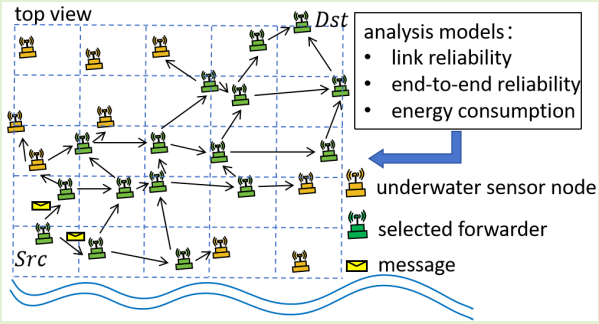
CPR: A Confined Percolation Routing for Distributed Underwater Acoustic Networks

Yuan Liu, *Member, IEEE*, Lin Cai, *Fellow, IEEE*, Fangyong Wang, Haiyan Wang, *Member, IEEE*, and Junhao Hu

***Abstract*—A routing protocol is essential for multihop relay to deliver data and reduce energy consumption in underwater acoustic networks (UANs). However, the low connectivity and high transmission power requirement of underwater acoustic channels pose critical challenges to network connectivity and the lifetime. In this article, we present a confined percolation routing (CPR) protocol that enhances connectivity and energy efficiency in UANs by exploring multiple paths to meet reliability requirements and optimizing transmission power and retransmission counts to reduce energy consumption. The analysis models designed assess the impact of activated node combinations and link reliability on end-to-end (ETE) reliability and energy use. The proposed protocol has been tested in various scenarios, and the results show that the proposed protocol improved the average ETE reliability by 17%, 21%, 7%, 12%, and 44% compared with the benchmark GEDAR, EEGNBR, Dflooding, Multi-SPR, and LEACH protocols in a 5 × 5 network, while consuming lower energy.**



***Index Terms*— Energy efficiency, routing protocol design, underwater acoustic network (UAN), underwater communication.**

I. Introduction

THE 21st century has opened a new era of comprehensive exploitation and utilization of marine resources. Underwater acoustic networks (UANs) have gained attention as an emerging information network. Real-time data collection near the sea floor is crucial for scientists and managers to respond in time when important events happen, especially in cases of tsunamis, hurricanes and storms, eddies, harmful algal blooms, and so on [2]. UANs typically comprise underwater sensor nodes, surface relay stations, ground data control center, etc. These sensor nodes are randomly deployed in the detection area to monitor all kinds of information. They autonomously organize into a network, with each node equipped with a low-bandwidth acoustic modem and multiple sensors [3]. Using the acoustic modem, they can communicate with each other, and the collected information is relayed progressively by neighboring nodes to the intended destination.

The preferred solution for achieving real-time monitoring in specific oceanic regions is to establish a network where various instruments interconnect by acoustic wireless links [4]. Accordingly, Li et al. [5] showed that properly introducing a relay can reduce the network energy consumption with a small increase on the end-to-end (ETE) delay.

However, the propagation delay in underwater acoustic channels is significant due to the slow speed of sound (1500 m/s), and the available bandwidth is narrow with a low data rate (measured in kb/s), resulting in large information transmission delays. This characteristic imparts severe spatiotemporal uncertainties to UANs. In addition, strong interference, time–frequency spreading, and time-varying fading lead to high bit error rates and packet loss rates in point-to-point underwater acoustic communications, rendering the network links to exhibit random ON–OFF characteristics.

The connectivity issue is predominantly influenced by the specifics of the UANs and the scenarios under consideration.

Manuscript received 29 June 2024; accepted 17 July 2024. Date of publication 7 August 2024; date of current version 13 September 2024. This work was supported in part by Hanjiang National Laboratory and in part by the Key Program of the National Natural Science Foundation of China under Grant 62031021. An earlier version of this paper was presented in part at the OCEANS 2021: San Diego—Porto [DOI: 10.23919/OCEANS44145.2021.9705955]. The associate editor coordinating the review of this article and approving it for publication was Prof. Jaime Lloret. *(Corresponding author: Yuan Liu.)*

Yuan Liu is with Hanjiang National Laboratory, Wuhan, Hubei 430061, China, and also with the Key Laboratory of Ocean Acoustics and Sensing, School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China (e-mail: yuan_l@mail.nwpu.edu.cn).

Lin Cai and Junhao Hu are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada (e-mail: cai@ece.uvic.ca; hujunhao0404@outlook.com).

Fangyong Wang is with Hanjiang National Laboratory, Wuhan, Hubei 430061, China.

Haiyan Wang is with the Key Laboratory of Ocean Acoustics and Sensing, School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: hywang@sust.edu.cn).

Digital Object Identifier 10.1109/JSEN.2024.3437211

Petroccia et al. [6] reported that a mere 12.5% of the links maintain a packet delivery ratio exceeding 0.75, with the majority showcasing ratios between 0.25 and 0.5. Such diminished packet delivery ratio poses substantial hurdles in achieving routing reliability. The network's resilience is put to the test particularly when standard communication paths are interrupted by unforeseen node or link failures, leading to substantial delays in re-establishing alternative routes. Thus, implementing multipath strategies emerges as a pivotal solution to bolster data transmission reliability amidst such network unpredictability.

Moreover, the nodes in UANs have limited energy reserves, which are extremely challenging to replenish. For the same communication distance, the power consumption of underwater acoustic transmitters is an order of magnitude higher than that of electromagnetic wave transmitters. Acoustic transmissions, in comparison to terrestrial microwave communications, demand significantly higher energy, with commercial underwater modems consuming between 1 and 40 W during transmission (Sendra et al. [7] and Khan et al. [8]). Notably, about 39.4% of these modems operate at power levels ranging from 1 to 8 W, 36.4% function within the 10–20-W range, and 24.2% expend energy within the 30–40-W spectrum according to the statistics [9]. The transmission power requirement tends to double as the transmission distance increases, showcasing a nonlinear relationship with distance. Consequently, optimizing packet retransmission counts and configuring appropriate transmission power levels are crucial for minimizing the energy expenditure.

While horizontal transmissions face multipath propagation effects, vertical transmission channels are comparatively less affected, exhibiting reduced signal distortion and thus providing a more stable communication route [10]. Considering the importance of horizontal transmission and ocean bottom monitoring application, this study zeroes in on the intricate routing challenges among sensor nodes located on the ocean floor. Emphasizing horizontal transmission strategies, our research endeavors to navigate past the complexities associated with horizontal transmissions, aiming to enhance both the reliability and efficiency of the data exchange processes within the UAN framework.

In this extended version of our work, key improvements include an expanded review of related literature to solidify the study's motivation, the introduction of a new network structure, and the development of a sophisticated ETE reliability analysis model based on graph decomposition. Furthermore, we introduced an energy consumption model and provided an in-depth description of the proposed protocol, alongside a comprehensive performance evaluation [1].

In this article, a confined percolation routing (CPR) for UANs is proposed to guarantee reliable packet delivery with less energy consumption using a way of multipath percolation. The relationship among the number of activated nodes, link reliability, and ETE reliability is developed based on graph decomposition, which guides each node to set an activated or inactivated status at local to satisfy ETE requirement and reduce energy consumption. The main contributions of this article are as follows.

1) First, under the characteristics of underwater acoustic communication, we develop a conjoint analysis for ETE reliability considering link reliability and energy consumption. This framework provides an essential basis for designing and optimizing routing protocols.
2) Second, a routing protocol is proposed by leveraging the above analytical framework. The link selection strategy with suitable link reliability is provided for packet percolation delivery, based on the joint analysis with link reliability, ETE reliability, energy consumption, and the optimal combination of activated nodes. Meanwhile, the configurations of transmission power and the maximum number of retransmission are optimized to minimize energy consumption.
3) Third, the performance of CPR is compared with five underwater routing benchmarks, including the Dflooding, Multipath SPR, GEDAR, EEGNBR, and LEACH. The results show that the proposed CPR outperforms benchmark schemes in terms of ETE reliability and energy consumption. In addition, to investigate the influence of the harsh underwater acoustic channel on CPR, the real-world channel measurement data have been used in the experiment. The results show the superior performance of the proposed CPR routing.

The remainder of this article is structured as follows. In Section II, we review related research. Section III outlines the system model. Section IV introduces the CPR routing protocol. Section V presents the performance evaluations, followed by concluding remarks and further research issues in Section VI.

II. Related Work

In this section, the existing routing protocols for UANs are introduced. We categorize these protocols into energy-based and reliability-based approaches, highlighting their strategies and limitations. By examining both traditional and contemporary methods, we underscore the unique challenges posed by underwater environments and the necessity for specialized routing protocols to address these challenges effectively.

The existing routing protocols for wireless sensor networks (WSNs) can be broadly classified into proactive and reactive categories [11]. In proactive protocols, nodes continuously maintain and broadcast routing information to update network routes. Reactive protocols, conversely, initiate route discovery only for new transmissions. Hu et al. [12] proposed a directed percolation routing to guarantee highly dependable and low-latency services in satellite networks, where the locations of all the satellites with a predefined orbit and speed at any time instant are known. However, the underwater environment's complexity and dynamism challenge such protocols. Underwater sensor nodes face limitations in computational power, storage, and energy, contending with issues such as high signal attenuation, prolonged transmission delays, unstable connections, and substantial power requirements. These factors contribute to inefficient packet delivery and elevated energy consumption in UANs, rendering conventional WSN routing protocols ineffective. This scenario underscores the

necessity for new routing protocols tailored to the unique demands of UANs.

Our investigation primarily focuses on two types of underwater routing protocols: energy-based and reliability-based. Energy-based protocols aim to optimize sensor nodes' energy consumption, thereby enhancing the network's longevity where possible. Reliability-based protocols, on the other hand, concentrate on efficient data forwarding from the source to the destination, ensuring consistent transmission integrity throughout the process.

The primary strategies for enhancing energy efficiency can be classified into two categories: energy-aware and cluster-based. Energy-aware strategies allow sensor nodes to select paths with minimal energy costs, leveraging the nodes' energy status. Protocols such as E-CARP, introduced by Zhou et al. [13], use a location-free, greedy, hop-by-hop forwarding strategy for energy efficiency. Zhou et al. [14] explored a Q-learning-based anypath routing, aiming to prolong UANs' lifespan and minimize packet routing delays. Gopi et al. [15] presented a protocol optimizing energy through strategic layering. To save overall energy consumption, the protocol optimizes the layering phase, and the packet delivery probability and energy consumption guide the selection of layer widths and transmission energy for nodes. Bouabdallah et al. [16] analyzed how underwater conditions influence energy consumption balance, showing that adjusting transmission power can distribute the load evenly across nodes.

Cluster-based strategies divide the network into clusters to boost efficiency, with cluster heads central to data relaying, thereby economizing data transmission and energy. Game-theory-based clustering schemes [17] and hierarchical protocols incorporating Q-learning [18] have been proposed to optimize network longevity and balance energy costs. Addressing network structure, energy cost, and balance, the authors combined the Q-learning with clustering routing to achieve cost reduction. In the paper [19], challenges regarding relay node placement and flow allocation within UANs were effectively resolved. The relay selection algorithm in this context contributes to energy conservation by reducing the number of packet relays. In recent studies, the EEGNBR protocol [20] was proposed as a reliable, efficient, and time-effective routing solution for UANs, emphasizing an energy-efficient guiding network.

On the other hand, reliable data forwarding also is a crucial goal within network routing protocol, emphasizing strategies like opportunistic routing (OR) and flood-based routing. OR involves selecting a neighbor subset for collaborative packet advancement toward the destination, using geographic attributes in UANs, as seen in focused beam routing (FBR) [21] and energy-efficient probabilistic depth-based routing (DBR) [22]. GEDAR [23], a combined geographic and OR protocol, enhances packet delivery to multiple sinks and incorporates topology control to solve void region problems. DBR [24] leveraged the broadcast nature of underwater nodes, using a greedy approach for depth-reduced packet forwarding, although it primarily focuses on energy efficiency without addressing connectivity and reliability comprehensively. Vector-based forwarding (VBF) [25], [26] created a virtual pipeline for selecting relay nodes, improving routing precision. Guan et al. [27] proposed an OR based on distance vector to solve the issues of the detouring forwarding and void region. In addition, directional flooding-based protocols [28] and machine-learning-infused methods, such as Q-learning and reinforcement learning [29], are being explored to enhance routing, particularly in sparse topologies, by optimizing for factors such as energy, delay, and link quality. With the widespread application of machine learning algorithms, reinforcement-learning-based OR has also emerged in underwater acoustic routing design [30], [31]. The OR based on reinforcement learning introduced a reward function that can consider several factors (i.e., energy, delay, and link quality) to search for the optimal result, especially for the sparse topology scenario.

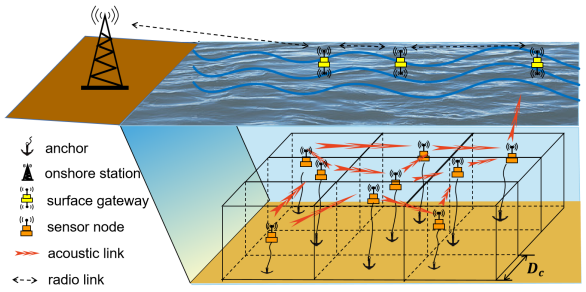


Fig. 1. System model of the UAN.

In summary, underwater routing literature suggests that strategies such as energy-aware and cluster-based approaches can mitigate high energy consumption to an extent. The energy-aware strategy requires knowledge of the next-hop candidate node set, selecting the subset with the lowest energy use, and continuously updating the residual energy. In the cluster-based approach, electing a cluster head wisely is crucial since it bears the majority of the data load. Moreover, OR and flooding-based protocols enhance reliability using more delivery opportunities and paths, albeit at the expense of increased delay and traffic. Improving routing effectiveness to achieve reliable and energy-efficient data delivery across various scenarios remains an open challenge.

III. System Model

In this section, the network architecture is presented. Subsequently, the link probability, the ETE reliability analysis model, and the energy consumption model proposed for protocol design are described.

A. Network Structure

UANs can be described as a set $G(V, E)$ where V is the set of nodes, and E consists of all the acoustic links. A reference architecture for UANs, depicted in Fig. 1, illustrates each sensor node anchored to the ocean floor within a unique cube unit for area monitoring. Nodes are deployed at intervals of $D_c \times D_c$ within their designated areas. Due to anchoring, the mobility of each node is confined to its cube unit, though dynamic movements occur due to ocean currents. This kind of movement property is hard to predict and leads to the unreliable connectivity of links. The routing selection is organized

TABLE I
SYMBOL ERROR PROBABILITIES

Modulation Schemes	Symbol Error Probabilities
FSK	$P_e = \sum_{k=1}^{M-1}(-1)^{k+1}\binom{M-1}{k}\frac{1}{k+1}e^{-\frac{kr}{k+1}}$
PSK	$P_e \approx 2Q\left(\sqrt{2kr}\sin\left(\frac{\pi}{M}\right)\right)$
QAM	$P_e \approx 4Q\left(\sqrt{\frac{3kr}{M-1}}\right)$

into discrete time periods, assuming hop counts remain stable within each initialization period. To accommodate variations in hop counts caused by ocean currents, topology discovery is periodically initiated by the sink node, allowing for hop count adjustments across different initialization periods. We introduce the multipath to explore multiple paths to overcome the connective unreliability. We explore the multiple paths to overcome the unreliable connectivity from the source to the destination.

In the routing procedure, nodes can be in one of the two statuses: activated and inactivated. Nodes in the set $\mathcal{N}_a$ are activated for message forwarding; others remain inactivated. Each node has a unique ID based on its position in the grid deployment. The connectivity probability between node pairs varies due to random deployment and ocean currents.

The percolation routing strategy: link reliability indicates the probability of successfully delivering a packet over a link. ETE reliability reflects the chance of forwarding a packet from source to destination successfully. Due to low link reliability in UANs, enhancing delivery ratio involves exploiting path diversity, allowing packets to be delivered in a percolation manner. The parameter P_{link} akin to the "edge open" probability in the percolation theory. To satisfy the expected ETE reliability, activated node sets are determined during link selection. Packets are forwarded within a subgraph composed of activated nodes $\mathcal{N}_a$ and links. This process not only reduces energy consumption by engaging only a subset of nodes but also further conserves energy as outlined in Section IV-A. Given the lattice UAN network topology, each forwarding node needs to forward a packet to its neighbors closer to the destination, leveraging multipath forwarding. As a result, a node may receive the same packet multiple times but processes each packet only once, discarding duplicates. To facilitate this, nodes temporarily store the IDs of received packets, discarding any with matching IDs. Given UANs' limited data volume, the process of ID comparison and redundant packet removal is practical.

B. Link Reliability Analysis

In underwater acoustic communication, the symbol error probabilities for different modulation schemes are shown in Table I, where M represents the number of different symbols used in the modulation scheme, and k denotes the number of bits carried by each symbol, with $k = \log_2 M$. In this section, we calculate the reliability of acoustic links, considering the characteristics of an acoustic channel and the adoption of frequency shift keying (FSK) modulation. Rayleigh fading with an additive Gaussian noise model has been widely used in UANs' routing protocol design [23], [32], [33], which is also adopted in our work.

Narrowband filtration and envelope detection are carried out individually for each frequency channel in FSK signal. Its bit error ratio (BER) is given by

$$P_e = \frac{1}{2}e^{-r/2}. \tag{1}$$

Here, r is calculated as $a^2/2\sigma_n^2$, with a representing signal amplitude and σ_n^2 denoting noise variance.

In the context of underwater acoustic detection, the receiver's detection index D is formally defined as per the notation provided in [34]

$$D = [M_s(s+n) - M_s(n)]^2/\sigma^2. \tag{2}$$

Here, $M_s(s+n)$ represents the average amplitude of the combined signal and noise, while $M_s(n)$ denotes the average amplitude of the noise component alone. The envelope distribution associated with Gaussian noise follows the Rayleigh distribution, with its mean value defined as $M_s(n) = E(X_0) = (\pi/2\sigma_n)^{1/2}$. When narrowband noise is combined with sinusoidal signals, the resulting distribution follows the Rician distribution. Its mean is given by

$$\begin{aligned} M_s(s+n) &= E(X_1) \\ &= \int_o^\infty X\frac{X}{\sigma_n^2}\exp\left[-\frac{1}{2\sigma_n^2}(X^2+a^2)\right]J_0\left(\frac{aX}{\sigma_n^2}\right)dx. \end{aligned} \tag{3}$$

When $a \gg \sigma_n^2$, $M_s(s+n)$ can be denoted as $E(X_1 = a)$, leading to $D = [E(X_1 - E(X_0))]^{2/\sigma_n^2}$. In underwater acoustic detection, when $(a/\sigma_n^2) \gg (\pi/2)^{1/2}$, the receiver defines the detection index D as follows:

$$D = a^2/\sigma_n^2. \tag{4}$$

Here, $r = a^2/2\sigma_n^2$. Therefore, $r = D/2$, $\text{SNR} = 10\log D$. Assuming that SL is the source level, TL is the transmission loss, NL is the noise level, DI_R is the directivity index, B is the working bandwidth, and the fluctuation of signal–noise ratio is denoted by F. From the underwater acoustic theory, SNR is given by

$$\text{SNR} = \text{SL} - \text{TL} - \text{NL} + \text{DI}_R - 10\log B - F. \tag{5}$$

For the detailed calculation procedures of these variables in (5), refer to the paper [1].

After accounting for fading fluctuations, we use matched filtering and envelope detection as the processing method, resulting in the calculation of the BER given by

$$P_e = 1/(\overline{Z}^2 + 2) \tag{6}$$

where

$$\overline{Z}^2 = \frac{a^2}{2\sigma^2} \tag{7}$$

$$\text{SNR} = 10\log\frac{a^2}{\sigma^2}. \tag{8}$$

Linear block codes are used for error correction, with a minimum Hamming distance of d. In this case, the number of correctable errors satisfies the condition $t = \lfloor (d-1)/2 \rfloor$.

The bit error rate at the input of error correction decoding is denoted as P_e, leading to the code word error rate after error correction of

$$P_{\text{we}} \approx \sum_{i=t+1}^{n_c} \binom{n_c}{i} P_e^i (1-P_e)^{n_c-i}. \tag{9}$$

The code element length is $n_c = K_{\text{inf}}+r$, where K_{inf} represents the length of information bits, and r is the length of the redundancy code. Consequently, the bit error rate after error correction can be expressed as

$$P_e' = 1 - \sqrt[K_{\text{inf}}]{1-P_{\text{we}}}. \tag{10}$$

BER is

$$\text{BER} = \frac{2^{k_o-1}}{2^{k_o}-1} P_e' \tag{11}$$

where $k_o = \log_2 M$. Finally, the link reliability or the link probability (used interchangeably in this article) can be expressed as

$$P_{\text{link}} = \exp(L \ln(1-\text{BER})) \tag{12}$$

where L, measured in bits, signifies the packet length.

P_{link} is the link reliability for one-time transmission. r_n represents the number of retransmissions. When $r_n > 1$, the link reliability is denoted by P'_{link} represented the probability of at least one successful transmission in r_n trials

$$P'_{\text{link}} = \sum_{i=0}^{r_n} (1-P_{\text{link}})^i \times P_{\text{link}}. \tag{13}$$

Thus, the reliability of a link is dependent on the number of transmission attempts and the probability of successful transmission in a single attempt.

C. ETE Reliability Analysis Based on Graph Decomposition

The recursive process is used to derive the reliability or ETE reliability (used interchangeably in this article). For a lattice with size $m \times n$, the probability that a message from node (m, n) can reach node $(0, 0)$ is given by [35]

$$P(0,0) = P(A) + \sum_{i=0}^{m+n-1} P(\mathcal{B} \mid S_i) P(S_i) \tag{14}$$

where A is the edge path in the upper left corner and $\mathcal{B}$ is the union of all other paths. It defines S_i as the event that the last i edges along A leading to the destination are all connected, but the last $(i+1)$th one is not, so $P(S_i) = P_{\text{link}}^i(1-P_{\text{link}})$, and P_{link} is the link reliability probability. $P(\mathcal{B}|S_i)$ is

$$\begin{aligned}
P(\mathcal{B}|S_0) &= P_{\text{link}} \cdot P(m, n-1) \\
P(\mathcal{B}|S_1) &= P(\mathcal{T}((1,0), \sqcup_{i=1}^{n-1}(0,m), m)) \\
&\cdots \\
P(\mathcal{B}|S_{m-1}) &= P(\mathcal{T}((m-1,0), \sqcup_{i=1}^{n-1}(0,m), m)) \\
P(\mathcal{B}|S_m) &= P(\mathcal{T}((m-1,0), \sqcup_{i=1}^{n-1}(0,m), m)) \\
P(\mathcal{B}|S_{m+1}) &= P\big(\mathcal{T}\big((m-1,0), (0,m-1), \\
&\qquad \sqcup_{i=2}^{n-1}(0,m), m\big)\big) \\
&\cdots \\
P(\mathcal{B}|S_{m+n-2}) &= P\big(\mathcal{T}\big((m-1,0), \\
&\qquad \sqcup_{i=2}^{n-2}(0,m-1), (0,m), m\big)\big) \\
P(\mathcal{B}|S_{m+n-1}) &= P\big(\mathcal{T}\big((m-1,0), \sqcup_{i=2}^{n-1}(0,m-1), m\big)\big)
\end{aligned} \tag{15}$$

where a universal structure $\mathcal{T}$ is introduced to articulate the recursion among graph decomposition. The source is positioned at the bottom-left vertex (m, n), while the target is located at the top-right vertex $(0, 0)$. On each layer, there are two types of structural blocks: triangles $\triangle$ and rectangles $\square$. The ith layer can be denoted by (t_i, r_i), where t_i is the number of $\triangle$ and r_i the number of $\square$, and the number of edges at the bottom layer is represented by b. The structure $\mathcal{T}$ can then be defined for all the layers as $\mathcal{T}((t_0,0), \sqcup_{i=1}^{K}(t_i,r_i), \sqcup_{i=K+1}^{N}(0,r_i), b)$, where $\sqcup$ signifies a series of layers, distinguishing four types of layers: 1) the topmost layer containing only triangles; 2) layer 1 to K are mixed layers of triangles and rectangles; 3) layers $K+1$ to N containing only rectangles; and 4) the bottom layer with b edges. For instance, $\mathcal{T}(\sqcup_{i=0}^{n-1}(0,m), m)$ represents the initial $m \times n$ network structure.

D. Energy Consumption Model

From Section III-A, link reliability P_{link} is a function of SNR. P_t and P_n are the transmission power and the noise power, respectively, where

$$\text{SNR}(R, f) = 10 \log\left(\frac{P_t}{P_n}\right) \tag{16}$$

$$P_t = 10^{\frac{\text{SL}-170.8}{10}}. \tag{17}$$

Meanwhile, P_t is increased with respect to the link distance R as P_t is a function of R. The increase in transmission power or retransmission can cause P_{link} increases.

The energy consumption for node i can be expressed as

$$E_i(R, r_n) = \frac{\left(P_t^i(R) \times (r_n^i + 1) + P_r^i\right) \times L}{R_b} \tag{18}$$

where $r_n^i \in N^+$ is the number of retransmissions, $P_t^i(R)$ is the transmission power, and P_r^i is the received power at node i. The total energy consumption is given by

$$E_{\text{total}} = \sum_{i \in \mathcal{N}_a} E_i(R, r_n). \tag{19}$$

IV. CPR Protocol

Given the constraints on ETE reliability, minimizing energy consumption emerges as a primary goal for routing protocols in UANs. The percolation-based multipath routing protocol explores path diversity to improve the delivery ratio. Energy costs are minimized by meticulously designing the percolation scope and transmission configurations at each hop. Based on the link reliability and energy analysis, the transmission power and the maximum number of retransmissions for each link can be configured to save total energy costs.

CPR involves topology discovery with energy optimization (TDEO) and link selection. In the TDEO phase, nodes exchange HELLO packets with neighbors to ascertain link reliability and hop count, thus forming the lattice topology. Subsequently, nodes calculate optimal transmission power and

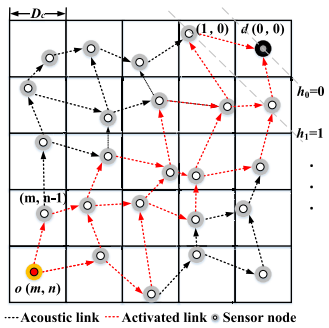


Fig. 2. Directed, loop-free graph can be constructed by initialization procedure.

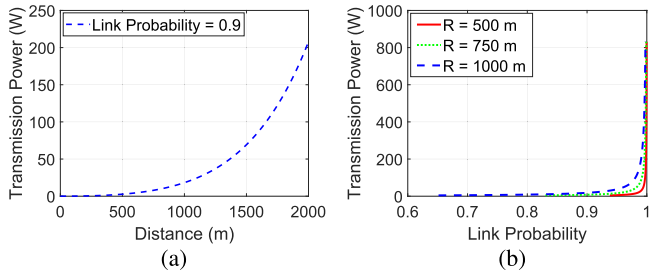


Fig. 3. Relationship among distance, link reliability, and transmission power: transmission power for different (a) distances and (b) link reliabilities.

retransmission counts to optimize energy use. In the link selection phase, informed by the initial phase's findings and the requisite ETE reliability, links are activated or deactivated to facilitate message delivery from the source.

A. Topology Discovery With Energy Optimization

After random deployment of each sensor node into a cube with length D_c, the network is simplified as an $m \times n$ lattice. Sensors communicate directionally with neighbors in adjacent grids via acoustic links. In percolation routing, nodes are labeled by their hop-count value and ID. The TDEO procedure labels nodes nearest to the sink with $h = 0$; those one hop closer to the sink are labeled with $h = 1$, and so forth. Nodes with higher h values are further from the sink. The sink triggers the TDEO process periodically. For instance, in the network depicted in Fig. 2, after topology discovery, node d $(0, 0)$ is assigned $h = 0$. Nodes directly connected to o (m, n) have $h = 1$, and the pattern continues. This approach constructs a directed, loop-free graph based on h, given the lattice topology of the UAN.

During the TDEO phase, gathering critical information about the network's nodes is essential. This includes node distances, clocks, and hop counts. The data collection process requires exchanging control packets among nodes. These control packets have a dual function: topology discovery and time synchronization. Time synchronization technology has been extensively researched and developed in the UAN domain, as highlighted in recent studies [36], [37], [38].

1) Topology Discovery: After network deployment, sensor nodes determine their hop count from the sink to construct the lattice network. The node nearest to the sink initiates the configuration process by sending a HELLO packet. This packet contains the packet type, node ID, packet sequence number, and hop count. It is forwarded hop by hop to the node farthest from the sink. As such, the node set $\boldsymbol{S} = \{\boldsymbol{h_0}, \boldsymbol{h_1}, \ldots, \boldsymbol{h_{m+n}}\}$ of hop can be established for the $m \times n$ lattice network, where $\boldsymbol{h_i}$ represents the set of nodes i-hop away from node $(0, 0)$. For instance, when $m \geq n$

$$\boldsymbol{h}_i = \begin{cases} \{(k, i-k) \text{ for } k = 0, 1, \ldots, i\}, & \text{if } i \leq n \\ \{(k, i-k) \text{ for } k = i-n, \ldots, i\}, & \text{if } n < i \leq m \\ \{(i-m+k, m-k) \text{ for } k = m+n-i, \ldots, 1, 0\}, & \\ & \text{if } m < i \leq m+n. \end{cases} \tag{20}$$

Likewise, when $m < n$

$$\boldsymbol{h}_i = \begin{cases} \{(k, i-k) \text{ for } k = 0, 1, \ldots, i\}, & \text{if } i \leq m \\ \{(k, i-k) \text{ for } k = 0, \ldots, m\}, & \text{if } m < i \leq n \\ \{(i-n+k, n-k) \text{ for } k = m+n-i, \ldots, 1, 0\}, & \\ & \text{if } n < i \leq m+n. \end{cases} \tag{21}$$

If a node belongs to set $\boldsymbol{h_i}$, the hop count of the node is i.

We define the number of nodes in the ith hop as n_i^h. When m and n are known, n_i^h can be acquired by

$$n_i^h = \begin{cases} i+1, & i \leq \min(m, n) \\ \min(m, n)+1, & \min(m, n) < i \leq \max(m, n) \\ m+n-i+1, & \max(m, n) < i \leq m+n. \end{cases} \tag{22}$$

2) Optimizing Energy Consumption: According to the association between distance and link reliability under constant transmission power, once the distance surpasses a certain threshold, P_{link} sharply diminishes to zero. Fig. 3(a) illustrates that to maintain P_{link} at a minimum of 0.9, the required transmission power escalates nonlinearly with respect to distance. Using high transmission power for all the nodes could lead to significant energy wastage. Conversely, using low transmission power may render long-distance links incapable of successful message transmission. As shown in Fig. 3(b), the transmission power increases rapidly to elevate P_{link} above 0.9, where the minimum transmission power is P_{init}. Therefore, P_t should be adjusted across various links to sustain P_{link} within an optimal range, ensuring ETE reliability.

To address the challenge of minimizing energy consumption, we formulate an optimization problem defined by the following equations:

$$\min_{R, r_n, P_t} E_i(R, r_n) \tag{23}$$

$$\text{s.t. } P_{\text{link}}(R, r_n) \geq \tau \tag{23a}$$

$$r_n \in N_+ \tag{23b}$$

$$R_- \leq R \leq R_+ \tag{23c}$$

$$P_t \geq P_{\text{init}}. \tag{23d}$$

This formulation aims to minimize the energy expenditure $E_i(R, r_n)$ for each node i, considering the transmission range R, retransmission count r_n, and transmission power P_t. The constraints ensure that the link reliability probability P_{link},

a function of R and r_n, meets or exceeds a threshold τ, with r_n taking positive integer values. R falls within a specified range $[R_-, R_+]$, and P_t should be no less than an initial (default) transmission power. To solve this formulation, threshold τ will affect ETE reliability, so we first determine the threshold of the link reliability after analyzing the relationship between link reliability and ETE reliability. Then, for a single link, by fixing the retransmission count and incrementally increasing the transmission power while traversing all possible retransmission counts as distance increases, the tradeoff between transmission power and retransmission efforts is addressed through traversal optimization.

The solution is divided into three parts by the communication range, given by

$$\begin{cases} P_t = P_{\text{init}},\ r_n = 1, & \text{if } R_- < R \le R_0 \\ P_t = 10^{\frac{\text{SL}-170.8}{10}},\ r_n = 1, & \text{if } R_0 < R \le R_1 \\ P_t = P_{\text{init}},\ r_n = \lceil \log(1-\tau)/\log(1-P_{\text{link}}) \rceil, & \\ & \text{if } R_1 < R \le R_+ \end{cases} \tag{24}$$

where P_{init} is the initial (default) transmission power, and τ is the threshold of link reliability. R_0 and R_1 are solved by (23). For $R_- < R \le R_0$, the CPR protocol uses a constant P_t with no retransmissions; for $R_0 < R \le R_1$, the CPR protocol adapts P_t with respect to the distance and no retransmissions; if the communication range is between R_1 and R_+, the CPR protocol uses a constant P_t and increases the number of retransmission according to the link reliability requirement. These three configurations are referred to as type 1—constant P_t, no retransmissions, type 2—adaptive P_t, no retransmissions, and type 3—constant P_t with retransmissions, respectively. Note that when a node needs to send a packet to both downstream neighbors, the sender configures the transmission power and retransmission count according to the worse link to ensure the reliability probability is above the given threshold. The corresponding numerical analysis will be discussed in Section V-A.

After the TDEO, all the nodes obtain the distance of the link reliability between their neighbors, the number of hops, and the total number of nodes in its hop n_i^h. Each node knows which hop it belongs to by (20) and (21). Using (12), a node can calculate the link reliability between it and its neighbors. Each node modifies its transmission power and the number of retransmission to maintain link reliability to its neighbors (i.e., the link reliability is above a given threshold). So, the link reliability is further updated by (13). The TDEO will be triggered again by the sink node in the next initialization (after a period of time).

TDEO phase represents the foundational step within routing protocols, a critical component featured in the benchmarks outlined within this article. To effectively evaluate the routing performance, we posit an equivalence in the costs of TDEO phase between the proposed protocol and the benchmark systems. Consequently, the costs incurred during this procedure shall not be expounded upon within the subsequent performance analysis section. For an in-depth exploration of the intricacies of this procedure, we refer readers to the comprehensive analysis presented in [39] and [40].

B. Link Selection

Two metrics are used to optimize the selection of activated nodes. The first metric, ζ evaluates the chosen activated nodes, as defined in [41]

$$\zeta = \frac{N_l}{N_v} \tag{25}$$

where N_l is the number of existed links after some nodes are inactivated, and N_v is the count of activated nodes in the network. If two routes traverse the same nodes, excluding the source and destination, we refer to these nodes as joint nodes. Given the same number of links, networks with more joint nodes offer a higher total number of paths. When comparing the reliability of two networks of identical size, both containing the same number of links, preference generally leans toward the one that has a greater number of joint nodes. Thus, solutions with a larger ζ are chosen to enhance reliability with the same number of links.

Another metric affecting the reliability is the variance in the number of nodes across each hop. The variance formula is

$$\begin{cases} \sigma_1^2 = \dfrac{\sum_0^{h_{\max}} \left(n_i^h - \overline{n}\right)^2}{h_{\max}+1} \\ \sigma_2^2 = \dfrac{\sum_0^{h_{\max}^{\mathrm{T}}} \left(\left(n_i^h\right)^{\mathrm{T}} - \overline{n}^{\mathrm{T}}\right)^2}{h_{\max}^{\mathrm{T}}+1} \end{cases} \tag{26}$$

where $h_{\max}$ is the maximum hop count, n_i^h is the number of nodes in the ith hop set, and $\overline{n}$ is the average number of nodes across all hops. σ_2^2 is the variance along the symmetry direction of σ_1^2, highlighted by orange dashed lines in Fig. 4. Lower variance often leads to higher reliability. Analysis of these metrics, considering different ON/OFF link states, suggests a preference for middle links. For illustrative clarity, combinations where the maximum node count s_a per hop equals 2 are depicted in Fig. 4, excluding symmetrical combinations for simplicity. The blue square represents the four vertexes of the square that have been activated. It shows that the first combination of the blue square has the largest ETE connectivity compared with the second and third combinations. Although these three kinds of combinations have the same value of ζ, the first combination has the lowest σ_1^2 and σ_2^2. Besides, bottleneck nodes shown in the fourth combination can increase the vulnerability of the network, and the two metrics in link selection avoid the existence of bottleneck nodes, as the combination included bottleneck nodes having the largest variance under the same number of activated nodes. The CPR protocol configures the links for message percolation in a network to achieve higher reliability, relying on two metrics ζ and σ^2.

After the TDEO phase, the source node $\text{ID}_o(j)$ generates and sends a message

$$\begin{aligned} \text{message} = \langle &\text{Type}, \text{ID}_{\text{sender}}(j), \text{ID}_o(j), \text{ID}_d(j) \\ &P_{\text{reliab}}^{\text{req}}, \text{seq\_num}, \text{data} \rangle \end{aligned} \tag{27}$$

to its neighbors. Type is the type of packet, and $\text{ID}_{\text{sender}(j)}$, $\text{ID}_o(j)$, and $\text{ID}_d(j)$ are the address of the sender node,

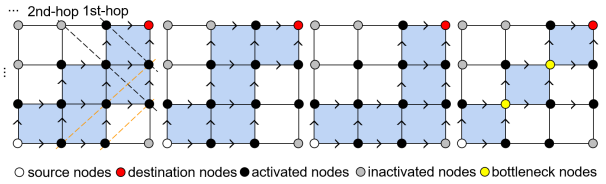


Fig. 4. Illustration of two metrics in the link selection.

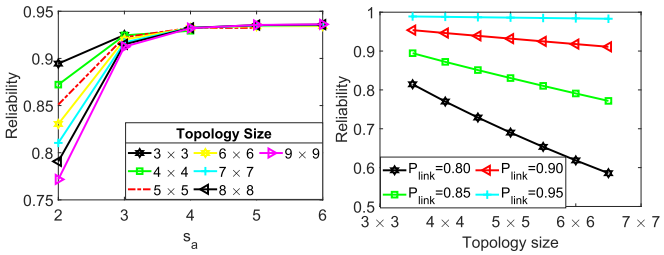


Fig. 5. Impact on reliability as topology size changes: (a) caused by different s_a values and (b) caused by different P_{link} values.

source node o, and destination node d for the jth transmission, respectively. $P^{\text{req}}_{\text{reliab}}$ represents the given ETE reliability demand specified by users.

Given a hop-based topology, when a node receives a message, it performs a link selection procedure to compute the activated node set $\mathcal{N}_a$ according to the requirement of ETE reliability $P^{\text{req}}_{\text{reliab}}$, the source node ID, and the destination node ID. If a node belongs to $\mathcal{N}_a$, it is active to forward the message, so its status will be set as 1; otherwise, its status $= 0$, and it will not forward the message. To accomplish the requirement of ETE reliability, the activated nodes need to satisfy the relationship among s_a, P_{link}, and $P^{\text{req}}_{\text{reliab}}$. The mapping relationships between s_a and P_{link} will be shown later in Fig. 5, which is stored at each node to support the autonomic calculation of $\mathcal{N}_a$. When $m \geq n$, the set of activated node is given by

$$\mathcal{N}_a = \begin{cases} \boldsymbol{h_i}, & \text{if } n_i^h < s_a \\ \left\{ \boldsymbol{h_i}[k] \text{ for } k = \left\lfloor \frac{n_i^h - s_a}{2} \right\rfloor + 1, \ldots, \left\lfloor \frac{n_i^h - s_a}{2} \right\rfloor + s_a \right\}, & \\ \quad \text{if } n_i^h \geq s_a, \left(n_i^h - s_a\right) \bmod 2 > 0 & \\ \left\{ \boldsymbol{h_i}[k] \text{ for } k = \frac{n_i^h - s_a}{2}, \ldots, \frac{n_i^h - s_a}{2} + s_a - 1 \right\}, & \\ \quad \text{if } n_i^h \geq s_a, \left(n_i^h - s_a\right) \bmod 2 = 0 & \end{cases} \tag{28}$$

where i is from 0 to $m + n$, and $\boldsymbol{h_i}[k]$ is a collection of nodes with index k in i-hop. Therefore, given $P^{\text{req}}_{\text{reliab}}$, the CPR protocol can build the arrangement pattern that satisfies the reliability requirement using activated nodes' set $\mathcal{N}_a$.

Algorithm 1 is the link selection and message forwarding used by CPR to determine activated nodes and handle received messages. Source node $\text{ID}_o(j)$ generates a message that will be forwarded to destination node $\text{ID}_d(j)$. Node q decides whether it is in the active node set for each received message in Lines 1–6 based on the TDEO information. According to

Algorithm 1 Link Selection and Message Forwarding

Data: *message* from the sender
Result: Active node set $\mathcal{N}_a$ and message forwarding process

1 **while** *node q receives a message* m_j **do**
2 upon the source node $ID_o(j)$ and $ID_d(j)$, node q chooses s_a according to the stored mapping relationship $(f : s_a, P_{link} \rightarrow P^{req}_{reliab})$;
3 computes $\mathcal{N}_a$ using 28;
4 **while** *node* $q \in \mathcal{N}_a$ **do**
5 set $status_j(q) = 1$;
6 **end**
7 **if** $status_j(q) == 1$ **then**
8 **if** *This message* m_j *is the first time received by node q* **then**
9 **if** *node q is the destination node* **then**
10 receive the message, no forwarding;
11 end the procedure;
12 **end**
13 **if** *node q is not the destination node* **then**
14 forward the message to downstream neighbors in $\mathcal{N}_a$, save *seq_num*;
15 **end**
16 **end**
17 **if** *This message has been received by node q* **then**
18 discard the message;
19 **end**
20 **else**
21 discard the message;
22 **end**
23 **end**

the mapping relationship among s_a, P_{link} and $P^{\text{req}}_{\text{reliab}}$, node q determines s_a upon source node $\text{ID}_o(j)$ and destination node $\text{ID}_d(j)$. If the ETE reliability can satisfy $P^{\text{req}}_{\text{reliab}}$, the minimum s_a is chosen. After link selection, the node knows its $\text{status}_j(q)$ in Lines 4 and 6. If its $\text{status}_j(q)$ is 1, it is an activated node; otherwise, it is inactivated node. According to Lines 7–22, the nodes whose $\text{status}_j(q) = 1$ will forward m_j if m_j is a new message and keep the message sequence information for a while to check whether the same message will arrive in the future; otherwise, the nodes will discard the message. In this way, the message can be percolated among the activated nodes until it arrives at the destination.

In the procedure of CPR, the packet is delivered by activated nodes, which have different transmission power and the number of retransmissions, through the confined routing paths in a percolation way.

C. Medium Access Control

CPR is used to efficiently manage transmission power and minimize the number of retransmissions, thus conserving energy. It leverages link quality metrics for estimating the optimal transmission power and the appropriate number of retransmissions.

TABLE II
PARTICULAR PARAMETERS [9], [44], [45]

Parameter	Value	Parameter	Value
Spreading factor	2	f	22.25 kHz
Initial SL	177.3 dB	α	4.5485 dB/km
Modulation method	FSK	B	20 kHz
P_{init}	4.5 W	R_b	500 bps
The length of packet	256 bits	P_r	0.5 W
The number of nodes	4×4-9×9	s_a	2
D_c (m)	600-800	τ	0.85

In the context of percolation routing, a scheduled TDMA-based medium access control protocol is used as a proactive measure to mitigate and prevent packet collisions. This strategic approach is used to enhance the overall reliability and efficiency of data transmission, reducing the likelihood of interference and collisions among packets during routing.

Given the lattice network topology of the UANs comprising sea floor with acoustic transceivers, there are two categories of interference for the receiver: the interference comes from the previous hop and the interference comes from the rest of the nodes. For the first type of interference, the zigzag decoding technique can decode collided packets by interference cancellation. For the second type of interference, using the time slot schedule can avoid collisions.

We assume packet length P_{size} is constant. To avoid interference and collisions, time is slotted and the duration of a time slot is t_{slot}, which is sufficient to transmit packets with the number of the maximum re-transmissions. T time slots constitute a time period. Nodes in the same hop count will be assigned one slot in each period to transmit. h_{total} is the number of hops from the source to the destination, and define $l_p = \lceil h_{\text{total}}/T \rceil$ as the number of periods needed from ETE. In addition, the same slot will be reused by nodes $(wT - 1)$ hops away, where $w = 1, 2, \ldots, l_p$. Thus, nodes in $N(h)$ are assigned the $(h \bmod l_p)$th time slot in each period to transmit. Therefore, we can appropriately select T for collision-free communication given the lattice topology.

V. PERFORMANCE EVALUATION

We implement and instantiate our proposed mathematical model in Sections III-B–III-C, as well as the peculiar characteristics of UANs using Python. The NetworkX library [42] is used for creating and managing the network topology and node interactions, and the SimPy library [43] is used for process-based discrete-event simulation. We consider the system model in Section III-A, where an arbitrary cube cell has the side length of D_c changed from 600 to 800 m. The number of sensor nodes ranges from 9 to 81. Each node is randomly deployed in one cube cell. Packets are generated by the source node at a constant rate and sent toward the destination. The parameter settings are shown in Table II.

For routing in UAN scenarios, energy consumption and ETE reliability are pivotal concerns. Section II discusses the existing protocols oriented toward energy and reliability. These protocols fall into five categories: flooding-based, shortest path, energy-aware, OR, and cluster-based protocols. We selected a representative benchmark protocol from each category for comparison: Dflooding [46], Multi-SPR [47], EEGNBR [20], GEDAR [23], and LEACH. These five protocols were compared with the proposed CPR protocol regarding energy efficiency and reliability.

Flooding-based protocols, such as Dflooding, aim to maximize reliability but at a high energy consumption. For fair evaluation, Dflooding strategy was configured to relay each packet to downstream neighbors once. Multi-SPR searches for multiple shortest paths to minimize energy consumption, selecting paths with the minimum number of shared links. To ensure a fair comparison, Multi-SPR and the proposed CPR were configured to use an equal number of active nodes for packet forwarding. EEGNBR, following node deployment, establishes a guiding network and uses distance vectors to prioritize candidates. The nodes with the minimum number of hop counts toward the sinks have the largest priority. GEDAR, an OR protocol, considers packet delivery probability to enhance routing decisions. LEACH, a cluster-based approach, reduces energy cost by creating and maintaining clusters, thereby extending the lifetime of network.

Three groups of experiments are designed in this article. First, we evaluate the effect of design parameters s_a and P_{link} on network ETE reliability based on the ideal homogeneous networks. We investigate the impact of the tradeoff between transmitting power and retransmitting in the performance of our proposed CPR routing protocol. These energy parameters are used to guide the following heterogeneous lattice network. Second, we further design heterogeneous networks and evaluate the performance of the proposed routing compared with five benchmarks, focusing on energy consumption, reliability, and delay. Third, the five benchmark protocols and CPR are compared using actual channel data to investigate their energy consumption and reliability in realistic harsh underwater acoustic environments.

A. Homogeneous Lattice Network

To determine the value of τ, we use a homogeneous network in this context, where the link success probability is equal for all links and is denoted as P_{link}. The situation of activated nodes for the homogeneous network is depicted in Fig. 4, with a certain value of the maximum number of activated nodes s_a for each hop. Given s_a for each hop, higher ETE reliability can be achieved when the activated link is closer to the line connecting the source node to the destination node.

According to the ETE reliability calculation model, we can establish a relationship between s_a for each hop and the network's ETE reliability, considering a value of P_{link}. When $P_{\text{link}} = 0.85$, we analyze the maximum network ETE reliability while varying s_a for each hop from two to six, as shown in Fig. 5(a). When s_a for each hop in the network is set to 2wo, the network's reliability decreases from 0.89 with 3×3 topology size to 0.77 with 9×9 as the network size increases. However, when s_a for each hop is set to 3, the reliability improves significantly, ranging from 0.92 to 0.91 for network sizes of 3×3 to 9×9. When s_a for each hop is set to 4, the network's reliability does not change much

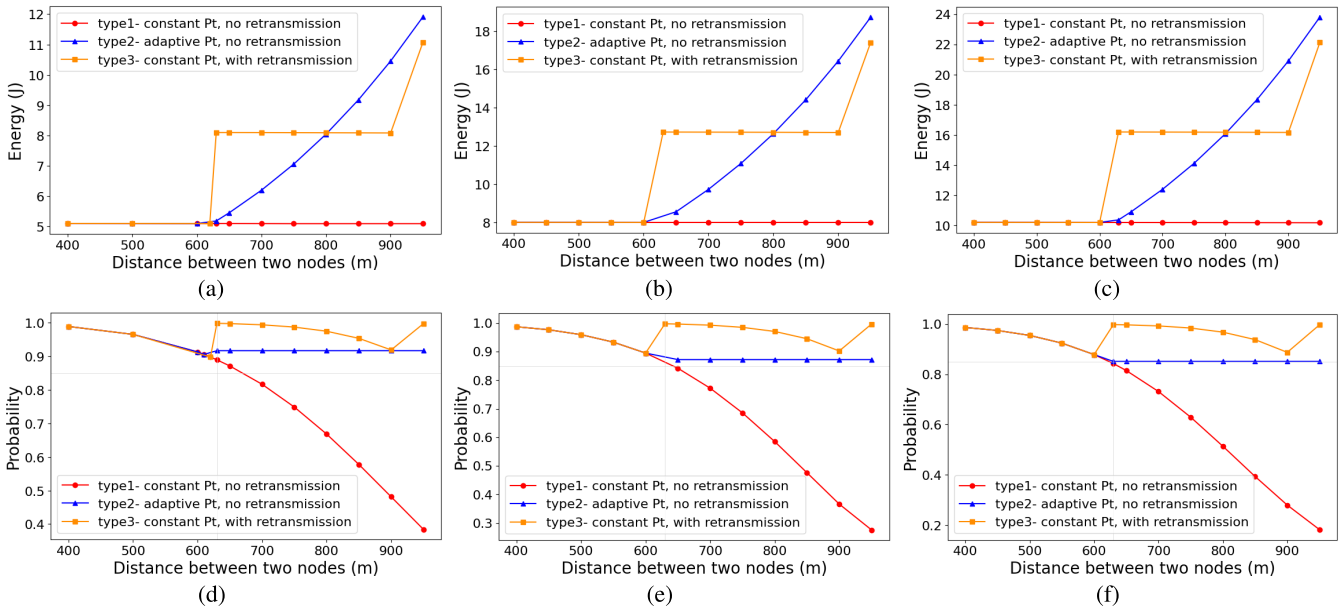


Fig. 6. Network average energy consumption and ETE reliability with different lattices versus per-hop distance. (a) Energy consumption (3 × 3). (b) Energy consumption (4 × 4). (c) Energy consumption (5 × 5). (d) Reliability with 3 × 3 lattice. (e) Reliability with 4 × 4 lattice. (f) Reliability with 5 × 5 lattice.

with the increase in network size, and all the values exceed 0.9. Given the limited energy supply for UANs, we aim to use as few links as possible while meeting the reliability requirements. Therefore, we consider the case where s_a for each hop is 2 in the homogeneous network scenario. Following this, we conduct further analysis to determine the suitable value for τ by examining the relationship between P_{link} and ETE reliability.

We calculate the impact of changes in P_{link} on network ETE reliability when s_a for each hop is set to 2. As the value of P_{link} decreases, the network reliability decreases more significantly. In Fig. 5(b), when $P_{\text{link}} = 0.8$, network reliability drops from approximately 0.81 in a 3 × 3 network to around 0.59 in a 9 × 9 network. When $P_{\text{link}} = 0.85$, network reliability is 0.77 for a 9 × 9 network. However, for $P_s = 0.9$, network reliability ranges from 0.95 to 0.9 as the network size increases from 3 × 3 to 9 × 9. When $P_{\text{link}} = 0.95$, network reliability consistently remains above 0.98.

Considering that the ETE reliability threshold $P_{\text{reliab}}^{\text{req}}$ is set above 0.9, when $P_{\text{link}} = 0.85$, and s_a for each hop is constrained to 2, it becomes apparent that the theoretical maximum network ETE reliability ranges from 0.89 to 0.77. Within this context, there exists greater flexibility in selecting the states of network links to improve reliability and manage remaining energy. Therefore, we impose the constraint $\tau = 0.85$ as the required success probability for individual links. Node energy control is carried out based on this requirement, striking a reasonable balance between node transmission power and transmission times.

Fig. 6 shows the results concerning energy consumption and reliability with an optimized power adjustment strategy. Power control and retransmission are effective ways to improve per-hop reliability. The transmission power P_t remains constant for type 1, shown by the red line with circles. For type 2, P_t increases with distance by the blue line with triangles. For type 3, shown by the orange line with squares, the transmission power is fixed, but the number of retransmissions increases to ensure a link probability over the threshold of 0.85. As shown in Fig. 6(a)–(c), the total energy cost of type 1 remains constant while its reliability drops quickly; that of type 2 exponentially grows with respect to distance to maintain the reliability of 0.9; that of type 3 piecewisely increases with the reliability varying between 0.9 and 0.99.

The reliability trend of the three types for different topologies is similar when n is changed from 3 to 5 as shown in Fig. 6(d)–(f). The results have shown that energy consumption rises with the increased network size. Types 2 and 3 can ensure link reliability with different energy costs depending on the distance. Meanwhile, when n is increased from 3 to 5 if s_a remains the same, ETE reliability decreases as n increases. In other words, s_a can be tuned to satisfy the reliability requirement for different values of n.

From the above results, to ensure the reliability with reduced energy, we can use the optimal transmission strategy. When the distance is below $R_0 = 630$ m, the transmission power is a constant value and no retransmission; when the distance is between R_0 m and $R_1 = 800$ m, the transmission power is increased with the distance and no retransmission; when the distance is greater than R_1, the transmission power remains constant, and the max number of retransmission is increased by one.

B. Heterogeneous Lattice Network

We simulate $n \times n$ heterogeneous lattice networks, where n varies from 3 to 9. The topology is generated with one node randomly deployed in a cube cell in each run where each node has random movement at 0.5 m/s in its cube cell by the way of drunkard random walk. For heterogeneous

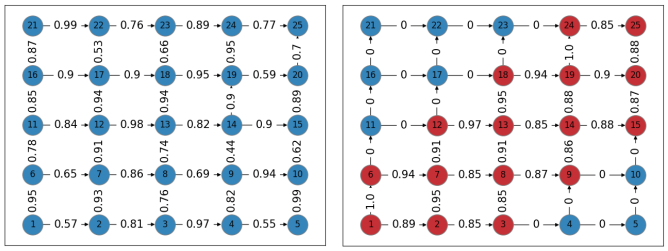


Fig. 7. Heterogeneous lattice network topology (5 × 5). (a) P_{link} after the TDEO. (b) P_{link} after the CPR.

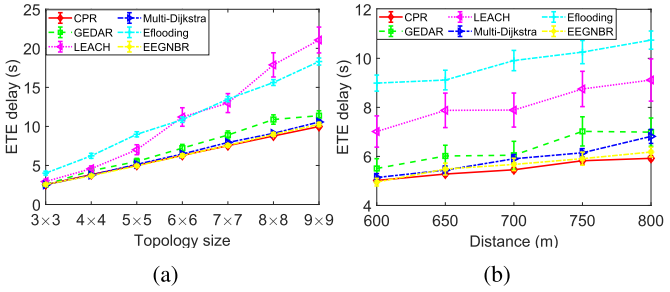


Fig. 9. Average ETE latency. (a) Average ETE delay versus network size. (b) Average ETE delay versus distance D_c.

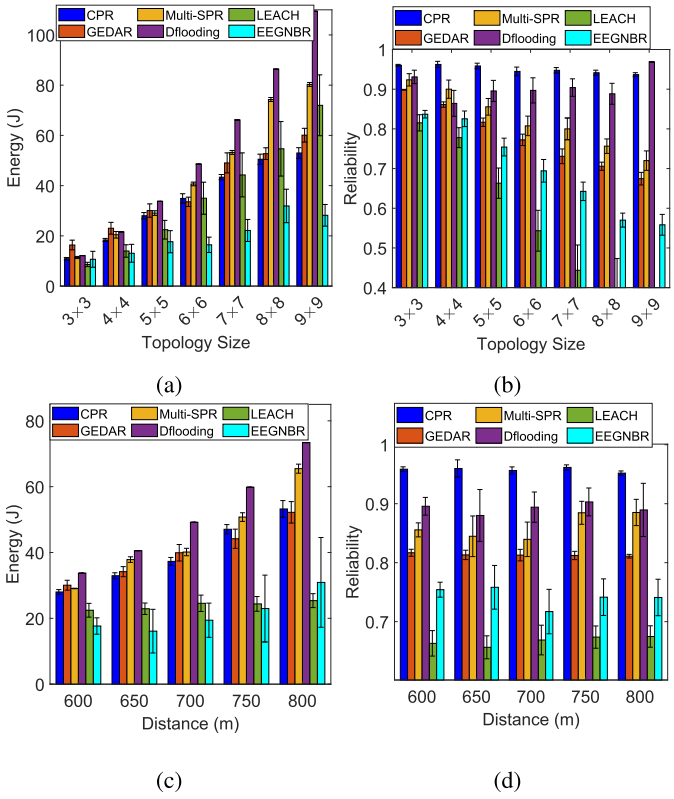


Fig. 8. Numerical results. (a) and (b) Average energy consumption and reliability in different sizes of network. (c) and (d) Average energy and reliability with different distances D_c.

networks, the link reliability of each link is different. For instance, a random topology with $n = 5$ is deployed in an area. In the network TDEO phase, nodes discover their neighbors and obtain the corresponding link connectivity. By exchanging these link information, each node learns the whole network graph. Fig. 7(a) shows the initial link probability among nodes after the TDEO phase. Then, after the link selection stage, some links are activated when s_a is set to 2, as shown in Fig. 7(b). According to the link selection scheme, the red nodes are active, and the blue nodes are inactivated. Using the CPR protocol with optimal power/retransmission configurations discussed in Section V-A, the routing path and the link probability are updated as shown in Fig. 7(b).

Fig. 8 shows the average energy consumption and reliability with different network sizes and single-hop distances with its 95% confidence interval.

For different topology cases, we take a random topology with n varying from 3 to 9. For each setting, we repeated 50 tests with random node deployments. We use the average energy and reliability to compare the performance of each protocol, as shown in Fig. 8(a) and (b). Obviously, the energy cost increases with the increase in the network size. When the size of the network is 3 × 3 , the energy consumption using the six protocols is similar, and CPR is the most energy-efficient. When n increases, the performance gap becomes more apparent. The CPR achieves the highest ETE reliability and remains stable at about 0.95. The reliability with Dflooding decreases when the network size increases, and the reliability decrement of the LEACH is the fastest. Meanwhile, CPR can use less energy than the Dflooding to achieve higher reliability, and the gain increases with respect to network size. As the distribution of underwater sensor nodes is not even in most cases, the reliability performance of Multi-SPR, GEDAR, LEACH, and EEGNBR is not satisfactory, sometimes below 0.75. Among the four, Multi-SPR performs the best, as it adjusts the transmission power and retransmission for improving link reliability. LEACH consumed slightly less energy than CPR for small-size networks and the trend is reversed when the network size is large. However, the reliability with LEACH is the worst for all the cases, below 0.85 for 3 × 3 networks, and drop to 0.4 for 9 × 9 networks.

To investigate the impact of D_c, we simulate the performance of six routing protocols with varying D_c from 600 to 800 m for 5 × 5 network size, as shown in Fig. 8(c) and (d). The increase in D_c means that the average distance between nodes in the network increases, so the transmission power needs to be increased to achieve the corresponding link probability. As the distance increases, the energy consumption of all the five protocols increases, while the reliability of LEACH drops the fastest. The reliability of CPR is the highest, around 0.95.

Fig. 9 presents the ETE delay across varying topology sizes for different routing protocols with its 95% confidence interval. As the topology size increases from 3 × 3 to 9 × 9 in Fig. 9(a), the ETE delay trends upward for all the protocols due to the extended paths that packets must traverse. The CPR shows a moderate increase in ETE delay, suggesting efficient packet forwarding mechanisms. The slope of the delay curve for CPR is relatively flat compared with others, which indicates better scalability with increasing network size. LEACH, with its cluster-based routing strategy, exhibits a

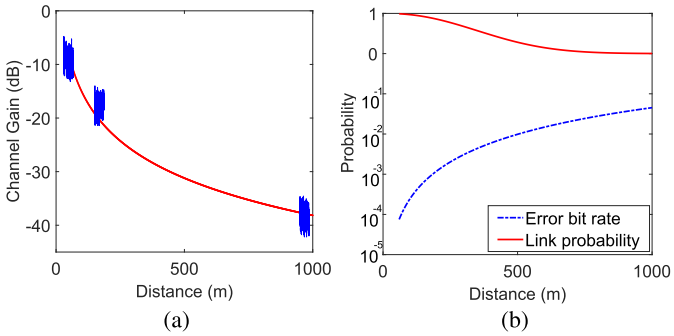


Fig. 10. Results of experimental analysis. (a) Gain versus distance and the fit model. (b) Error bit ratio and the link probability.

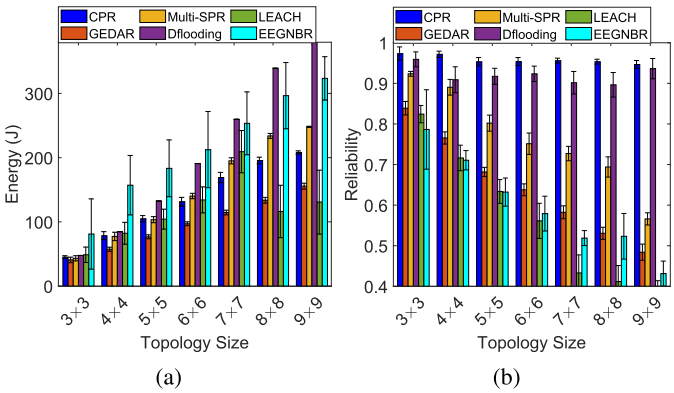


Fig. 11. Performance comparison in different sizes of network using the realistic channel data. (a) Average energy consumption. (b) Average ETE reliability.

steeper rise in delay, particularly notable in larger topologies. This indicates potential inefficiencies in cluster-head communication or increased intracluster distances as the network expands. Dflooding demonstrates a high delay that increases significantly with the network size, a consequence of its exhaustive broadcasting approach that, while robust, is not energy-efficient. The protocols Multi-SPR and EEGNBR, while still experiencing an increase in delay with topology size, maintain lower delays compared with Dflooding and LEACH, showing better scalability and possibly more effective path selection strategies. GEDAR, which uses OR considering packet delivery probabilities, maintains a relatively stable delay across different sizes, indicating effective management of dynamic routing decisions.

Fig. 9(b) shows the results concerning the average delay with different density for 5×5 networks. The CPR shows a progressive, yet modest, increase in ETE delay as distance extends, indicating effective management of delay even over longer hops. The relatively stable growth pattern of CPR's delay curve suggests that its routing strategy copes well with increased single-hop distances. The LEACH protocol demonstrates a more significant delay increase, indicative of potential inefficiencies in cluster communication over longer distances. This may be due to the extended range between cluster heads and cluster members or the sink, which becomes more pronounced with distance. Dflooding exhibits an upward trend in ETE delay consistent with its energy-intensive nature, where each node's transmission contributes to the cumulative delay. This trend is expected, given the protocol's design, which does not consider distance directly. Both GEDAR and Multi-SPR protocols show intermediate increases in delay, indicating better scalability with distance than LEACH and Dflooding but less so than CPR.

C. Experimental Results

We investigate the feasibility of CPR for UANs using data recorded during the Surface Processes and Acoustic Communications Experiment (SPACE08) [48], which was conducted near the coast of Martha's Vineyard in MA, USA, in fall 2008. The channel gain and the fit model are shown in Fig. 10(a). According to the experiment fit channel gain model, the link probability and the error bit ratio can be obtained in Fig. 10(b). Based on the obtained link probability, CPR sets the transmission power and the number of retransmission and calculates the ETE reliability.

The proposed algorithm is evaluated using the above channel datasets. Fig. 11(a) and (b) shows the energy cost per successful delivered packet and the ETE reliability relative to the network size with 95% confidence interval. As expected, the Dflooding exhibited a robust reliability performance due to its redundant packet transmissions. Therefore, Dflooding consumed more energy than the other schemes (even worse with increased network size). CPR exhibited the highest and stable reliability due to the link selection and optimal energy scheme.

VI. Conclusion

In this article, we proposed a routing scheme designed for UANs considering the characteristics of underwater communications. Given the physical layer configurations, the link probability is obtained, which is applied to calculate the network connectivity. The link selection rule and energy optimization are analyzed based on the reliability analysis model of UANs. This insight was then used to design a CPR protocol. Performance evaluation was carried out first in simulated homogeneous and heterogeneous networks and subsequently in a realistic network scenario where the realistic channel data were used. The results confirmed that the proposed routing scheme achieves much larger ETE reliability and energy efficiency by exploring path diversity and optimizing power/retransmission settings.

There are some open issues we need to address in future research. First, how to consider the impact of different MAC layer protocols and node mobility on link selection for heterogeneous UANs. How to apply a flexible and adaptive link selection scheme to avoid traffic congestion and address the energy balance problem needs to be further investigated. In addition, we can extend our work by applying the related other noise models for link performance analysis and estimation, and then use the results to optimize routing in our future work.

References

[1] Y. Liu, H. Wang, L. Cai, and X. Shen, "An efficient percolation-based routing protocol for underwater acoustic networks," in *Proc. OCEANS, San Diego Porto*, Sep. 2021, pp. 1–4.

[2] H. Khan, S. A. Hassan, and H. Jung, "On underwater wireless sensor networks routing protocols: A review," *IEEE Sensors J.*, vol. 20, no. 18, pp. 10371–10386, Sep. 2020.

[3] Y. Liu, H. Wang, L. Cai, X. Shen, and R. Zhao, "Fundamentals and advancements of topology discovery in underwater acoustic sensor networks: A review," *IEEE Sensors J.*, vol. 21, no. 19, pp. 21159–21174, Oct. 2021.

[4] A. Al Guqhaiman, O. Akanbi, A. Aljaedi, and C. E. Chow, "A survey on MAC protocol approaches for underwater wireless sensor networks," *IEEE Sensors J.*, vol. 21, no. 3, pp. 3916–3932, Feb. 2021.

[5] Y. Li, Y. Zhang, H. Zhou, and T. Jiang, "To relay or not to relay: Open distance and optimal deployment for linear underwater acoustic networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3797–3808, Sep. 2018.

[6] R. Petroccia et al., "Deployment of a persistent underwater acoustic sensor network: The CommsNet17 experience," in *Proc. OCEANS MTS/IEEE Kobe Techno-Oceans (OTO)*, May 2018, pp. 1–9.

[7] S. Sendra, J. Lloret, J. M. Jimenez, and L. Parra, "Underwater acoustic modems," *IEEE Sensors J.*, vol. 16, no. 11, pp. 4063–4071, Jun. 2016.

[8] A. Khan et al., "Modem design for underwater acoustic networks: Taxonomy, capabilities, challenges, applications and future trends," *J. Intell. Fuzzy Syst.*, vol. 39, no. 6, pp. 8161–8171, Dec. 2020.

[9] M. Y. I. Zia, J. Poncela, and P. Otero, "State-of-the-art underwater acoustic communication modems: Classifications, analyses and design challenges," *Wireless Pers. Commun.*, vol. 116, no. 2, pp. 1325–1360, Jan. 2021.

[10] L. Freitag, M. Johnson, and D. Frye, "High-rate acoustic communications for ocean observatories-performance testing over a 3000 m vertical path," in *Proc. MTS/IEEE Conf. Exhib. Conf.*, Sep. 2000, pp. 1443–1448.

[11] D. Pompili, T. Melodia, and I. F. Akyildiz, "Distributed routing algorithms for underwater acoustic sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2934–2944, Sep. 2010.

[12] J. Hu, L. Cai, C. Zhao, and J. Pan, "Directed percolation routing for ultra-reliable and low-latency services in low earth orbit (LEO) satellite networks," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, Nov. 2020, pp. 1–6.

[13] Z. Zhou, B. Yao, R. Xing, L. Shu, and S. Bu, "E-CARP: An energy efficient routing protocol for UWSNs in the Internet of Underwater Things," *IEEE Sensors J.*, vol. 16, no. 11, pp. 4072–4082, Jun. 2016.

[14] Y. Zhou, T. Cao, and W. Xiang, "Anypath routing protocol design via Q-learning for underwater sensor networks," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8173–8190, May 2021.

[15] S. Gopi, K. Govindan, D. Chander, U. B. Desai, and S. N. Merchant, "E-PULRP: Energy optimized path unaware layered routing protocol for underwater sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3391–3401, Nov. 2010.

[16] F. Bouabdallah, C. Zidi, and R. Boutaba, "Joint routing and energy management in UnderWater acoustic sensor networks," *IEEE Trans. Netw. Service Manage.*, vol. 14, no. 2, pp. 456–471, Jun. 2017.

[17] G. Xing et al., "Game-theory-based clustering scheme for energy balancing in underwater acoustic sensor networks," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9005–9013, Jun. 2021.

[18] Y. Yuan, M. Liu, X. Zhuo, Y. Wei, X. Tu, and F. Qu, "A Q-learning-based hierarchical routing protocol with unequal clustering for underwater acoustic sensor networks," *IEEE Sensors J.*, vol. 23, no. 6, pp. 6312–6325, Mar. 2023.

[19] L. Liu, M. Ma, C. Liu, and Y. Shu, "Optimal relay node placement and flow allocation in underwater acoustic sensor networks," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2141–2152, May 2017.

[20] Z. Liu, X. Jin, Y. Yang, K. Ma, and X. Guan, "Energy-efficient guiding-network-based routing for underwater wireless sensor networks," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21702–21711, Nov. 2022.

[21] J. M. Jornet, M. Stojanovic, and M. Zorzi, "Focused beam routing protocol for underwater acoustic networks," in *Proc. 3rd ACM Int. Workshop Underwater Netw.*, 2008, pp. 75–82.

[22] M. Zhang and W. Cai, "Energy-efficient depth based probabilistic routing within 2-Hop neighborhood for underwater sensor networks," *IEEE Sensors Lett.*, vol. 4, no. 6, pp. 1–4, Jun. 2020.

[23] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Loureiro, "Geographic and opportunistic routing for underwater sensor networks," *IEEE Trans. Comput.*, vol. 65, no. 2, pp. 548–561, Feb. 2016.

[24] H. Yan, Z. Shi, and J. Cui, "DBR: Depth-based routing for underwater sensor networks," in *Proc. 7th Int. IFIP-TC6 Netw. Conf. Ad Hoc Sensor Netw., Wireless Netw., Next Gener. Internet*, Singapore. Berlin, Germany: Springer, May 2008, pp. 72–86.

[25] P. Xie, J. Cui, and L. Lao, "VBF: Vector-based forwarding protocol for underwater sensor networks," in *Proc. 5th Int. IFIP-TC6 Netw. Conf. Netw. Technol., Services, Protocols, Perform. Comput. Commun. Netw., Mobile Wireless Commun. Syst.*, Coimbra, Portugal. Berlin, Germany: Springer, May 2006, pp. 1216–1221.

[26] N. Nicolaou, A. See, P. Xie, J.-H. Cui, and D. Maggiorini, "Improving the robustness of location-based routing for underwater sensor networks," in *Proc. OCEANS Eur.*, Jun. 2007, pp. 1–6.

[27] Q. Guan, F. Ji, Y. Liu, H. Yu, and W. Chen, "Distance-vector-based opportunistic routing for underwater acoustic sensor networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3831–3839, Apr. 2019.

[28] D. Hwang and D. Kim, "DFR: Directional flooding-based routing protocol for underwater sensor networks," in *Proc. OCEANS*, Sep. 2008, pp. 1–7.

[29] Y. Zhang, Z. Zhang, L. Chen, and X. Wang, "Reinforcement learning-based opportunistic routing protocol for underwater acoustic sensor networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2756–2770, Mar. 2021.

[30] R. Zhu, Q. Jiang, X. Huang, D. Li, and Q. Yang, "A reinforcement-learning-based opportunistic routing protocol for energy-efficient and void-avoided UASNs," *IEEE Sensors J.*, vol. 22, no. 13, pp. 13589–13601, Jul. 2022.

[31] C. Wang, X. Shen, H. Wang, H. Zhang, and H. Mei, "Reinforcement learning-based opportunistic routing protocol using depth information for energy-efficient underwater wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 15, pp. 17771–17783, Aug. 2023.

[32] Y. Noh, U. Lee, P. Wang, B. S. C. Choi, and M. Gerla, "VAPR: Void-aware pressure routing for underwater sensor networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 895–908, May 2013.

[33] Y. Noh et al., "HydroCast: Pressure routing for underwater sensor networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 333–347, Jan. 2016.

[34] R. Urick, *Principles of Underwater Sound*. New York, NY, USA: McGraw-Hill, 1983.

[35] L. Zhang, L. Cai, J. Pan, and F. Tong, "A new approach to the directed connectivity in two-dimensional lattice networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 11, pp. 2458–2472, Nov. 2014.

[36] J. Liu, Z. Wang, M. Zuba, Z. Peng, J.-H. Cui, and S. Zhou, "DA-sync: A Doppler-assisted time-synchronization scheme for mobile underwater sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 3, pp. 582–595, Mar. 2014.

[37] O. Pallares, P.-J. Bouvet, and J. del Rio, "TS-MUWSN: Time synchronization for mobile underwater sensor networks," *IEEE J. Ocean. Eng.*, vol. 41, no. 4, pp. 763–775, Oct. 2016.

[38] Y. Chen, Z. Jin, Q. Zeng, and Q. Yang, "A collision-avoided MAC protocol with time synchronization and power control for underwater sensor networks," *IEEE Sensors J.*, vol. 22, no. 19, pp. 19073–19087, Oct. 2022.

[39] R. Zhao, Y. Liu, X. Shen, S. Ma, J. Chang, and X. Wang, "An efficient topology discovery protocol for underwater acoustic networks," in *Proc. OCEANS Marseille*, Jun. 2019, pp. 1–4.

[40] R. Zhao, Y. Liu, O. A. Dobre, H. Wang, and X. Shen, "An efficient topology discovery protocol with node ID assignment based on layered model for underwater acoustic networks," *Sensors*, vol. 20, no. 22, p. 6601, Nov. 2020.

[41] J. Hu, L. Cai, and J. Pan, "Mesh network reliability analysis for ultra-reliable low-latency services," in *Proc. IEEE MASS*, Oct. 2021, pp. 198–206.

[42] A. Hagberg, P. J. Swart, and D. A. Schult. *Networkx*. Accessed: Jun. 12, 2024. [Online]. Available: https://networkx.github.io

[43] N. Matloff. *Simpy: Discrete Event Simulation for Python*. Accessed: Jun. 12, 2024. [Online]. Available: https://simpy.readthedocs.io

[44] A. Waite, *Sonar for Practising Engineers*. Hoboken, NJ, USA: Wiley, 2002.

[45] R. Petroccia, "DIVE: A distributed ID assignment and topology discovery protocol for underwater acoustic networks," *Ad Hoc Netw.*, vol. 122, Nov. 2021, Art. no. 102610.

[46] R. Otnes and S. Haavik, "Duplicate reduction with adaptive backoff for a flooding-based underwater network protocol," in *Proc. MTS/IEEE OCEANS Bergen*, Jun. 2013, pp. 1–6.

[47] R. M. Gomathi and J. M. L. Manickam, "Energy efficient shortest path routing protocol for underwater acoustic wireless sensor network," *Wireless Pers. Commun.*, vol. 98, no. 1, pp. 843–856, Jan. 2018.

[48] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, Oct. 2013.