An Overview Survey of Recent Routing Protocols for Underwater Wireless Sensor Networks

Faiza A. Al-Salti *, Nasser M. Al-Zeidi, Khaled Day, Bassel Arafeh and Abderezak Touzene

Department of Computer Science, College of Science, Sultan Qaboos University, P.O. Box 123, Al-Khod, Muscat, Sultanate of Oman. *Email: s81851@student.squ.edu.om

ABSTRACT: With the recent advances in underwater sensor devices and technologies, underwater wireless sensor networks (UWSNs) enable a variety of applications such as underwater exploration and monitoring, disaster prevention, and military surveillance and reconnaissance. However, these kinds of networks face a number of challenges induced by the nature of the underwater environment and its influence on the physical media. Therefore, new routing protocols are proposed specifically for such networks to mitigate these challenges. This paper surveys some of the recent routing protocols for UWSNs. Specifically, the idea of each protocol is presented as well as its advantages and disadvantages. Furthermore, the presented protocols are classified into different categories. The paper is concluded with some open research issues.

Keywords: Underwater wireless sensor networks (UWSNs); Acoustic channels; Routing protocols; Location-based routing; Location-free routing.

1. Introduction

One oceans, seas, rivers and lakes cover nearly three quarters of the Earth’s surface. They provide a prime source for nourishment, trade and commerce. In addition, oceans help in producing the world’s oxygen and absorbing carbon dioxide. With the increasing role of the oceans and seas in human life, there is a strong demand to explore and investigate the unexplored regions to make use of their treasures. However, due to the harsh and high-pressure underwater environment, unmanned technologies have become vital to deep-sea exploration and coastal monitoring. Underwater Wireless Sensor Networks (UWSNs) are considered to be one of the promising candidates for investigating the oceans. This kind of networks consists of a collection of sensors deployed in an environment to perform a collaborative task. A sensor node is a micro-electronic device having the capability to detect, measure and transmit a variety of parameters such as temperature, salinity and pressure [1]. UWSNs are envisioned to enable various applications including scientific, military and commercial applications.
For example, they can be used for coral reef monitoring [2], military surveillance and reconnaissance [3], oil drilling and exploration [4], and underwater pipeline monitoring [5][6].

The characteristics of terrestrial wireless sensor networks (WSNs) and UWSNs are different. Thus, the protocols, standards and technologies used in terrestrial WSNs cannot be applied directly for UWSNs. The main difference is that acoustic media is the most acceptable physical layer technology in UWSNs [7]. Other counterparts such as electromagnetic and optical signals cannot be exploited efficiently in the underwater environment. This is because electromagnetic waves suffer from high attenuation and absorption, and hence, propagate for short range [8]. Optical signals, on the other hand, suffer from scattering, absorption, and very short-range propagation [8] [7].

Nevertheless the acoustic signal is the one suitable medium that works satisfactorily in the underwater environment; it is considered one of the toughest communication media in use today [9]. Generally, underwater wireless sensor networks (UWSNs) experience a number of challenges induced by the nature of the environment and the transmission media. Particularly, acoustic speed in water is five orders of magnitude slower than the speed of radio signals in the terrestrial environment [10]. In addition, underwater channels are affected by multi-path and fading. Underwater sensor nodes are equipped with batteries, which are difficult to replace or recharge. The available bandwidth is limited and inversely proportional to the transmission distance [11] [12]. In addition, the three-dimensional (3D) deployment and the dynamic environment due to the passive movements of sensors with water currents cause extra challenges when designing protocols in the physical, MAC, and network layers [11][13].

Due to these challenges, the routing protocols used for terrestrial WSNs are not suitable for UWSNs, and hence, new routing protocols have been proposed specifically for relaying data in UWSNs. These routing protocols aim to address these characteristics and challenges, and mitigate their effects on the efficiency of the networks. There are a number of works found in the literature surveying and classifying these routing protocols from different perspectives. In [8], the authors provided a comprehensive description of several routing protocols. In addition, they categorized the described protocols based on their network architecture, data forwarding and protocol operation. The survey in [14] detailed a number of routing protocols; however, it did not provide any classification of the summarized protocols. The surveys in [15], [16] and [17] each concentrate on one category of routing protocols. In [15], geographic based routing schemes are reviewed and categorized based on forwarding strategy, location service and design goal. The work reported in [16] surveyed pressure based routing protocols. In [17], greedy routing protocols were summarized and classified into location-based and location-free protocols.

The aim of this paper is to provide an up-to-date survey of routing protocols. Basically, the idea of each protocol is clearly presented. The advantages and disadvantages of each protocol are discussed. Furthermore, the paper classifies the presented protocols based on location information dependency and the number of transmitted copies of each data packet in each hop.

The rest of this paper is organized as follows. Section 2 describes some of the architectures that can be used for UWSNs. Recent routing protocols are surveyed and classified in section 3. Section 4 outlines the main performance metrics commonly used in the evaluation of routing protocols for UWSNs. Finally, section 5 concludes the paper and proposes some open research issues.

2. Architectures for UWSNs

A well-designed architecture and the deployment method of sensor nodes can effectively help in maximizing network capacity, minimizing energy consumption and improving network reliability. According to [18], the architectures for UWSNs can be classified based on the coverage of the network and the mobility of the sensor nodes. For the network coverage, UWSNs can be designed to cover a two-dimensional (2D) or three-dimensional (3D) space within the underwater environment. Regarding their mobility, sensor networks can be stationary, mobile or hybrid. In a stationary UWSN, nodes are attached to surface buoys and/or anchored to the ocean floor. In a mobile UWSN, on the other hand, sensor nodes have the capability to move, for example, with water currents or as autonomous vehicles. Examples of such nodes are Autonomous Underwater Vehicles (AUVs), Unmanned Underwater Vehicles (UUVs), drifters and gliders. A hybrid network, from its name, incorporates both stationary and mobile nodes. Here we review the architectures proposed for UWSNs according to their coverage. In addition, the pros and cons of each type are outlined.

2.1. 2D architectures

Sensor nodes in this category are kept static by means of anchors or surface buoys. They are deployed on the ocean floor (as depicted in Figure 1) and are responsible for monitoring 2D environments. In [19], the authors proposed a static 2D architecture for ocean bottom monitoring. In this architecture, sensor nodes are anchored at the bottom of the ocean, and they communicate directly with the underwater sinks on the surface. Monitoring of underwater plates for tectonic studies is an example of an application that can use this architecture [19].

Although the direct link is the simplest compared to multi-hop paths, it requires high transmission power because the distances between sensors and sinks can be long. Moreover, packets from different sensors may interfere with each other and thus reduce the throughput. The longer the distance that the signal should travel, the higher the energy consumption [20] and the more interference with other signals. Generally, a 2D UWSN architecture is suitable only for applications where the depth of the monitored space is less than the transmission range of sensor nodes, so that nodes
can communicate directly with sink nodes. Therefore, it is not applicable for environments with greater depths than nodes’ transmission ranges, such as deep oceans [13] [21]. In addition, it is not suitable for applications that require 3D monitoring or tracking.

### 2.2. 3D architectures

The 3D UWSN architectures have attracted the attention of the research community [22]. In this kind of architecture, sensor nodes are distributed at different depths to perform a collaborative monitoring task over the 3D monitored region as illustrated in Figure 2. It may consist of static, mobile or hybrid sensor nodes depending on the application requirements. While 2D architectures are used for ocean bottom or surface monitoring, 3D architectures are mainly used for ocean column monitoring.

In [23], the authors proposed a 3D architecture where sensor nodes are assumed to be able to adjust their depths. The sink node has the capability to move, and exchange data and control packets with the sensors via single-hop communications. The sink node is also assumed to know the locations of the sensor nodes. In order to minimize the distance that the sink node should travel, the authors adopted a partitioning scheme in which the network is divided into clusters. Then, the sink node traverses each cluster and exchanges packets with sensors belonging to each cluster. The gathered data is then transmitted to the base station at the surface. The problem with this architecture is that the sink node is assigned complex tasks, and it is a single point of failure. Another concern is that the architecture is not suitable for delay-sensitive applications, as the nodes in each cluster should buffer the packets until the sink arrives to that cluster.

Akyildiz et al. [19] proposed a static 3D UWSN architecture for ocean column monitoring. In this architecture, sensors float at different depths. They are anchored to the bottom of the ocean by means of wires that link sensors to the anchors. The sensors are also equipped with floating buoys to pull them toward the surface. However, this architecture might not reflect real environments in which nodes may move from one point to another due to the water currents. The authors proposed a 3D mobile architecture consisting of a collection of fixed sensor nodes deployed at different depths and mobile AUVs. The AUVs can be used to accomplish different tasks (e.g. deploying sensors and forwarding data).

Another 3D architecture has been presented in [24]. It consists of bottom nodes, underwater nodes, automatic mobile nodes (e.g. AUVs and UUVs) and surface sinks. The bottom nodes are anchored to the bottom of the ocean to sense the 2D area. The underwater nodes float at different depths and are equipped with floating buoys to adjust their depths. The bottom nodes, underwater nodes and automatic mobile nodes work together to achieve 3D monitoring. The automatic mobile nodes float at the surface to receive GPS signals and then dive and move among underwater nodes following predefined trajectories to help in the localization process or in information gathering. The sensed information is transmitted to the surface via multi-hop paths or via automatic mobile nodes. Typical applications of the 3D architecture may be surveillance applications, oceanography and monitoring of ocean phenomena (e.g. water streams and pollution) [19].

The challenges for such architectures are that sensors should regulate their depths such that the network is always connected, so that every sensor can send its data to the surface sink via multi-hop paths. In addition, ensuring the required coverage and energy efficiency is another challenge [21]. The mobility feature of sensor nodes may help in maximizing the coverage with a limited number of nodes since they can be equipped with different types of sensors and have the capability to reach points that might not be reachable by static nodes. However, mobility raises new challenges such as energy consumption, localization and connectivity maintenance [21].

It is worth mentioning that no single architecture is good for all application types. Hence, network designers and engineers should select the appropriate architecture based on the requirements of their applications and the nature of the targeted environment.

![Figure 1. A view of a 2D architecture.](image1)

![Figure 2. A view of a 3D architecture.](image2)
3. Recent routing protocols for UWSNs

Routing protocols for UWSNs can be broadly classified into location-based and location-free. Location-based routing protocols utilize location information of sensor nodes to relay data packets. Since sensor nodes do not employ any addressing system (e.g. IP) [1] and because sensors are randomly deployed in the environment, location information can be used to route data packets. Location-free routing protocols, on the other hand, employ other information such as hop count or link quality. The rest of this section summarizes some of these protocols.

3.1. Location-based routing protocols:

3.1.1. Energy efficient routing protocols

Energy conservation is one of the most critical tasks in UWSNs; thus, careful attention is needed when designing a routing protocol. Huang et al. [25] proposed a Power Efficient Routing (PER) protocol to reduce the energy consumed by sensor nodes. PER consists of two modules. The first module consists of four sub-components; the fuzzifier, the fuzzy rule base, the fuzzy inference engine and the defuzzifier. Each forwarder node uses a fuzzy logic technique to select two candidate nodes to relay the packet. It requires three inputs including the distance and the angle between the current forwarder and one of its neighboring nodes, and the remaining energy of the forwarder, and it works as follows. The fuzzifier takes the three inputs of the module and converts them to linguistic values. The distance value is converted to either short, medium or long, while the angle is converted to either small, medium or big. Low and high are used to convert the remaining energy value. The fuzzy rule base consists of some linguistic rules and control goals used by the fuzzy inference engine along with the output of the fuzzifier to determine the appropriateness of a node to forward a data packet. The output of the fuzzy inference engine is classified into excellent, good, medium, bad and weak. The defuzzifier converts these linguistic values into non-linguistic values and based on the results, the forwarder node selects the two best candidates and sends the packet to them. The selection of two candidates by each forwarder can cause a massive growth of the forwarding tree. Thus, the second module is used to prevent this unnecessary growth of the tree, and it works as follows. The forwarder node does not forward the packet to the second best selected node unless the number of duplicate packets received by the forwarder is less than a predefined threshold (δ). This will prevent unnecessary power consumption of the sensor nodes while forwarding the packets. The performance evaluation shows that PER reduces the energy consumption in dense networks. However, the packet delivery ratio of PER is highly dependent on the selection of the duplicate-packet’ threshold. Choosing a high threshold can improve the delivery ratio but at the cost of increasing power consumption.

The LE-VBF [26] routing protocol was proposed to enhance the network lifetime. It is an improvement over the Vector-Based Forwarding (VBF) [27]. In VBF, each packet carries the location of the source, sink, and forwarder (i.e. the node from which the packet is received). The idea of the VBF approach is that a virtual routing pipe from the source to the sink is constructed and the radius of the pipe is used as a threshold to determine whether a node is a candidate for forwarding a packet. When a node receives a packet, it checks whether it is located within the routing pipe and if so, it updates the packet’s header by putting its location value in the packet and then forwards the packet. Otherwise, it discards the packet. In order to reduce the number of forwarders, when a dense network is used, the authors have introduced a self-adaptation algorithm to determine the density of the network within its region. Upon receiving a packet, nodes within the routing pipe compute a desirability factor to weigh their suitableness to transmit the packet. The desirability factor is based on the position of the receiver node relative to the source node and the previous forwarder, the transmission range and the radius of the pipe. If the node is eligible to forward the packet, it holds it for a period of time before forwarding it. Otherwise, it discards the packet. Simulation of VBF shows that the data delivery ratio increases by increasing the density of the network and/or the radius of the routing pipe. However, nodes closer to the routing vector from the source to the destination have a high chance to be frequently selected as eligible candidates for forwarding the packet, which drains their energy and reduces the network lifetime.

LE-VBF suggests a solution to balance the energy consumption between nodes by including residual energy of the nodes as a factor in relay selection. In addition to the position information, each packet carries the residual energy information of the previous forwarder. When a node receives a packet, it checks whether it is located within the routing pipe and if so, it calculates its suitableness based on its position from the routing vector and its residual energy. It holds the packet for a period of time and if it does not overhear the transmission of that packet after the period elapses, it broadcasts the packet. However, the packet may be delivered with extra delay due to the waiting time at each hop.

The authors in [28] proposed an Energy Efficient Fitness based routing (EEF) to achieve improved energy efficiency. EEF is a depth-based routing in which the forwarder candidates are selected based on the depth information. Residual energy and the distances to the sink node and to the previous forwarder are additional factors to determine the best eligible candidate. In EEF, each packet carries the position and the fitness value of the sender. When a node f receives a packet from a node (say s), it discards the packet if it is located within a depth less than that of s. Otherwise, it calculates its fitness value as given in (1) [28] where \( E_d \) is its residual energy, \( \text{depth}_{	ext{diff}} \) is the difference between its depth and the depth of the previous hop, \( d_{sf} \) is its distance to the previous hop and \( d_{sd} \) is its distance to the destination. If the fitness value is less than the fitness included in the packet, it discards the packet. Otherwise, it holds the packet for a certain period of time. If the period expires without overhearing the transmission of that packet, it includes its position and fitness value in the packet and broadcasts the packet. The process continues until the packet reaches its destination.
Although the evaluation of the protocol shows that it can save energy and improve the network lifetime and end-to-end delay, EEF has a number of issues. First, in a dense network, a high number of nodes will receive the packet, which consumes extra energy. Second, the calculation of the fitness value and the waiting time is done at every hop by each candidate node, which increases the complexity and end-to-end delay when a high traffic network is used. Third, in high traffic conditions, nodes need to keep track of a large number of packets that might overflow their buffers.

Javid et al. [29] proposed two routing protocols called Efficient and Balanced Energy consumption Technique (EBET) and Enhanced EBET to achieve efficient and balanced energy consumption. The protocols assume that the network is divided into concentric circles called ring sectors and nodes’ energy is divided into levels called energy levels. Each of the protocols consists of two phases, namely a route establishment phase and a data transmission phase. During the route establishment phase of EBET, nodes share their location and energy levels. Each node builds a set of routes toward the sink node based on the distance information. These routes are used in the data transmission phase to select the optimal relay based on the neighbors’ energy levels. The node with the highest energy level is selected from the set of candidates. Although EBET tries to balance energy among nodes by selecting neighboring nodes with high energy levels to forward the packet, the selection might cause transmission loops, and hence waste energy without delivering data packets to their destinations. The problem is solved by the EEBET such that depth information is used in building the routes. Consequently, backward transmissions are avoided by always selecting those forwarders that are located in depths less than that of the node. It is worth noting that neither EBET nor EEBET are suitable for highly dynamic networks due to the frequent need for updating the routes, which might incur extra overhead.

The authors in [30] proposed a Level-Based Adaptive Geo-Routing (LB-AGR) protocol for energy efficiency and low delay. In LB-AGR, sensor nodes are able to obtain their levels (i.e. the hop distances between the sensor nodes and the sink) and location information via flooding control packets initiated by sink nodes. In addition, each node maintains a neighbor table for each sink, storing some information about neighbors within two hops from the specific node (e.g. node ID, its level, location, and remaining energy). Data packets are routed based on the gathered information (i.e. level, location, residual energy and node density). When a node has a packet to be transmitted, it searches its neighbor table for nodes whose level is less than its own (i.e., nodes closer to the sink than itself). Since more than one candidate might satisfy the condition and in order to avoid redundant transmissions of the same packet which might interfere with each other and waste energy, LB-AGR calculates a desirable factor based on the remaining energy and node density (the number of one-hop neighbors whose level is closer to the sink than the node itself) to select the optimal node. The node with the highest desirable factor is selected. However, LB-AGR has some issues. First, the dissemination of control packets to update neighbor information can incur extra overhead and drain nodes’ energy, especially in dense and dynamic networks. Second, nodes need to store one neighbor table for each sink. This will require extra storage when multiple sinks are used or when a moderately or highly dense network is used.

### 3.1.2. Depth-controlled routing (DCR)

In UWSNs, geographic routing protocols are preferred over proactive and reactive counterparts, because they do not require the establishing and maintaining of complete routes, which incur overhead and consume resources due to the dynamic nature of the underwater environment. Geographic routing makes an optimal selection of the next hop with no concern about the global view of the entire path. This greediness in the relay selection creates void communication regions (i.e., holes) in the network, which occur when a node is the closest node to the sink but cannot communicate directly with it [31]. Nodes located in these regions are called void nodes.

In the Depth-Controlled Routing (DCR) protocol [32], the authors suggested a solution for the void node problem by adjusting the depth of such nodes. The authors assumed that the network is static, and there is a center for monitoring that has a global vision of the network. They also assumed that the nodes have the ability to move in the vertical direction. The protocol consists of two phases: an initialization phase and a network operation phase. In the initialization phase, the nodes are localized with the help of the AUVs, which transmit the location information to the sink nodes. The sink nodes send this information to the monitoring center. Then, the topology control algorithm is executed by the monitoring center to determine the void nodes and their new suggested depths so that they can communicate with the sink nodes via multi-hop communication. Consequently, the AUVs dive towards their suggested depths to inform the nodes about their new depths. In the network operation phase, when a node has a packet to send, it selects the closest node to the sink from the set of its neighbors and forwards the packet to it. Upon receiving the packet, this node forwards the packet in the same manner and the process continues until the packet reaches the destination node. Although the protocol can reduce the number of disconnected nodes and improve the delivery of the packets to their destination, the assumption that nodes become static after the depth adjustment phase is not practical due to the dynamic nature of the environment. Moreover, nodes that are close to the sink nodes are frequently selected which causes a quick depletion of their energy, hence reducing the network lifetime and creating new void nodes in the network. Another issue is that it is not clear how nodes determine their neighboring nodes along with their distances. In addition, the availability of a monitoring center with a global vision of the network might not be practical.
In [31], the authors presented a Distributed Topology Control (DTC). In DTC, each node determines if it is located in a communication void region, and, if so, determines its new depth and starts moving to it. This operation is achieved by disseminating control packets between nodes. Nodes located near the sinks are adjusted first, followed by those nodes at higher depths. As in DCR, the depth adjustment is performed before sending data packets, and, after the depth adjustment phase, the nodes are assumed to be static.

### 3.1.3. Grid-based routing

Routing protocols that depend on constructing complete paths between nodes to route data packets require the exchanging of control packets in order to establish and maintain the routes. The overhead increases in the underwater environment because of the dynamic topology, which requires continuous maintenance of the routes. Moreover, flooding-based routing protocols suffer from the broadcast storm problem. The increase in the number of packets propagated in the network causes a quick drain of its resources (e.g. bandwidth and energy). Due to the limited resources in UWSNs, there is a significant demand to reduce the amount of transmission and reception of the packets. In [33], a Multi-path Grid-based Geographic Routing (MGGR) protocol is developed to reduce the number of the required control packets and prevent some nodes from forwarding data packets. MGGR assumes that the network is divided into 3D logical. Each cell can have up to 26 neighboring cells based on the relationship between the length of the cell side and the transmission range of the nodes. Data packets are routed in a cell-by-cell manner via neighboring gateways (i.e. sensor nodes elected as cell heads). Since the routes are constructed based on grid cells, there is no need to maintain the routes after construction.

The protocol consists of three main components: namely, a gateway election algorithm, a mechanism for updating neighboring gateways’ information, and a packet forwarding dealing with holes. The gateway election algorithm is responsible for electing gateways based on their locations and residual energy level. A node that is closer to the center of the cell it belongs to and which has high residual energy has higher priority to be elected.

The second component, as its name implies, is responsible for keeping the nodes updated with gateways in local and neighboring cells. The packet forwarding mechanism is responsible for constructing disjoint paths (i.e. paths with no common intermediate cells) from source cells to destination cells, forwarding packets to the destination and dealing with holes (i.e. cells with no gateways) in the network. The construction of the paths is performed by gateways, and only on demand when there is a need to forward data packets and there are no paths constructed yet. Once the paths are constructed, the gateway stores the paths until it moves out of its cell or a new gateway is elected for that cell. When a node has a data packet to be forwarded, it looks for a gateway in the local cell. If there is a gateway, then it forwards the packet to it; otherwise, it discards the packet. When a gateway receives a packet, it selects a path from its table and forwards the packet to the gateway in the next hop. Upon receiving that packet, this next gateway checks the path in the packet header and looks for a gateway in the next hop to forward the packet to it. If the gateway encounters a hole in the next-hop, it tries to re-route the packet through another path or send a negative acknowledgment to the previous hop if all routes are broken. Then, the gateway in the previous hop will do the same thing as a recovery process from the hole. If the packet reaches the source gateway and all paths are broken, the packet is dropped. Since nodes are continuously moving with water current, the holes might become filled with nodes and thus paths can be reused.

The reachability of MGGR was evaluated analytically under the effect of varying network density (i.e. number of nodes per cell), reliability level (i.e. the number of the available paths to be used), and the transmission range of the nodes. Results indicate that the reachability of the protocol increases by increasing these three factors. However, the evaluation assumes a static network, which is not practical in an underwater environment.

EMGGR [34] is an extension of MGGR. The extension includes three main points: (i) updating the gateway election algorithm, by allowing nodes to serve as gateways for one or more periods, depending on their remaining energy levels; (ii) modifying the construction of the routing paths; and (iii) evaluating the performance of the new protocol in a simulation package by testing its performance under the effects of varying network density, traffic load and node mobility. The performance evaluation demonstrates that EMGGR is an energy-efficient protocol because data packets are forwarded only by gateways and there is no need to maintain paths. The evaluation also shows a good delivery ratio of the data packets. On the other hand, EMGGR incurs extra delay in packet delivery due to the constructed long paths. Moreover, the performance degrades in a highly dynamic topology and in sparse networks due to the nature of grid-based protocols.

Jiang et al. [35] proposed two grid-based routing protocols called GFGD and GGFGD. Besides the assumption of the grid topology, the authors assumed that the channel link is symmetric and the nodes are stationary. In addition, they incorporated a duty-cycle mechanism (i.e. some nodes are scheduled to sleep for some time) to save energy. As in MGGR and its successor, the number of neighboring cells is determined based on the relationship between the transmission range and the length of the cell side. In GFGD, a cell can have up to 32 neighboring cells; while in GFGD, it can only have up to six neighboring cells. The proposed schemes consist of two phases; selecting the next cube to forward the packet and choosing the best candidate node from the selected cube based on the distance, residual energy and path loss. The GFGD and GGFGD protocols incorporate a duty cycle, which can balance and save energy. In addition, they use path delay, path loss and remaining energy for relay selection; however, it is not clear how nodes acquire this information. Moreover, the assumption that the channel link is symmetric is not practical since acoustic channels underwater are known to be asymmetric [36]. In addition,
the protocols assume a 3D stationary deployment of the nodes, which is also not practical due to the passive movement of nodes with water currents.

3.1.4. Chain-based routing protocol

In [37], 4-chain, 2-chain and single-chain based routing schemes (i.e. a network with a cylindrical shape) are proposed to improve network lifetime and throughput. The protocols consist of two main phases; namely, an initialization phase and a protocol operation phase. In the initialization phase, nodes broadcast their location information to be used in forming chains and optimal paths. In the protocol operation phase, the chains are formed based on the location information starting from the farthest node from the sink. A local optimal path between nodes in each chain is constructed. Then, each chain is connected to the nearest node in the next chain forming a global optimal path. Finally, data packets are transmitted through the global optimal path. Simulation results show that the 4-chain routing scheme performs better than its counterparts do. However, the proposed protocol is not suitable for mobile underwater networks because nodes continuously change their positions. Hence, location information needs to be rebroadcasted in order to reconstruct the chains and optimal path. This results in extra overhead and drains energy.

3.2. Location-free routing protocols

3.2.1. Channel aware routing protocols

Due to the limited resources such as energy and bandwidth and to the other challenges such as the high error rate of acoustic channels, there is a high demand to consider link quality when designing a routing protocol to alleviate these challenges and to avoid retransmission of the packets. Basagni et al. [38] proposed a Channel Aware Routing Protocol (CARP) for UWSNs to take the link quality into account when selecting the next data forwarder. Nodes in CARP need to know their hop count to the sink (i.e. number of hops to reach the sink). This is achieved by broadcasting HELLO control packets initiated by the sink node in the initialization phase. When a node has one or more data packets to be forwarded, it broadcasts a PING control packet containing its hop count and the list of data packets’ IDs to be forwarded, and waits for replies. Upon receiving such a packet, a node drops it if its hop count to the sink is greater than the hop count included in the packet. Otherwise, it replies with a PONG control packet including its hop count, available buffer space, residual energy, link quality to its neighbors, and IDs of the packets (from the list included in the corresponding PING packet) already received by the sink and those already received by the node itself. After receiving the PONG replies, the node selects the one with the best link quality (the one that exhibited most successful transmissions recently) and lowest hop count. If there is a tie, then residual energy, buffer space and node ID are used to select the optimal relay. After selecting the forwarder, the node sends only the packets that are not received by the sink and those that are not received by the selected relay. After receiving the data packets, the node sends an acknowledgement packet indicating the IDs of the received packets.

An enhanced version of CARP called E-CARP is proposed in [39] to reduce energy consumption and prolong the network’s lifetime. Since some applications require that the sensed information is transmitted only if the changes between the previous and recent sensed information is higher than a certain threshold, E-CARP employs this requirement. It achieves this by caching the information sensed by each node in the sink. In addition, each node also caches the information it senses to compare it with the recent sensed information and forwards the information only when the difference is higher than the predefined threshold. This can reduce the number of data packets to be forwarded especially if the threshold is large. In addition, E-CARP enhances the relay selection mechanism by realizing that when the changes in the position of the nodes are relatively small, the previous candidate forwarder may be the candidate for forwarding the current packet. The performance of the protocol has been evaluated and compared with its predecessor CARP. The results show an improvement in energy consumption especially when the data packet size is increased and when the bias threshold is high. However, this protocol it is not suitable if the link quality is changing frequently and if the nodes change their positions frequently. Furthermore, it has high storage space (e.g. buffers) requirements.

3.2.2. Routing protocols based on physical distance and residual energy

Bandwidth limitation and path loss are two factors that affect the design of UWSNs especially when increasing the distance between the sender and the receiver. Therefore, considering the physical distance towards the sink is important when selecting the next candidate to forward data packets. In [40], the authors proposed an energy-efficient routing protocol (ERP\(^R\)) that utilizes physical distance from the sensor nodes to the sink to forward data packets towards the destination. In addition, residual energy of the nodes is taken into account to prolong the network lifetime (e.g. the time when the first node in the network loses its full energy [40]). ERP\(^R\)
consists of two phases. In the first phase, the sink node broadcasts a HELLO packet. Nodes that receive this packet compute their distances to the sink using the Time of Arrival (ToA). Then, they embed their distances to the sink and their residual energy in the packet and rebroadcast it. Upon receiving this packet, nodes compute their distance to the forwarders using ToA and accumulate it to the distance stored in the packet as its own distance to the sink. Then, they rebroadcast the packet after updating the distance and residual energy. The process continues until all nodes have computed their distances to the sink. By the end of this phase, each node knows its own physical distance and residual energy of its own and those of the neighbors. The protocol assumes that the vertical movement of the nodes in underwater is negligible, and consequently, the depth information remains the same. However, the residual energy is continually changing due to the different activities performed by sensor nodes (e.g., transmission and reception of the packets); thus, nodes periodically check their remaining energy. If the difference between the current remaining energy of a node and the previous value is larger than a threshold, it broadcasts a HELLO packet to inform its neighboring nodes. Moreover, the cost establishment phase is performed periodically.

The second phase is responsible for selecting the forwarders. During this phase, when a node has a data packet to be transmitted, it selects those nodes closer to the sink than itself and sorts them in decreasing order of their residual energy. The node checks the list and discards the packet if the ID of that node is not in the list. Otherwise, it holds the packet for a certain amount of time based on its position in the list to avoid redundant transmissions that might interfere with each other. If the holding time of a packet expires without overhearing the transmission of the packet, the node transmits the packet. However, if a node overhears the transmission of a packet it holds, it suppresses the transmission of that packet. This reduces the number of copies of the same packet propagated in the network and helps preserve energy. The idea of taking into account the residual energy of the nodes when forwarding data packets balances the energy among nodes and improves the network lifetime. However, the protocol has some issues. First, if a node is in the forwarding list of a packet, then, it needs to keep monitoring the transmission of that packet, and needs to store the packet until the holding time expires or it overhears the transmission of the packet by another node. In a high traffic network, the buffer might be overflowed and packets might get dropped. Second, in highly dense networks, a node needs to keep track of the distance and residual energy of a large number of nodes, thus requiring a huge storage. Third, when more than one sink is used, each node needs to keep track of a number of neighboring nodes for each sink and the overhead of the cost establishment phase increases. In addition, as the network density increases, the number of HELLO packets increases, which increases the energy consumption and reduces the network lifetime. Fourth, using ToA to compute distances requires the nodes to synchronize their clocks, but the paper does not mention anything regarding clock synchronization.

**R-ERP**[2] [41] is a successor of **ERP**[2]. **R-ERP**[2] adds the link quality as a third metric in the relay selection. Due to the high error-rate of the acoustic signal, which leads to high packet losses and degrades the performance of the routing protocol, the authors incorporate link quality to improve reliability and throughput. They use the Expected Transmission Count (ETX) [42] as an estimate for link quality. ETX is based on the forward and reverse packet delivery ratio between two points. The link quality is computed in the cost establishment phase. In the data forwarding phase, the sender selects a node that has better link quality and high residual energy from those closer to the sink than itself, and forwards the packet to it instead of the remaining group of candidates as in the **ERP**[2]. Upon receiving the packet, a node checks the ID included in the packet. If the ID is its own ID, it follows the same process in forwarding the packet to the next hop; otherwise, it discards the packet. This process continues until the packet reaches the destination. Although incorporating the link quality in the forwarding process helps in selecting a good candidate for packet relay, link quality in UWSNs is unstable due to the nodes’ mobility. This link instability might lead to packet loss; thus, the authors incorporate an acknowledgment strategy into the protocol. Upon sending a packet, a source node buffers a copy of the packet and if it overhears the packet transmission by the selected candidate, it removes the packet from its buffer. On the other hand, if it does not overhear the transmission of the packet after a period of time, it retransmits the packet. The packet is dropped after a certain number of retransmission trials. The evaluation of the **R-ERP** shows an improvement over its predecessor, **ERP**[2], in terms of network lifetime, energy consumption, end-to-end delay and delivery ratio.

The acoustic link is asymmetric, in that the strength of the link in one direction might be different from the strength in the reverse direction. Therefore, the ETX measurement of link quality, which depends only on forward and reverse delivery ratio between two nodes, might lead to inaccurate estimation of the link quality. For example, assume we have two links SX and SY where S, X and Y are three nodes. Assume that the forward delivery in SX is 100 and the reverse delivery is zero, and in the second link the forward delivery in SY is 50 and the reverse delivery is 50. To forward data from the source S, the ETX estimator always favors the link SY over SX although the forward delivery is much better in the link SX than in SY. The authors of [36] proposed a Distance based Reliable and Energy Efficient (DREE) routing protocol, trying to improve **R-ERP** by using another estimate for the link quality. They used the Fuzzy logic based link Quality Estimator (F-LQE) [43] instead of ETX. The F-LQE depends on a number of parameters including the stability and the asymmetry of the link, and the delivery ratio. As ETX in **R-ERP**, the F-LQE in DREE is computed in the cost establishment phase and it is updated periodically. Another difference between DREE and **R-ERP** is that in DREE the physical distance towards the sink is not just for finding the possible candidates, but also for calculating the cost of transmission to each neighbor. In other
words, a node with high link quality, high residual energy and less distance has a high priority to be the next forwarder. DREE was evaluated and compared with R-ERP\textsuperscript{R}, and the evaluation shows that DREE shows a slight to small improvement over R-ERP\textsuperscript{R} in terms of network lifetime, energy consumption end-to-end delay and delivery ratio.

3.2.3. Depth-based routing protocols

As mentioned earlier, providing a location-service in UWSNs can be expensive due to the inability to use the GPS system, and to the continuous movement of nodes with water currents. Therefore, depth-based routing protocols have been proposed as an alternative to the need for location information in data transmission. In such protocols, data packets are forwarded based on depth information, which can be obtained via inexpensive depth sensors built into sensor nodes [44]. Perhaps the first depth-based routing scheme for UWSNs is the one proposed in [44] and called DBR. In DBR, data packets are forwarded in a greedy manner towards the sink on the surface using only depth information. Basically, when a source node has a data packet to be transmitted, it includes its depth information in the packet and then broadcasts it. A node, upon receiving that packet, compares its depth with the number included in the packet. It discards the packet if it is located deeper than the source node. Otherwise, it replaces the depth information in the packet with its own and then broadcasts the packet. The process continues until the packet reaches one of the deployed sinks. The main advantage of DBR is its simplicity, in that nodes need to know only their depth information. There is no need for location information and route discovery to deliver data packets to the destinations. However, the main drawbacks of DBR are void problems in sparse networks, and severe collisions between packets due to the broadcast nature of the protocol.

In [45], a Void Aware Pressure Routing (VAPR) protocol is proposed to handle the void problem of the DBR protocol. It consists of two main components, which are enhanced beaconing and opportunistic data forwarding. The aim of the first component is to build a directional route from each node to the closest sink node. Its main idea can be summarized as follows. Each sink node initiates a beacon packet containing its depth, sequence number of the packet, minimal hop count and its directional forwarding. Nodes upon receiving such a beacon packet set their information accordingly and update the information in the packet by replacing the depth with their own depths, incrementing the hop count and setting the direction of the current node towards the sink. Then, the packet is retransmitted. The directional forwarding is set to upward if the packet is received from a node with a depth less than that of the receiver, and downward if it is received from a deeper node. The second component uses the directional route constructed in the first component to forward data packets toward the sink nodes. Although VAPR succeeds in handling the void problem, executing the enhanced beaconing periodically can increase the network overhead.

Li et al. proposed a DBR-MAC [46] to reduce the collisions in the DBR protocol. DBR-MAC is a cross-layer approach that integrates a depth-based routing and a handshaking-medium access control (MAC). Furthermore, it aims to give those nodes with extra load (i.e. key nodes) priority to access the channel. DBR-MAC uses the depth, angle and overheard one-hop neighboring nodes’ transmissions for scheduling the transmissions and giving the key nodes higher priority. The node backs-off for a certain period of time if its packet will collide with another transmission. The protocol has been evaluated via simulation. It gives a good performance in terms of throughput, energy consumption and delay at the cost of fairness between nodes.

3.2.4. Diagonal and vertical routing protocol (DVRP)

Forwarding packets horizontally between nodes can sometimes be an extra step that increases the routing path and burdens the nodes’ energy. Tariq et al. [47] proposed a Diagonal and Vertical Routing Protocol (DVRP) for UWSNs to avoid such horizontal communication between nodes. DVRP is an enhanced flooding-based protocol where the flooding zone is controlled by $\theta = 90 \pm 10 \, \text{K}$, where $1 \leq K \leq 8$ ensures that data packets are propagated only vertically or diagonally. When a node has a data packet to be forwarded, it first calculates its flooding zone. For example, if the selected value of $K$ is 1, then the angle of the flooding zone ranges from 80$^\circ$ to 100$^\circ$. Then, the node broadcasts a Hello packet (HP) including the flooding zone. The nodes within the flooding zone reply with Hello Reply (HR) packets. If the waiting time expires without any reply, then the node rebroadcasts the HP packet with an increased flooding zone. If, however, one or more nodes are located within the flooding zone, they should calculate and reply with their priority based on their remaining energy and depth. If the node receives more than one HR packet, then the node with highest priority is selected as the next relay which will follow the same process until the packet reaches its destination. The protocol suffers from the problem of flooding control packets at every hop of packet propagation. The source node keeps trying to send HP packets until it receives one or more HR packets. This not only consumes high energy and reduces the network lifetime, but can also induce extra delay on packet transmission. Moreover, it is not clear how the nodes calculate their priority to be the next forwarder.

Figure 3 classifies the above described protocols based on two criteria: location information dependency and the number of transmitted copies of each data packet in each hop. We can conclude from the figure that most of the recent routing schemes developed for UWSNs are location-based routing. However, the main problem with this type of protocols is the communication void region as mentioned earlier. Some of the protocols suggest recovery processes to handle this problem, such as retransmission or node adjustment as mentioned in some of the above
summarizes. Chen et al. [48] have surveyed some of the available techniques for void handling in location based protocols for WSNs. It is also worth mentioning that most of the above protocols avoid transmitting more than one copy of each packet to preserve the limited resources available and to avoid interference between packets. Table 1 further summarizes these protocols. Each of the location-based and location-free categories is further classified into sender-based and receiver-based. These are also subdivided based on the number of sinks assumed by the protocols.

![Figure 3. Classifications of routing protocols for UWSNs](image)

Table 1. Summary of the routing protocols

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Evaluation method</th>
<th>Static/dynamic topology</th>
<th>Requirements &amp; assumptions</th>
<th>Maintained knowledge</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sender-based</td>
<td>PER [25]</td>
<td>Simulation (random-walk)</td>
<td>Dynamic</td>
<td>Location info. is available</td>
<td>Counting the number of copies of each received packet</td>
</tr>
<tr>
<td>MGGR</td>
<td>Analytical analysis</td>
<td>Static</td>
<td>The network is viewed as 3D logical grids</td>
<td>Own and sink location</td>
<td>Considered as the first grid-based routing for UWSNs</td>
</tr>
</tbody>
</table>
## AN OVERVIEW SURVEY OF RECENT ROUTING PROTOCOLS

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Type</th>
<th>Simulation</th>
<th>Location info.</th>
<th>Works better in dense networks</th>
</tr>
</thead>
</table>
| **EMGGR**       | Simulation    | Static & Dynamic (random-walk) | • Location info. is available  
                    • Sink is static and located at the top-middle of the grid  
                    • Packets are forwarded through disjoint paths via gateways | • Same as MGGR  
                    • Same as MGGR | • An enhanced version of MGGR  
                    • Works better in dense networks |
| **GFID & GGGID**| Simulation    | Static                   | • Location info. is available  
                    • Static sink | • Location info., residual energy, layer number for calculating the priority  
                    • History of the forwarded data packets | - |
| **EBET**        | Simulation    | -                        | • Location information is available | • Distance to neighboring nodes & sinks  
                    • Relay nodes  
                    • Energy levels of neighbors | • uses distance & energy levels of neighbors |
| **EEBET**       | Simulation    | -                        | • Location and depth info. is available | • Distance to sinks & neighbors & their depths  
                    • Relay nodes  
                    • Energy levels of neighbors | • Solves the transmission loop problem in EBET |
| **CARP**        | Simulation & Real experiments | Static | - | • Hop distance to the sink  
                    • IDs of data packets received by the node & the sink | • Utilizes link quality in relay selection |
| **E-CARP**      | Simulation    | -                        | • Topology is relatively steady  
                    • Applications does not require transmitting packets with low bias from previous sensed information | • Hop count | • Enhanced CARP  
                    • Suitable for specific applications & with static topologies |
<table>
<thead>
<tr>
<th>Protocol</th>
<th>Type</th>
<th>Nodes</th>
<th>Location &amp; Energy Info</th>
<th>Main Features</th>
</tr>
</thead>
</table>
| DCR [32]      | Simulation | Static | -                      | ▪ Static network architecture  
▪ Nodes are equipped with depth-adjustment mechanism  
▪ Location info. is available  
▪ Global vision of the network owned by a monitoring center  
▪ Non-void Neighboring nodes’ location  
▪ Considered as the first protocol that takes into account the nodes’ vertical movement ability |
| LB-AGR [30]   | Simulation | -     | ▪ Location info. is available | ▪ Location, level & residual energy of neighbors  
▪ Enhanced ERP'R  
▪ Reliability improved by including link quality in relay selection  
▪ Data packet is sent to a single node |
| R-ERP'R [41]  | Simulation | -     | ▪ Nodes can compute their distances to their neighbors through ToA/TDoA | ▪ Distance, ETX & residual energy of its own & its neighbors  
▪ Enhanced R-ERP'R  
▪ ETX is replaced by F-LQE |
| DREE [36]     | Simulation | Dynamic | ▪ Nodes can compute their distances to their neighbors through ToA | ▪ Distance, F-LQE & residual energy of its own & its neighbors  
▪ Enhanced R-ERP'R  
▪ ETX is replaced by F-LQE |
| VAPR [45]     | Simulation | Dynamic | ▪ Nodes knows their depth information | ▪ Node’s depth  
▪ Hop count to the closest sink  
▪ Directional forwarding route toward the sink  
▪ An enhanced DBR routing  
▪ Handles void problem of greedy protocols |
| DBR-MAC [46]  | Simulation | Dynamic | ▪ Nodes knows their depth information | ▪ depth, angle & overheard one-hop neighboring nodes’ transmissions  
▪ Integrates a handshaking MAC with the DBR |
| DVR [47]      | Simulation | Static sinks & mobile sensor nodes | - | ▪ a priority queue of the transmitted packets  
▪ An angle-based flooding protocol |
| LE-VBF [26]   | Simulation | Dynamic | ▪ Location info. is available  
▪ Location & energy info. are included in the packet | ▪ Own & sink location  
▪ Residual energy  
▪ Enhanced VBF  
▪ Network lifetime is improved |
AN OVERVIEW SURVEY OF RECENT ROUTING PROTOCOLS

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Energy consumption</th>
<th>Network lifetime</th>
<th>PDR</th>
<th>End-to-end delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>PER [25]</td>
<td>Fair</td>
<td>Fair</td>
<td>Fair</td>
<td>Low</td>
</tr>
<tr>
<td>MGGR [33]</td>
<td>Low</td>
<td>Fair</td>
<td>Fair</td>
<td>Low</td>
</tr>
<tr>
<td>EMGGR [34]</td>
<td>Low</td>
<td>Fair</td>
<td>Fair</td>
<td>Low</td>
</tr>
<tr>
<td>GFGD [35]</td>
<td>Fair</td>
<td>Fair</td>
<td>Fair</td>
<td>Fair</td>
</tr>
<tr>
<td>GGFGD [35]</td>
<td>Fair</td>
<td>Fair</td>
<td>Fair</td>
<td>High</td>
</tr>
</tbody>
</table>

4. Performance metrics

According to the above presented routing protocols, there are four main performance metrics that are generally used to evaluate routing protocols. These are energy consumption, network lifetime, packet delivery ratio (PDR) and end-to-end delay. Perhaps energy is one of the most critical resources in underwater sensor networks, for the reasons mentioned earlier. End-to-end delay is an important metric for real time applications, and PDR is important for applications that require full information about the sensed environment. They can be defined as follows:

4.1. Energy consumption

This is the total energy consumed by all nodes during the simulation. It includes the transmission power, the reception power and the idling power consumed by all nodes [49]:

\[
\text{Total Energy} = \sum_{i=1}^{N} (\text{transmission power}_i + \text{reception power}_i + \text{idling power}_i)
\]  

(2)

where \(N\) is the total number of nodes used during the simulation.

4.2. Packet deliver ratio (PDR)

The ratio of the number of distinct data packets that are successfully delivered to the sink nodes to the total number of data packets generated at the source nodes [49]. Formally, it can be written as:

\[
PDR = \frac{\sum \text{(number of packets received)}}{\sum \text{(number of packets sent)}}
\]  

(3)

4.3. End-to-end delay

The average time taken by a data packet to arrive to the destination [49]. It is computed from the time a packet is generated until it reaches the destination. Only the data packets that have been successfully delivered to destinations are counted. Mathematically, it can be calculated as:

\[
\text{Delay} = \frac{\sum_{i=1}^{\text{pkts}} \text{(arrival time}_i - \text{sending time}_i)}{\text{pkts}}
\]  

(4)

4.4. Network lifetime

The time span when the sensors are deployed to the time when the first node in the network loses all its energy [40].

Table 2 demonstrates a performance comparison of the above surveyed protocols under the four performance metrics mentioned (Energy consumption, network lifetime, PDR and end-to-end delay).

Table 2. Performance comparison of the presented routing protocols
5. Conclusion and future research directions

At present, routing protocols for underwater sensor networks (UWSNs) is a hot topic that has attracted a lot of attention in the research community. This paper has surveyed recent routing protocols for UWSNs. The basic idea of each protocol is presented, as well as its advantages and disadvantages. In addition, the protocols presented are classified in different ways.

Although the presented routing protocols have aimed to alleviate the challenges encountered in communication in the underwater environment, none of them has achieved good performance with respect to all metrics or tackled all aspects in the real environment. For example, the effect of node mobility was absent in most of the evaluation studies. Thus, one of the possible future research directions is incorporating the effect of node mobility in the evaluation of the routing protocols. This will help in obtaining more realistic evaluation results. In order to achieve this, a realistic mobility model that can reflect the physical movements of sensors with water currents is needed to simulate the movements of the nodes underwater. Moreover, improving routing performance of UWSNs by means of a prediction mechanism on the basis of the study of node movement characteristics is another research direction.

Another possible future research direction is to optimize the routing protocols via cross layer design. In a non-cross layer system, communication is restricted between adjacent layers. However, this restriction has been released in cross layer design. In other words, the different layers of the conventional Open System Interconnection (OSI) model share the information and interact with each other regardless of their positions in the model. Cross-layer design has been widely used in WSNs to eliminate the effects of the limited resources available, such as memory and energy [50].

Third, studying the effects of duty-cycle on the performance of UWSNs and the methods of applying duty-cycle mechanisms to such networks is another research direction. In a duty-cycle mechanism, some nodes may be scheduled to sleep for some time [51]. This can preserve the energy of the nodes, but might have adverse effects on the delivery ratio and the end-to-end delay.

Finally, the study of how data aggregation techniques may affect the performance of UWSNs should be addressed. Data aggregation aims to reduce the amount of data to be transmitted by summarizing or compressing them before transmission [52]. Such techniques are used in terrestrial WSNs [52-54] for energy efficiency. It would be interesting to investigate whether they have similar benefits on the performance of UWSNs.

Acknowledgment

This work is supported by The Research Council (TRC) of the Sultanate of Oman under the open research grant number RC/SCI/COMP/15/02.

References

4. Ribeiro, F.J.L., de Castro Pinto Pedroza, A. and Costa, L.H.M.K. Underwater monitoring system for oil...
AN OVERVIEW SURVEY OF RECENT ROUTING PROTOCOLS