol. 13, no. 8, pp. 9132–9140, 2018.

[32] Y. Chu, P. D. Mitchell, and D. Grace, "Reinforcement learning based ALOHA for multi-hop wireless sensor networks with informed receiving," in *Proc. IET Conf. Wireless Sensor Syst. (WSS)*, London, U.K., 2012, pp. 1–6.

[33] H. Chang, J. Feng, and C. Duan, "Reinforcement learning-based data forwarding in underwater wireless sensor networks with passive mobility," *Sensors*, vol. 19, no. 2, p. 256, Jan. 2019.

[34] D. Ye and M. Zhang, "A self-adaptive sleep/wake-up scheduling approach for wireless sensor networks," *IEEE Trans. Cybern.*, vol. 48, no. 3, pp. 979–992, Mar. 2018.

[35] L. Jin and D. Huang, "A slotted CSMA based reinforcement learning approach for extending the lifetime of underwater acoustic wireless sensor networks," *Comput. Commun.*, vol. 36, no. 9, pp. 1094–1099, 2013.

[36] Z. Zhang, W. Shi, Q. Niu, Y. Guo, J. Wang, and H. Luo, "A load-based hybrid MAC protocol for underwater wireless sensor networks," *IEEE Access*, vol. 7, pp. 104542–104552, 2019.

[37] Y. Zhou, K. Chen, J. He, and H. Guan, "Enhanced slotted aloha protocols for underwater sensor networks with large propagation delay," in *Proc. IEEE Veh. Technol. Conf.*, May 2011, pp. 1–5.

[38] B.-M. Seo, J. Cho, and H.-S. Cho, "A signaling-free underwater code division multiple access scheme," *Electronics*, vol. 8, no. 8, p. 880, Aug. 2019.

[39] E. M. Sozer, M. Stojanovic, and J. G. Proakis, "Underwater acoustic networks," *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 72–83, 2000.

[40] *903 Series ATM-903*. Accessed: Jan. 15, 2021. [Online]. Available: http://www.teledynemarine.com/903-series-atm-903

FAISAL AHMED (Student Member, IEEE) received the B.Sc. degree in electrical and electronic engineering (EEE) from East West University (EWU), Bangladesh, in 2017. He is currently pursuing the M.Sc. degree in electronic and electrical engineering with Kyungpook National University (KNU), Daegu, South Korea. He is also working as a full-time Graduate Research Assistant (GRA) with Mobile Communication Laboratory (MCL). He was a Graduate Teaching Assistant (GTA) with the Department of Electrical and Electronic Engineering (EEE), EWU. His research interests include underwater acoustic sensor networks, medium access control protocol, and reinforcement learning. He is a Student Member of KICS.

HO-SHIN CHO (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 1992, 1994, and 1999, respectively.

From 1999 to 2000, he was a Senior Member of the Research Staff with the Electronics and Telecommunications Research Institute (ETRI), where he was involved in developing a base station system for IMT-2000. From 2001 to 2002, he was a Faculty Member with the School of Electronics, Telecommunications, and Computer Engineering, Hankuk Aviation University. In 2003, he joined the School of Electronics Engineering, Kyungpook National University, as a Faculty Member, where he is currently a Professor. His current research interests include traffic engineering, radio resource management, medium access control protocol for wireless communication networks, cellular networks, underwater communication, and molecular communication. He is a member of IEICE, IEEK, KICS, and ASK. He received the Rising Researcher Fellowship of NRF, in 1998.

• • •