Journal of Network and Computer Applications 75 (2016) 385–397

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

PRADD: A path reliability-aware data delivery protocol for underwater acoustic sensor networks

Nusrat Nowsheen[a,*], Gour Karmakar[a,b], Joarder Kamruzzaman[a,b]

[a] Faculty of Information Technology, Monash University, Australia
[b] Faculty of Science and Technology, Federation University, Australia

ARTICLE INFO

Keywords:
Underwater Acoustic Sensor Networks (UASNs)
Path reliability
Link transmission reliability
Node movement

ABSTRACT

Underwater Acoustic Sensor Networks (UASNs) are becoming increasingly promising to monitor aquatic environment. However, reliable data delivery remains challenging due to long propagation delay and high error-rate of underwater acoustic channel, limited energy and inherent mobility of sensor nodes. To address these issues, we propose a protocol called Path Reliability-Aware Data Delivery (PRADD) to improve data transfer reliability for delay tolerant underwater traffic. Data delivery reliability is significantly improved by selecting the next hop forwarder *on-the-fly* based on its link reliability, reachability to gateways and coverage probability through probabilistic estimation. Data forwarding solution is coupled with delay tolerant networking paradigm to improve delivery with reduced overhead. PRADD does not require active localization technique to estimate the updated location of a sensor node except its initial coarse location. The movement of an anchored node is exploited to estimate its coverage probability. Mobile message ferries are used to collect stored data from one or more nodes, called gateways. A strategy for gateway selection is devised to maximize their lifetime. Simulation results show that PRADD achieves significant performance improvement over competing protocols using low overhead and less energy.

1. Introduction

Covering more than 70% of the earth's surface, the ocean is still largely unexplored. During the last several years underwater networking has attracted rapidly growing interests from both researchers and industry due to its variety of applications, such as oceanographic data collection, pollution monitoring, offshore exploration, disaster prevention, assisted navigation, tactical surveillance, etc. (Akyildiz et al., 2005; Cui et al., 2006; Heidemann et al., 2006; Partan et al., 2006; Lanbo et al., 2008; Chitre et al., 2008). To make such applications viable, an underwater system with networked wireless sensors, referred to as Underwater Wireless Sensor Network (UWSN) is deployed. UWSNs mainly rely on acoustic, rather than radio or optical communication. Radio waves propagate to significant distance through conductive sea water at low frequencies (30–300 Hz), but at the cost of large antennae and high transmission power (Pompili and Akyildiz, 2009). Transmitting optical signals is possible but requires high precision in pointing the narrow laser beams (Pompili and Akyildiz, 2009). Thus acoustic signal is the only practical and feasible method for underwater communications (Stojanovic, 2001). However, some unique characteristics of underwater communications impose many challenges in the design of networking topology and protocol compared with terrestrial radio frequency (RF) communications. First, the propagation speed of acoustic signal in water (~1500 m/s) is five orders of magnitude lower than that of RF wave (3×10^8 m/s) (Akyildiz et al., 2005; Cui et al., 2006). This incurs large propagation delay in underwater communications. Second, the available bandwidth of underwater is extremely limited and depends on both transmission range and frequency (Akyildiz et al., 2005). Third, high bit error rate and temporary loss of connectivity (shadow zones) can be experienced due to the extreme characteristics of underwater channel (Akyildiz et al., 2005; Pompili and Akyildiz, 2009). Fourth, the power needed for acoustic communication in underwater is much higher than that for terrestrial RF communication because of the lower efficiency of the physical technology (Harris et al., 2009). Fifth, the characteristics of the communication medium vary with time and space. These intrinsic characteristics and some fundamental differences between UWSN and its terrestrial counterpart have great impacts on network protocol design and make existing data forwarding protocols (Perkins and Royer, 1999; Perkins and Bhagwat, 1994; David et al., 2001) designed for terrestrial radio networks unsuitable in underwater communication.

* Corresponding author.
E-mail addresses: nusrat.nowsheen@monash.edu (N. Nowsheen), gour.karmakar@federation.edu.au (G. Karmakar), joarder.kamruzzaman@federation.edu.au (J. Kamruzzaman).

http://dx.doi.org/10.1016/j.jnca.2015.11.021
Received 2 March 2015; Accepted 23 November 2015
Available online 16 April 2016

Due to the error-prone nature of the acoustic channel, difficulties in battery replacement and natural/inherent node mobility, reliable data delivery for large-scale UASNs is challenging and remains an open research problem. A link which is able to deliver data successfully at one moment may not be able to do so in near future. Compared with other communication paradigms such as ad hoc networks, wireless coverage and link quality are more susceptible in UASNs due to the disruptive nature of the communication medium and mobility pattern of sensor nodes. Because of high acoustic propagation delay and probability of connectivity loss among nodes during network operation in such harsh and dynamic deployment context, a fault tolerance paradigm embedding delay tolerant networking (DTN) needs to be introduced to improve the data transfer reliability. Besides, for such disruptive environment, to further improve the data transfer reliability, a strategy to collect the data *on-the-fly* is also essential. For this Autonomous Underwater Vehicles (AUVs) and other ocean (mobile) vessels such as ships, fishing boats can be exploited to enhance connectivity, coverage area and data collecting facility. Therefore, an energy-efficient, reliability-aware and opportunistic data delivery technique with embedded DTN facility is required for a large-scale UASN.

To enable large-scale and long-term underwater monitoring applications in a 3D architecture, quasi-stationary sensor nodes are usually preferred rather than mobile free-floating nodes. This is because mobile free-floating nodes can only penetrate into a very small depth (usually up to 100 m) and can quickly drift away from the communication range due to ocean tide/current (Jaffe and Schurgers, 2006). This makes the free-floating nodes suitable for short-term ocean monitoring systems only. Quasi-stationary sensor nodes, on the other hand, can be anchored with a cable/wire to the ocean bottom restricting their movements within the surface area of a hemisphere. Therefore, till now the quasi-stationary is the most feasible and accepted technology for long-term ocean monitoring system.

The new challenges for data delivery in underwater communications have recently inspired a significant amount of research efforts (Xie et al., 2006; Jornet et al., 2008; Shin et al., 2012; Chen et al., 2012; Erol and Oktug, 2008; Yan et al., 2008; Lee et al., 2010; Prasan and Mahapatra, 2012; Wahid and Kim, 2012; Domingo, 2011; Seah and Tan, 2006; Zhou and Cui, 2008). These works are detailed in the next section. Most of these studies assume that sensor nodes move randomly and base their works on localization techniques. Even though there exist some localization techniques underwater (Zhou et al., 2011, 2010, 2007; Chandrasekhar and Seah, 2006; Bian et al., 2009), yet localization is a difficult task to accomplish. Most existing underwater localization algorithms require reference nodes at known locations to assist localization of unknown nodes. In most cases, sensor nodes are assumed to be perfectly clock synchronized. However, achieving precise time synchronization is difficult in underwater environment (Han et al., 2012). The localization problem is compound by the fact that locations of underwater sensor nodes may change due to water current. However, for an anchored sensor node, it is possible to estimate an updated location using its movement information in a particular environment. The idea is to detect the changes in velocity of water current and estimate the location of sensor nodes without requiring any signaling method for localization.

Motivated by the above facts, in this paper, we propose a protocol, called Path Reliability-Aware Data Delivery (PRADD), to transfer data in a reliable manner for large-scale UASNs. We adopt an anchor-based UASN and introduce *accumulate-and-forward* approach at every hop. Each forwarding node accumulates a train of data packets before forwarding them to the next hop, and the process continues till data reach local gateway(s). We incorporate mobile nodes (termed as "ferry") that collect data from local gateways opportunistically within their one hop transmission range. During data forwarding, the next hop node is chosen *on-the-fly* rather than pre-determined by the delivery protocol. In PRADD, the path reliability is maintained by selecting the next hop node explicitly based on its link transmission reliability, reachability to the destination (gateway) and coverage probability. The link transmission reliability is estimated by overhearing next hop forwarders' transmission. Overhearing technique has been studied and used in wireless ad hoc networks for co-operative caching algorithm (Wu et al., 2013) to improve the performance of data caching, and routing protocols (Camp et al., 2002; Chatterjee et al., 2007) to construct route efficiently. We use the movement pattern of an anchored node to estimate its new location and then apply it in computing the coverage information of a neighbor node to improve data forwarding reliability. The updated location of a sensor node is calculated using an approximated initial location, the cable length, and the relevant forces that include ocean current, gravity, tension of the cable and buoyancy force.

The key contributions of this paper over the existing UASN protocols are as follows:

- Design of a data delivery protocol for UASNs to achieve path reliability in data transfer without requiring the support of active location estimation technique. The protocol attains higher successful data delivery ratio with less energy and message overhead.
- Probabilistic estimation of link transmission reliability, reachability to destination and coverage probability to the next hop.
- Inclusion of node movement due to ocean current and other forces exerted on a node in protocol operation, making the protocol more robust to deal with environmental factors.
- Selection and management of gateway nodes in a manner that improves their lifetime and network reliability.

Finally, we investigate the performance of PRADD through extensive simulations using ns2 and compare it with the existing contemporary methods. The rest of this paper is organized as follows. Section 2 surveys the network architectures and existing routing solutions along with their suitability in underwater communications. In Section 3, we describe the basic components and data forwarding mechanism of PRADD, while the probabilistic approach for selecting next hop forwarder, and accumulation and delivery of data to the selected forwarder are presented in Section 4. Section 5 shows the experimental setup and simulation results. Finally, Section 6 concludes the paper.

2. Background and related works

2.1. Network architecture

Based on the mobility of an underwater node, sensing technologies can be categorized into three groups (i) stationary (static 2D UASNs for ocean bottom monitoring), (ii) quasi-stationary (attached to surface buoy or anchored with wire to create a 3D UASNs), and (iii) mobile nodes (free floating or AUVs) (Akyildiz et al., 2005, 2004, 2007). In static 2D network (Akyildiz et al., 2005, 2004), a group of sensor nodes are anchored to the bottom of the ocean. However, this network is responsible only for monitoring the ocean bed. In quasi-stationary 3D architecture, a network of underwater sensors floating at different depths is deployed to perform cooperative sampling of the 3D ocean environment. In this architecture (Akyildiz et al., 2005, 2004), underwater sensor nodes are either attached to the surface buoys by means of wire or anchored to the ocean bottom. The former approach of deployment is easy but multiple floating buoys may obstruct to navigate ships on the surface. Moreover, floating buoys are vulnerable to weather condition and can easily be identified by the enemies in military settings (Akyildiz et al., 2005). In the latter, as underwater nodes are connected to anchors through cables, their movement is restricted in a limited area. As nodes move with ocean current, there is no assurance that the selected forwarders will remain within the communication coverage of their respective senders during data transmission. Both approaches allow depth regulation by adjusting the length of the wire used to connect the sensor to the anchor or buoy.

The main objective of AUV based architecture is that these devices rely on their local intelligence and are less dependent on communication from onshore station. However, these autonomous devices require control strategies for coordination, obstacle avoidance and steering strategies. They can be used as temporary relay nodes for anchored sensor devices. If the likelihood of coverage could be improved, 3D architecture with anchored nodes would be more viable for monitoring underwater environment. In addition, the integration of mobile message ferries will enhance its capability and coverage.

2.2. Data delivery

Existing message forwarding protocols such as Ad hoc On-demand Distance Vector (AODV) routing (Perkins and Royer, 1999), Destination Sequenced Distance Vector (DSDV) routing (Perkins and Bhagwat, 1994), Dynamic Source Routing (DSR) (David et al., 2001) for terrestrial ad hoc networks are not suitable for UWSNs due to the different nature of underwater communication medium and network characteristics (Akyildiz et al., 2005; Pompili and Akyildiz, 2009). The existing ad hoc routing protocols are classified into two main categories: (i) proactive, and (ii) reactive. The former incurs a large signaling overhead to establish route for the first time and propagate route update to all the nodes in the network. To reduce such overhead, reactive protocols maintain information about active routes only. However, the reactive protocols require source initiated flooding of control packets for path establishment and also cause a high latency which can be amplified underwater due to slow propagation speed of acoustic signal (Akyildiz et al., 2005). The communication overheads associated with these protocols are extremely high due to the high latency incurred to exchange control packets. Thus, to reduce the control overhead as much as possible in UWSNs, the number of control packets should be minimized. Moreover, the effect of ocean current in 3D environment, which changes the relative sensor position and causes connectivity loss in the network, should also be taken into account.

The message forwarding techniques for UWSNs can be broadly categorized into two classes: (i) location unaware and (ii) location aware. A very few location unaware techniques have been investigated in the literature. Distributed Underwater Clustering Scheme (DUCS) (Domingo, 2011) is a hierarchical routing protocol which does not need geographical information of sensor nodes. The whole network is divided into clusters using a distributed algorithm. To eliminate redundant information, the protocol incorporates an energy-aware cluster head selection algorithm as well as data aggregation. DUCS achieves high packet delivery ratio, considerably reduces the network overhead, and consequently throughput increases. However, the structure of its clusters is affected with node movement caused by the water current, which consequently decreases the cluster life (Ayaz et al., 2011). During the network operation phase, a cluster head can only transmit its data to another cluster head, which restricts its availability and flexibility. This may lead to inter-cluster communication failure as water current moves two cluster heads away. Seah and Tan (2006) proposed a virtual sink architecture where multiple sinks are connected to a network via high speed links. A data packet is sent simultaneously over multiple paths towards multiple sinks that collectively form the virtual sink architecture. Data delivery is considered successful as long as a copy of the packet reaches one of these sinks. This technique does not require the location of sensor nodes to be known. However, it suffers from multiple redundant transmissions that consume critical underwater resources (Ayaz et al., 2011). In Zhou and Cui (2008), a location unaware Multi-path Power control Transmission (MPT) protocol is introduced. The protocol is a cross layer approach which combines the power control with multipath routing. However, MPT suffers from redundant packet transmissions due to multipath routing and incurs large overhead like other path setup based protocols (i.e., reactive protocols mentioned above).

Most existing data delivery protocols for underwater such as Vector-Based Forwarding (VBF) (Xie et al., 2006), Focused Beam Routing (FBR) (Jornet et al., 2008), Directional Flooding-based Routing (DFR) (Shin et al., 2012), Depth-Adaptive Routing Protocol (DARP) (Chen et al., 2012; Erol and Oktug, 2008) are location-based. Most of them reduce the flexibility of the network assuming that sink nodes are located on the sea-surface and their locations are known a priori to all other nodes in the network. The first routing protocol proposed for UASNs is VBF (Xie et al., 2006). Data packets are forwarded along redundant and interleaved paths from the source to sink, which helps handling the problem of packet losses, node failures, and hence increases robustness. Each data packet carries the location of the source, forwarding nodes and the sink. The forwarding path is guided by an imaginary line/vector from source to destination, and nodes that are close enough to that line are candidates for relaying packets. However, the virtual routing pipe from source to destination in VBF affects the routing performance of the network with different node densities as the protocol is very sensitive about the routing pipe radius threshold. In FBR (Jornet et al., 2008), the source node is aware of the location of itself and the destination node, but not of intermediate nodes. Routes are established dynamically during the traversing of data packets from the source towards the destination. Similar to VBF, FBR also does not perform well in sparse areas where no node lies within the forwarding cone of angle. DFR (Shin et al., 2012) is a packet flooding technique, which assumes location knowledge of each node, its one-hop neighbors and destination. Limited numbers of sensor nodes participate in forwarding a packet in order to control its flooding over the whole network. Forwarding nodes are decided according to the link quality to increase reliability. However, multiple nodes can forward the same data packet in areas where the link quality is not good. As a consequence, this increases the consumption of network resources, e.g., energy. DARP (Chen et al., 2012), takes the acoustic speed at different depths into consideration to find the shortest end-to-end delay path. Each sensor node calculates the distances between itself and the sink, and the sender and the sink using their location information. The candidate node which is closer to the sink and the straight line between the sender and the sink has higher priority to forward data. Similarly in Erol and Oktug (2008), a localization and routing framework has been proposed for mobile UWSNs. Localization and routing are done in two consecutive rounds in this approach. A limited number of Mobile Beacon and Sink (MBS) nodes are used for localization and routing. MBSs move vertically in the water column to localize sensor nodes and collect their sensed data. However, the protocol assumes that all nodes are clock synchronized throughout the network which is feasible only for short-term applications, and for long-term applications, it needs additional synchronization mechanisms (Erol and Oktug, 2008).

To limit the use of 3D location information, depth-based routing protocols such as Depth-Based Routing (DBR) (Yan et al., 2008), hydraulic pressure based routing (e.g., HydroCast) (Lee et al., 2010), Depth Based Multi-hop Routing (DBMR) (Prasan and Mahapatra, 2012) and Energy-Efficient Depth-Based Routing (EEDBR) (Wahid and Kim, 2012) have been proposed. In DBR (Yan et al., 2008), data packets are forwarded greedily from higher to lower depth sensor nodes until the data sink is reached. The protocol is able to handle the dynamic networks with good energy efficiency. However, DBR cannot handle void zone, thus fails to achieve high packet delivery rate in sparse areas. The void zone problem of DBR is addressed in another depth-based protocol, namely HydroCast (Lee et al., 2010). The protocol considers wireless channel quality for simultaneous packet reception among the neighbor nodes and this enables opportunistic forwarding by a subset of the neighbors that have received the data packet correctly. Even though this increases data delivery rate but consumes significant network resources such as energy, because redundant copies of the same data packet can be delivered to the destination. However, both DBR and HydroCast are not applicable to horizontal networks where nodes are deployed at the same depth.

Based on DBR, DBMR protocol is proposed in Prasan and Mahapatra (2012). The protocol reduces communication cost by using multi-hop mode at each node. Each node stores the depth information of the next hop in the routing table, and thus the protocol reduces the redundant data forwarding and channel occupancy. Unlike DBR which employs only the depth of sensor nodes, EEDBR (Wahid and Kim, 2012) utilizes both depth and residual energy for forwarding data packets. However, these protocols also suffer from the same disadvantages as those of DBR caused due to the consideration of vertical communication architecture.

3. Path reliability-aware data delivery (PRADD) protocol

In this section, we present PRADD, a path reliability-aware data delivery protocol for UASNs. Before presenting an overview and details on how PRADD works, its proposed network topology is presented in the following section.

3.1. The proposed system model

There are three types of nodes in the proposed architecture as shown in Fig. 1. Ordinary sensor nodes, represented by a number from 1 to 8, are responsible for sensing data. One or more ordinary sensor nodes are selected to act as gateway nodes, indicated by 'GW' (shown as shaded circle) to collect sensor data from other ordinary nodes. One or more mobile ferry nodes (e.g., two ships and one AUV shown in Fig. 1) are utilized to collect data from *GW* nodes within one hop transmission and offload those data to onshore station for further processing. A simplified version of the system model was initially proposed in Nowsheen et al. (2013). As discussed in Section 2.1, both ordinary and *GW* nodes are anchored to the ocean bottom and move with ocean current. The movement trajectory of these nodes remains within the surface area of a hemisphere centered at the anchor and is regulated by the ocean current and other forces applied to it. The network can be represented by a graph $G(V, E)$, where, is a set of n sensor nodes, i.e., $n=/V/$ and Ni is a set of positive integers. $E=\{(i,=j)\ A\ (1 < i,j < n)\}$, where, nodes i and j are within each other's communication range. As mentioned earlier, each link is characterized by three parameters: (i) link transmission reliability (Q_ij), measured as the number of i's data packets successfully forwarded by node j; (ii) reachability to gateway from i, measured as the inverse of the average number of hops (Hij) to reach a gateway using j as the next hop; and (iii) coverage probability (C_ij), defined as the probability that a node j will be within the coverage range of node i at a particular time instant. Node i maintains $\{Q_ij, H_ij, C_ij\}$ for each of its neighbors, j, where, $j\ G\ \{1, 2,m\}$ in its neighborhood table and m denotes the number of neighbors.

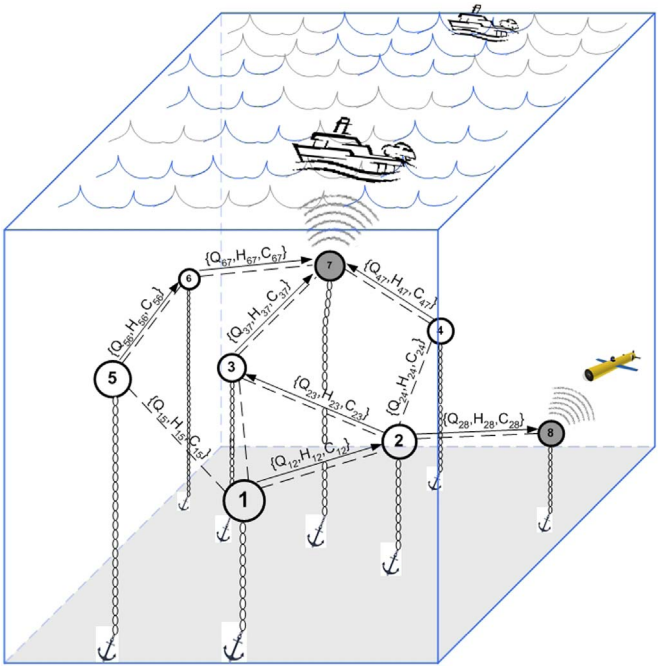

Fig. 1. System architecture and protocol overview.

3.2. Overview of PRADD

In PRADD, mobile ferry node(s) periodically meet the *GW* nodes to collect data stored in their local buffers. A ferry node (i.e., AUV, Boat, Ship) is aware of the deployment topology of the fixed sensor network. First, a ferry node moves around the network, stops at some key points for enough time to gather information about candidate *GWs* that receive beacon from them. A candidate *GW* sends information to the ferry that contains its residual energy, per bit transmission and reception energy, number of source nodes and their data generation rates and neighbors' remaining energy. Based on this information, it selects *GWs* among the candidates (details in Section 3.3). Once selected, a *GW* announces to the rest of the network about its ability to act as a *GW*. An overview of PRADD is shown in Fig. 2. After receiving announcement from *GW*, each ordinary sensor node creates path to the *GW* directly or through intermediate sensor nodes by multi-hop communication. Ordinary sensors and *GW* nodes form a connected sensor network as shown in Fig. 1, which indicates that it is possible to have multiple routes between the source and the GWs. Each ordinary node works in three phases: (i) receiving and storing data, (ii) selection of a forwarding neighbor and (iii) data forwarding to the selected next hop. During the first phase, each node accumulates data packets for a certain amount of time, called "*hold time*". In the second phase, each sensor node initiates forwarder discovery through which the best next hop is selected and sends data through it (details in Section 4). The protocol employs opportunistic forwarding where nodes take independent and local forwarding decision *on-the-fly* at each hop. Each node measures the probability of a forwarder being selected given its link transmission reliability, destination reachability and coverage probability with an aim to ensure path (the routing path between a source node and one of the gateways reachable to it) reliability in data transfer towards the gateways. The link transmission reliability is updated by overhearing next hop forwarder's transmission. Details on link reliability calculation are discussed in Section 3.4. An ordinary sensor node can receive *GW* announcement directly or from its neighbors, and updates its minimum hop count to *GW* from the information of packet. The reachability to the *GW* is estimated from this information. A detail on reachability estimation is discussed in Section 3.5. The coverage probability to neighbors is estimated from the initial coarse location (which is assumed to be known during the deployment time) and the movement information (occurred due to environmental factors) of the neighbors. This is discussed in more detail in Section 3.6. Once a forwarder is selected, each node transmits a train of data packets to the selected next hop. This process continues until data reach one of the *GWs*. Finally, the ferry node collects data from *GWs*.

3.3. Gateway selection and announcement

A ferry node transmits a beacon signal when it arrives close to the network and stops at a key point of its movement path. Upon receiving a beacon signal from the mobile ferries, both ordinary and previously selected *GW* nodes (if any) become candidates to act as *GWs* by sending back their replies to the ferry nodes. It is imperative that the data transmission reliability to the ferry nodes depends on the lifetime of the *GWs*. Therefore, we have developed an algorithm presented in Algorithm 1 to select a set of *GWs* on maximization of respective lifetime considering their own and neighbors' energy resources (e.g.,

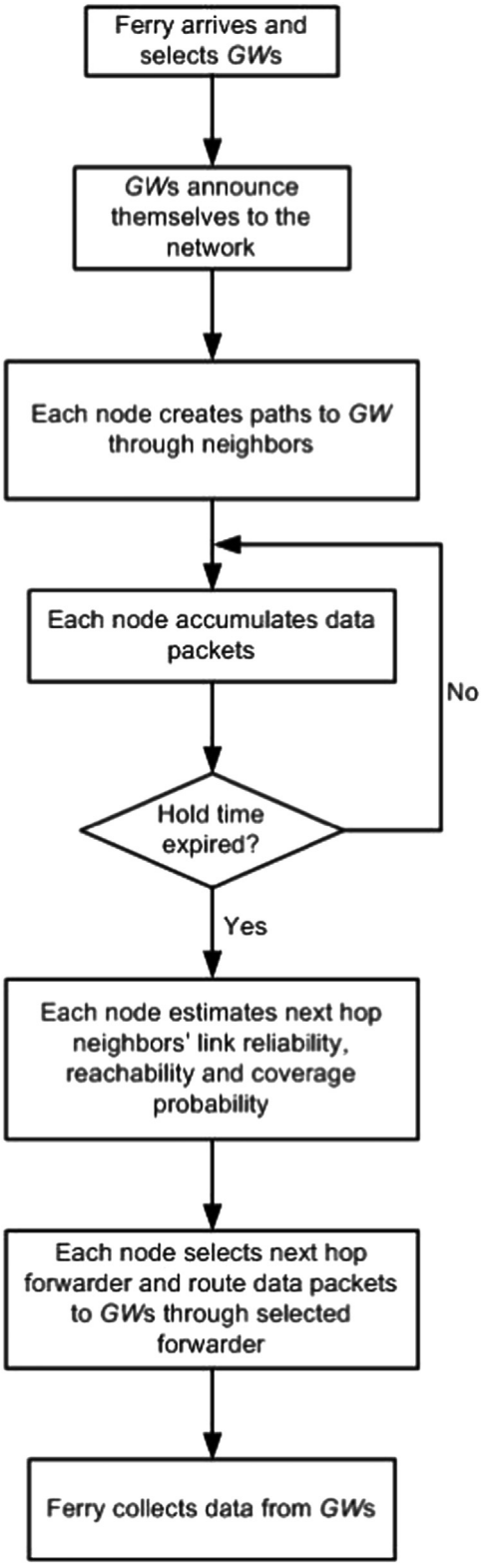

Fig. 2. Overview of PRADD protocol.

remaining energy) and consumption (e.g., data reception and transmission). Demise of neighboring nodes of a *GW* disconnects it from the rest of the network, and therefore, expected lifetime of neighbors should play a part in *GW* selection. Note that the higher the number of *GWs,* the less is the amount of data to be handled per *GW*. The variables used in this algorithm are initialized in Lines 1 and 8. Lines 2–22 articulates the ways to select a set of *GWs* that maximize their lifetime. Line 4 calculates the data that each *GW* needs to handle. The time that a *GW* must remain alive to receive and deliver its collected data is determined in Line 5. In Lines 6–21, a *GW* is added to the existing ones only if it increases the lifetime of other *GWs*. Line 7 determines the expected number of bits that any neighbor of a particular candidate *GW* needs to forward. Line 9-14 calculates the number of neighbors for a particular candidate *GW* that remain connected to the *GW*. Line 10 estimates the time that a neighbor of any particular *GW* requires to receive and forward its data calculated in Line 7. The selected set of *GWs* is termed as *gw*.

Once selected, the nodes in *gw* announce to the whole network about their ability to act as *GW*. Each "ANNOUNCE" packet from a particular *GW* is uniquely identified by < gateway id, announce id >. The message can be received by a node directly or via intermediate nodes and is later used to select multiple candidate forwarders to send packets to a particular GW. After receiving first "ANNOUNCE" packet, each

Algorithm 1. Selection of GWs.

Require:
C_g: A sequence containing remaining energy of the candidate *GWs* in descending order
S_k: A sequence containing remaining energy of all neighbors of kth candidate GW
N_s: Number of source nodes
m_k: Number of kth candidate *GW's* neighbors
n_k: Number of kth candidate *GW's* neighbors required to maintain its connectivity
S_g: Data generating rate of source nodes
E_r: Energy required per bit reception E_t: Energy required per bit transmission
E_{th}: Energy threshold (depends on application/network requirements), which is used to leave a certain amount of energy for sensing and relaying after acting as a *GW*
Ensure: gw: The set of the selected GWs
1: Initialize k=0, gw=ϕ and T=0
2: **while** k < *size*(C_g) **do**
3: k=k+1
4: Calculate the number of bits expected to be received by each GW per unit time,
5: Calculate the time to act as a
6: if temp > T then
7: Calculate the expected number of bits to be received by any of the neighbors of kth candidate GW,
8: Initialize *count* = 0
9: **for** j=1 **to** m_k **do**
10: Calculate time required for jth node of kth candidate GW to act as its neighbor,
11: **if** T_k (j)≥ temp **then**
12: *count* = *count* + 1
13: **end if**
14: **end for**
15: **if** *count* ≥ η_k **then**
16: T=*temp*
17: gw = gw U {k}
18: **end if**
19: **else**
20: **stop**
21: **end if**
22: **end while**

node waits for a pause time t_p to receive more "ANNOUNCE" packets from other neighbors. During this period, a node may receive more "ANNOUNCE" packets and creates paths to GWs by adding an

entry $\{Q_{ij}, H_{ij}, C_{ij}\}$ in the table for each one-hop neighbor j. After t_p expires, each node stops receiving "ANNOUNCE" packets from its neighbors and rebroadcasts the first "ANNOUNCE" packet that it has already received. This reduces the chance of a packet being processed around in a closed loop by preventing a node to receive its own transmitted "ANNOUNCE" packet. When an ordinary node needs to send data to a ferry node, it transmits the data to the GW which has better reliability in terms of path quality.

3.4. Link transmission reliability estimation

As discussed in Section 1, due to the high error-rate of acoustic channel, the link transmission reliability at any particular moment can be changed in near future, and therefore, needs to be re-assessed before a node forwards data after "*hold time*" on that link. To calculate link transmission reliability with the next hop forwarder, each node keeps track of the number of packets sent to and forwarding of those by its next hop node, during its data "*hold time*". As mentioned in Section 3.2, a node i calculates its link transmission reliability Q_{ij}- by overhearing the data forwarding information of its next hop forwarder j. For a given packet P, if the next hop forwarder receives P successfully, it will forward the packet further to its next hop along the path to the *GW*. The node js transmission of this packet heard by previous forwarder i is taken as a passive acknowledgment. The link transmission reliability is calculated for a train of data packets by overhearing their forwarding by next hop neighbors. A good link quality means all packets sent by i are received and forwarded by j and thus overheard by the sender. When the link quality is poor, some of sent packets would not be overheard because either they are not received or their passive acknowledgments are not overheard. Thus, the portion of self-transmitted packets overheard can be an indicative of link quality. At a particular time instant t, if node i wants to send a set of data packets, it first calculates its current link success rate as

where S_{ij} and F_j are the number of packets sent from i to j and j's forwarded of those packets overheard by i, respectively. Since the success of a packet to be forwarded towards the *GW* depends on the quality of its uplink forwarding path, represents the traffic status of the uplink forwarding path as well as its link transmission reliability.

The same equation is applied for each of m candidate forwarders of node i to calculate . Data packets are stored in the local buffer of node i until a suitable next hop forwarder is found. When a node i is willing to send data packets, it broadcasts a "REQUEST" control packet to its one hop neighbors. Each "REQUEST" packet is uniquely identified by a < source id, request id > to track the "REPLY" packet for a particular request. A node j receiving "REQUEST" packet checks if it has valid next hop information to reach the gateway. If it does, it calculates the transmission reliability of all its outbound links. To improve reliability we need to ensure that, other parameters (e.g., reachability, coverage probability) remaining the same, the next hop node j with a link having the highest transmission reliability is selected as a forwarder.

In calculating link reliability of a candidate forwarder j, we consider the quality of its all outgoing links leading towards a *GW*. First, all outgoing links of a node are ranked in descending order of transmission reliability and each ranked value is weighted by a positional value of decimal system, which gives node transmission reliability as

where Q_{jl}; is the transmission reliability from node j to its next-hop node l. For example, a forwarder j can reach *GW* by using m outbound links. After receiving "REQUEST" packet from node i, candidate forwarder j first sorts the transmission reliability of its outbound links in descending order as $Q_{j1}, Q_{j2}, \ldots Q_{jm}$, each with one digit precision after the decimal point. In this order, the kth link has a multiplying factor of 10^{m-k}. Node j then calculates its uplink transmission reliability as in Eq. (2).

After calculating Q_j, node j sends it to node i through its "REPLY" packet. As the *GW* does not forward the packet until a ferry is met, the reliability of its all inbound links is updated from an "ACKNOWLEDGMENT" packet received from it. Node i receives Q_j for each candidate forwarder j and identifies the forwarder having the maximum number of outbound links. Let the maximum number of outbound links be m'. For neighbors having lesser number of outbound links than m', their Q_js are padded with zeros at the right so that the lengths of all Q_js are equal to m. Therefore, Q_j is modified by node i as

This strategy ensures that, the node with an outbound link having the highest value in ranked order will always be selected, and in case of a tie, the one with the highest subsequent ranked value is selected. For example, a node i has two candidate forwarders 1 and 2, where, forwarder 1 has three outgoing links with transmission reliability 0.9, 0.6, 0.6, respectively, while forwarder 2 has two outgoing links with transmission reliability 0.8 and 0.7, respectively. Both forwarders 1 and 2 calculate their uplink transmission reliability as, $Q_1=0.9\times10^2+0.6\times10^1+0.6\times10^0=96.6$ and $Q_2=0.8\times10^1+0.7\times10^0=8.7$. Therefore, after receiving Q_1 and Q_2, node i estimates m =3. As, Q_2 has one less digit than Q_1, node i recalculates Q_2 as, $Q_{\mathbf{2}}=8\ 7\times10^1=87.0$. The above example vindicates that $Q_{\mathbf{1}}$ will be selected as the next hop forwarder as it has the outgoing link with the highest transmission reliability (0.9). The value of Q_j calculated in Eq. (3) is then normalized to a range between 0 and 1 as

The transmission reliability Q_j also represents the data transmission reliability of uplink forwarding path. Node i can receive many "REPLY" packets from its neighbors for a particular < source id, request id > and calculates the link reliability of each candidate forwarder j and its uplink forwarding path as

where is defined in Eq. (1). For the remaining sections of this paper is assumed as . Candidate forwarder j also sends its initial coarse location and movement information in the "REPLY" packet to let i know whether it will be within the communication range of i after a certain time interval. The expected number of hop count to reach *GW* using j is also incorporated in the "REPLY" packet. All of the information is used to select a forwarder from the candidate ones in Section 4.

3.5. Reachability estimation

The reachability in the context of a data communication network can be defined as whether any routing path exists from the source to the final destination nodes, i.e., whether a packet generated by a source can reach the destination. However, reachability in this paper is defined in terms of hop count required to reach the destination. If other relevant factors, such as, reliability and coverage of all paths are same, the higher the hop count, the lower is the reachability of that path. If there is no path from a source node to the destination node, the required hop count is assumed to be infinity, and hence its reachability is zero. In our scheme, the number of one-hop steps to reach the destination is already known from the neighborhood table consisting of during the *GWs'* announcements (referred to Section 3.3). A candidate forwarder may have many outgoing links, each having different hop counts to reach the destination. Since we do not know in advance which of the outgoing links of the selected neighbor will be used to forward data, neither minimum nor maximum hop counts of all outgoing links can be considered as they represent the best and worst case scenarios, respectively. For this reason, we need to use the expected value of hop count for a next hop during its selection process. The expected number of hop count for a data packet to reach *GW* from j would be the expected value of the hop counts required for its all outbound links. This can be defined as

where

andwhere represents the probability of l being selected by node j as the next hop forwarder, is the number of hops required to reach *GW* from node j using l as the next hop, and m is the number of outbound links of j. The calculation of is detailed in Section 4. Node j calculates Hj and sends it to node i as a part of "REPLY" packet, which then calculates its required hop count to reach a *GW* using forwarder j as

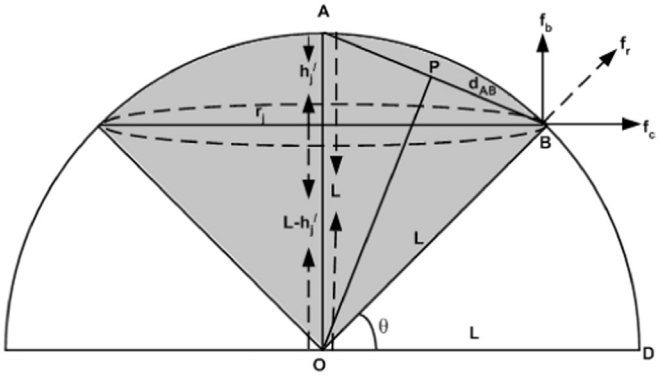

Fig. 3. Movement area of an underwater sensor node.

3.6. Node movement and coverage probability estimation

Now that we have an understanding of the basis of the proposed data forwarding solution, we proceed to the detailed discussion of node movement and coverage probability estimation technique. The coverage probability (C_{ij}) indicates whether node j is within the coverage range of node i and how close it is after its movement. This section calculates C_{ij} for any candidate forwarder j from its estimated displacement due to movement. Sensor nodes anchored to the ocean bed float in the water and move with water current along the surface area of a hemisphere as shown in Fig. 3. Here, A is the initial position of any sensor node j in the water and O is the anchor position in the ocean bottom. The initial position of each sensor node can be estimated from the coarse location of its anchor during deployment and the ocean current and other factors at that time. In a 3D sensor network, the position of each node is computed by its xyz co-ordinates. The motion of underwater mobile node is mainly affected by ocean current. However, for quasi-stationary nodes, the movement is impacted by the ocean current but limited by the tension of the cable. There are four forces: gravity, buoyancy, water current and tension of the cable that control the displacement of a node. The gravitational force and tension force are downward, while the buoyant force is upward. The buoyant force must balance the sum of the tension and gravitational forces to keep the sensor node in its place, i.e., the magnitude of the buoyancy force , where is the gravity force and is the force due to tension of the cable. Hence, the buoyancy control keeps the tension of the cable straight and thereby, keeps the movement of an anchor node limited to the surface of a hemisphere.

Each candidate forwarder j calculates its displacement along the surface area of a hemisphere centered at the anchor position. As shown in Fig. 3, is the time-varying ocean current which acts along the x–y plane and is the upward buoyancy force along the z-axis which pushes the node up towards the surface. Hence, and are always perpendicular to each other irrespective of the direction of . The magnitude and direction of these forces are used to calculate the movement of the candidate forwarder j at a particular time instant. The upward buoyancy force of a sensor node is defined as (Munson et al., 2009)

where p is the density of water, g is the gravitational acceleration and V is the volume of a sensor node. In Fig. 3, the dotted line indicates the circumference of the base of a spherical cap along the x–y plane. As shown in Fig. 3, after movement, the position (indicated by B) of the sensor node j can be located on the circumference.

Movement of objects depends on many environmental factors such as water current and water temperature (Zhou et al., 2011). Mobility characteristics of underwater object is different in different underwater environments. For example, the movement of an object shows certain semi-periodic property in sea-shore environment. This is mainly due to tide and bathymetry in that particular region (Zhou et al., 2011). The kinematic model proposed in Beerens et al. (1994) can be used to model node movement in such environment. However, it cannot represent the nature of ocean current in all environments.

As there is no generic model to represent the ocean current at all places, the mobility of underwater node in a particular region can be estimated from the velocity of the current of that area. Oceanographers can use tools such as Acoustic Doppler Current Profiler (ADCP) (Acoustic doppler current profiler) to monitor the ocean current. The device can calculate the speed and direction of the ocean current from the frequency of the acoustic signal, the distance it travels and the time it takes to travel. A velocity data set of ocean current can be found in (Measuring deep ocean currents, 2006) where the global ocean was observed with more than 2400 profiling floats. Researchers at IPRC's Asia-Pacific Data-Research Center (APDRC) (Asia-pacific data-research center) developed a method by which they estimated the velocity of ocean current using these floats. The instantaneous velocity of ocean current at a particular time can be estimated by using a sensor package that uses information from an electromagnetic current meter, a pressure gauge and data from a three-axes accelerometer to determine the directional velocity field in a cost effective way (Elwany and Mahr, 2003). The instrument can be integrated to an underwater node to provide immediate information about wave characteristics in deep water. Once the velocity (both speed and direction) of the ocean current v_x and v_y can be found either from accelerometer or by any other means, the force along the x- and y-axes can be calculated as (Bagtzoglou and Novikov, 2007)

where C is the drag coefficient of the sensor node and A' is its effective area hit by water force. The magnitude and direction of the resulting force (shown in Fig. 3) can be expressed as, and .

The radius of the hemisphere is controlled by the cable length (L) and anchor position (O). The angle (θ) of the resultant force as shown in Fig. 3 can be calculated as

As alluded before, sensor node j moves from its original position to a new position on the surface area of the hemisphere due to and . The new displacement of node j along the surface area of the hemisphere can be calculated as

where L is the cable length of sensor node.

For each new movement, a spherical cap can be formed by cutting the sphere off along the x–y plane according to node displacement. Suppose the hemisphere is cut into a number of spherical caps each with different radius and height. If h is the height of a cap, the surface area of the spherical cap is denoted by . For node j, the height of a particular spherical cap with displacement , defined in Eq. (12), can be expressed as

The candidate forwarder calculates its new estimated location as

In the similar way, the new location of node i can be estimated. The candidate forwarder then sends a "REPLY" packet to node i with its H_j, Q_j and new location at the time of data transmission. The direction of v_x and v_y will determine the sign of and co-ordinates, respectively. After receiving "REPLY" packet from a candidate forwarder j, node i calculates whether j is within the coverage range of node i. The Euclidean distance between nodes i and can be calculated as

If the communication radius of i is R, the coverage probability of j is expressed as

The higher the value of , the higher the probability that node j will be within the coverage range of node i at a particular time instant.

4. Next hop selection and data forwarding

As mentioned in Section 3, PRADD ensures path reliability by selecting the next hop forwarder at each hop towards *GW(s)* based on three parameters: (i) , (ii) and (iii) of neighbors at a particular time instant. Node i has already for each of its m neighbors. There are many ways (e.g., weighted average, joint-estimation, probabilistic method, etc.) to select the next hop forwarder. Since the underwater acoustic communication link is error-prone and dynamic, we need to consider both current and historical information of a link in its selection process. In this case, probabilistic method, namely Bayes' theorem is

the most suitable candidate as it can efficiently evaluate the conditional probability and incorporate historical information. Details of next hop forwarder selection process using the Bayes' theorem and data forwarding techniques are discussed in the following sections.

4.1. Phase I: Local next-hop forwarder selection

For an environment which is susceptible to contextual changes, the local approach is better than the global approach. Consequently, the hop-by-hop communication is better than the end-to-end communication for underwater (Cui et al., 2006) and hence, we have exploited the local information i.e., the information up to two-hop neighbors to select the next hop forwarder. The calculation of the posterior and likelihood probabilities of a next hop neighbor considering its local information is presented in the following sections.

4.1.1. Posterior probability estimation

The conditional (posterior) probability of a neighbor j of node i being selected given using the Bayes' theorem can be defined as

The interest is only in the numerator of the fraction because the denominator is used only for normalization purpose (Rish, 2001). According to the "naive" conditional independence, each feature , and are conditionally independent with each other. Then

Eq. (20) is evaluated for each candidate forwarder j. The candidate forwarder having the maximum posterior probability is selected as the next hop. In a particular time interval, if a packet is inferred as a successful transmission, entries and are updated and stored in a neighborhood table. The posterior probability is based on likelihood probabilities and and prior probability $P(j)$ and they need to be estimated before the calculation of Eq. (20).

4.1.2. Likelihood probability estimation

To calculate the likelihood probabilities related to node j, node i takes into account of the current status of the link as well as its deviation from the historical mean. Factors affecting undersea environment, e.g., sound speed fluctuation, surface waves, turbulence, volcanic eruption and other phenomena cause random signal variation (Stojanovic and Preisig, 2009). Therefore, we assume link quality to be normally distributed. For node i, the probability of the link quality given node j can be defined as

Initially when time is the normalized value of the current link quality . But for the other time is defined as the normalized weighted value of . The historical information is encoded here as, , i.e., the probability that the historical mean of the link quality becomes greater than its current value . is a weight for the link quality and can be regarded as a reward or punishment. Similar to the credit assessment for a loan application in the banking sectors, the links that exhibit better quality are rewarded, while the others are punished. A lower value of indicates punishment. Considering the fact of social norms for equity, we cannot assume its fixed value, rather the reward/punishment should reflect the degree of creditability that a link deserves. As a consequence, it should be proportional to the value of . The higher the value of , the higher the reward of that link. For this reason, can be defined aswhere Z is defined as

The average link quality is estimated using the exponential moving average (EMA) aswhere the coefficient $0 < \alpha < 1$ makes a trade-off between the current and previous values of . The higher the value of α, the more importance is on the recent value of link reliability, while its smaller value gives more emphasis on historical information. The value of a depends on the amount of the past information intended to be considered and can be derived as, $\alpha=2/(N+1)$, where, N is the observation window size. Since in is compared with , the standard deviation of all historical link quality values, σ_{ij} is calculated from as

As mentioned before, the reachability is inversely proportional to the hop count, Hij**.** Therefore, following the similar way of defining Eq. (21), the probability that a packet from node i will reach to a gateway via node j can be formulated as

where . Similar to (Eqs. (23) and 24) for , the formula for Z and the average hop count can be formulated. In the same way of defining in Eq. (21), the coverage probability between nodes i and j can be expressed as

The equations for can be derived in the same manner as defined in (Eqs. (22) and 24), respectively.

4.2. Phase II: Data forwarding

After forwarder discovery, a node starts transmitting its accumulated data packets to the selected forwarder. For any node, if it does not receive a "REPLY" packet from candidate forwarders, it will re-initiate forwarder discovery process by sending another "REQUEST" packet until the maximum number of transmission attempts is exceeded. Data packets are forwarded until one of the *GWs* is reached. Finally, data collected by *GWs* are delivered to the ferry node for offloading.

5. Performance evaluation

5.1. Simulation settings

Experiments were conducted to evaluate the performance of PRADD against the existing and contemporary underwater protocols such as VBF (Xie et al., 2006), DBR (Yan et al., 2008) and DBMR (Prasan and Mahapatra, 2012) using three widely used metrics: (i) packet delivery ratio (PDR), (ii) routing overhead, and (iii) energy consumption. PDR is regarded as the ratio between the number of packets successfully received by the gateways and those sent by the source nodes. Routing overhead is calculated as a portion of the total traffic that is expended as control messages. Energy consumption is the total energy spent for transmission and reception of data and control packets. To reflect the characteristics of underwater acoustic environment in the simulation, the channel modeled articulated in Harris and Zorzi (2007) was used in the physical layer of ns2 simulator. The ns2 implementation includes the propagation time model, the bandwidth-distance relationship, attenuation and the signal to noise ratio for 3D UASNs. The path loss is calculated based on the spreading loss and the absorption loss in this study. The Thorp's approximation for absorption for the frequency range of interest defined in Brekhovskikh and Lysanov (2003) was used in the simulation.

For all experiments in this paper, Constant Bit Rate (CBR) application traffic model was used as the application layer protocol. The total transmission loss due to absorption and spreading experienced at a particular distance was calculated using the Urick propagation model articulated in Urick (1983). Non-persistent Carrier Sense Multiple Access (CSMA) protocol was used for medium access control (Forouzan, 2003). Physical layer characteristics as well as other simulation parameters are listed in Table 1. Studies have shown that the velocity of ocean current remains within 0–2.5 m/s (Jaffe and Schurgers, 2006; Zhou et al., 2011). The pause time, t_p for "ANNOUNCE" packet was set as 1 s. Each simulation ran for 900 s and for all protocols, a set of similar scenarios was used. For all

Table 1
Simulation settings.

Network parameters	
Propagation speed	1500 m/s
Transmit power	2 W
Antenna model	Omni-directional
Packet size	200 bytes
Initial energy	1000 J
Traffic model	CBR
Volume of deployment	450×450×450 m^3
Deployment strategy	Uniform

topologies, the results were averaged from 10 simulation runs.

5.2. Simulation results

To assess the scalability and the effectiveness of PRADD in a variety of applications such as pollution monitoring, weather forecasting and so on, we evaluated each metric for a wide range of packet generation intervals (interval between packets are generated by source nodes), number of data sources and network sizes. The values of the physical layer parameters that reflect a high frequency short range acoustic channel were chosen based on the work done in Nowsheen et al. (2010a, 2010b). The communication parameters were set based on a high data rate acoustic modem with data rate of 100 kbps and the maximum transmission range of 120 m.

5.2.1. Network load

An application may need to sense multimodal data or different applications may require sensing of different attributes (e.g., temperature, water salinity, sound, picture, etc.). Therefore, different applications may have different data generation rates. Packet generation rate increases in reality when the activity of monitored event increases. For this reason, in this set of simulations, the impact of varying packet generation interval (i.e., the inverse of data generation rate) on each of the performance metrics was evaluated. In each case, the network size was set to 50 nodes, and at a given instance, 5 nodes were randomly selected as the data generating sources. Packets were generated at an interval varying from 1 s to 10 s. This interval is commensurate with the traffic generation advertised by different applications in UASNs. However, this value may vary depending on the application requirements and the type of parameters sensed (Mohamed et al., 2011).

Fig. 4 shows the PDR of VBF, DBR, DBMR and PRADD at different packet generation intervals. In all cases, the PDR is relatively low when the data generation interval is small (i.e., high data generation rate). However, it improves with increasing data generation interval. The reason is that faster generation rate increases the chance of collision in the network and thus the PDR drops. The performance of VBF and DBR is almost similar because they use broadcast technique for forwarding data packets. This enhances the chance of collision as packets are transmitted more frequently and redundantly which causes PDR reduction.

The PDR of DBMR falls below that of the VBF and DBR as the packet generation rate decreases. This is because at slower generation rate, the already known next hop neighbors in DBMR can move out of range due to water current before the next data packets are actually forwarded, giving a false impression about their availability. On the other hand, our protocol achieves high PDR (e.g., 0.73–0.94 (PRADD) vs 0.47–0.68 (DBMR)) as it exploits the concept of DTN and selects the

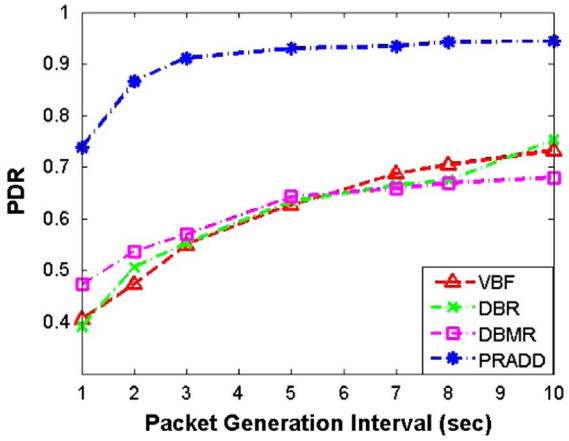

Fig. 4. PDR vs packet generation interval.

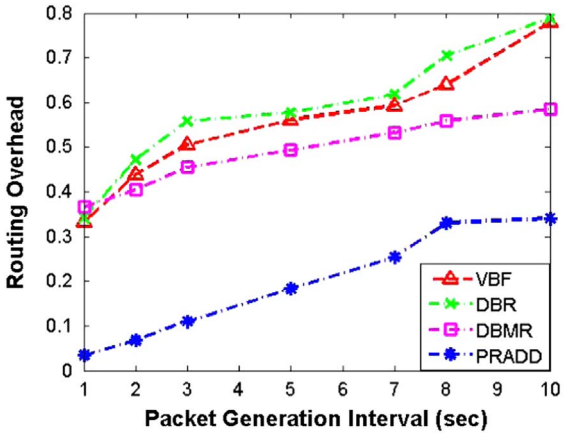

Fig. 5. Routing overhead vs packet generation interval.

next hop opportunistically to provide better data transfer reliability, coverage probability and reachability to the destination. Even at high network load (at 1 s interval), our scheme achieves considerably higher PDR (0.73) than the other protocols. During the forwarder selection process, reachability and reliability of a next hop link are statistically estimated using their historical information. Furthermore, our scheme selects the forwarder that has better coverage probability which aids to ensure that the next hop node will remain within the coverage (communication range) during data transmission. This increases the chance of reliable data delivery and consequently the PDR.

Fig. 5 shows that the routing overhead of all the protocols increases with higher packet generation interval. However, VBF and DBR suffer from higher overhead than DBMR, while PRADD yields least overhead. The reason behind this is that each data packet in VBF and DBR contains routing information (i.e., 3D location in VBF, depth in DBR). The lower the packet generation rate, the lower the chance of packet collisions. Less collision results in higher data delivery which is evident from Fig. 4, but the broadcasting nature of these protocols generate more redundant packets in intermediate nodes creating greater overhead than other protocol. The routing information in each data packet as well as redundant packets contribute to increase the overhead in these protocols. Though DBMR supports unicast routing, its route discovery process incurs more routing overhead. As mentioned before, this occurs because in-between two consecutive data generation time, a next hop could be moved out of the communication range due to water current creating a false impression of a valid route. This is further compounded by the absence of any mechanism in the protocol to recover from this situation. The significantly lower overhead in our scheme is due to the fact that PRADD transmits a stream of data packets for the same set of control packets (request-reply) and thereby, reduces routing overhead per packet, which coupled with higher PDR further improves the overall overhead. This higher PDR indicates more successful data forwarding which reduces the overhead (ratio between routing packets and data packets).

Fig. 6 shows the total energy consumption in each protocol with respect to the packet generation interval. In all cases, the energy consumption decreases as the packet generation interval increases. As the packet generation rate gets slower, the overall network load decreases causing the entire network to consume less energy. From Fig. 6, we can see that the decreasing rates of VBF and DBR are higher than the other protocols due to their broadcast nature in data forwarding. On the other hand, DBMR does not take void zone (Yan et al., 2008) into account and generates routing messages frequently (i.e., spends more energy) when void zone is encountered. However, as shown in Fig. 6, the energy consumption of PRADD is less than other three protocols. This is due to the following reasons: (i) consideration of reachability in our protocol tends to select shorter route to gateway, which means less energy dissipation in transmission and reception;

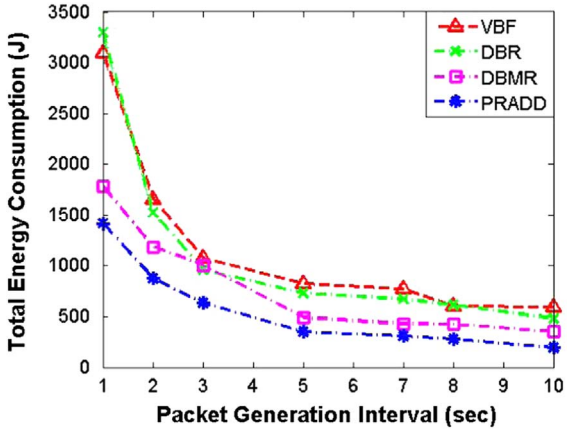

Fig. 6. Energy consumption vs packet generation interval.

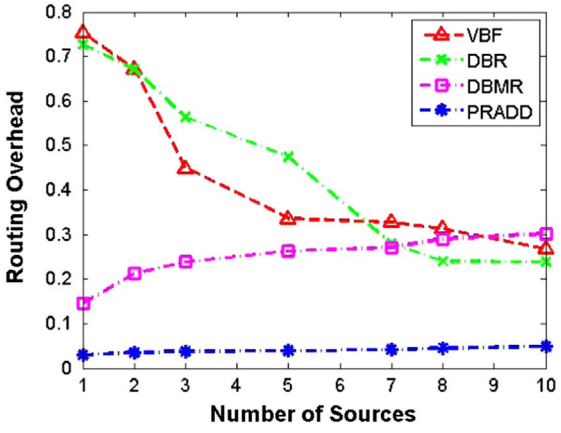

Fig. 8. Routing overhead vs number of sources.

and (ii) less routing overhead in our protocol also contributes to energy saving. Options (i) and (ii) combinedly indicate that PRADD achieves higher PDR with less number of transmissions. The combined effect of all these factors results in energy saving in this protocol.

5.2.2. Source density

In this set of experiments, we investigated the performance of all protocols for varied number of data sources. Depending on the types of applications, the number of data generating sources may vary. For example, when multiple distributed events occur in a region of ocean, many sensor nodes can detect them and report their sensed data to the gateway nodes. For this purpose, the network size was set to 50 nodes and the number of active sources varied from 1 to 10. Each source generated one packet per second. The other parameters were kept the same as specified previously.

In Fig. 7 all the protocols show a decreasing trend in PDR as the number of source nodes increases. As the number of sources increases, so does the network traffic, which in turn leads to increased packet loss mainly due to collision. But the decrement in PDR is considerably higher in other three protocols than PRADD. The superior performance of PRADD is due to its operational characteristics as articulated in the explanation of the results for PDR vs packet generation interval (shown in Fig. 4). Results shown in Figs. 4 and 7 substantiate that our protocol is more effective than the others for applications that require a high number of geographically distributed data sources.

The routing overhead presented in Fig. 8 for VBF and DBR decreases with increasing number of sources. This is because as the number of sources increases, packet collisions increase; this in turn implies that fewer packets are being re-broadcasted which reduces packet redundancy in VBF and DBR and thus, the overall routing overhead. Similar phenomenon is also observed in Fig. 5 which shows lower overhead with smaller packet generation interval, i.e., at high load. The routing overhead for DBMR rises when the number of data sources grows. This is because more and more control packets are needed to find the increased number of routing paths required as the data sources increase. This produces more routing overhead for DBMR. Though the routing overhead in our protocol increases only slightly, it has far less overhead than others. PRADD is based on the mechanism which uses only one routing packet exchange for a bunch of data packets and hence, shows little increase with increasing sources.

Fig. 9 shows the total energy consumption of all the protocols for varied number of data sources. The energy consumption increases in all cases due to the increased traffic load. But the increase is significantly higher in other three protocols than ours. The energy efficiency of PRADD is achieved due to the same factors that attained its energy saving compared to others in Fig. 6.

5.2.3. Scalability

In this experiment, we evaluated scalability of the protocols by increasing the network size. In reality, it is expected that many underwater sensor network applications will require deploying a large number of sensor nodes over a certain geographical area of interest. We ran the simulations varying the network size from 5 nodes to 100 nodes. In this experiment, 20% nodes were randomly chosen as active data sources to generate data at an interval of 1 s. Thus network density was kept same for different network sizes, i.e., the geographic area ($150\times150\times150$ m^3 to $600\times600\times600$ m^3) of the network was increased proportionally according to the number of nodes.

As shown in Fig. 10, the PDR decreases as the network size increases for all the protocols. This is due to the similar reasons

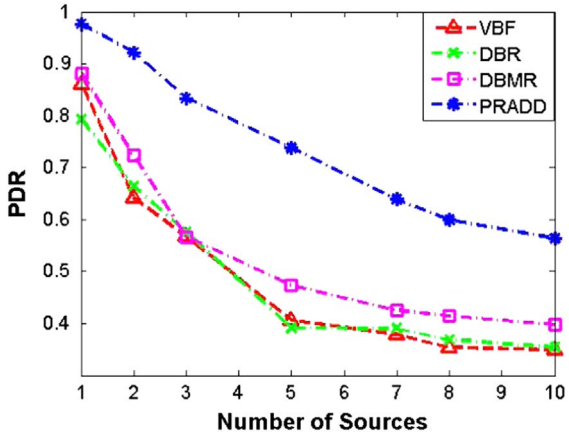

Fig. 7. PDR vs number of sources.

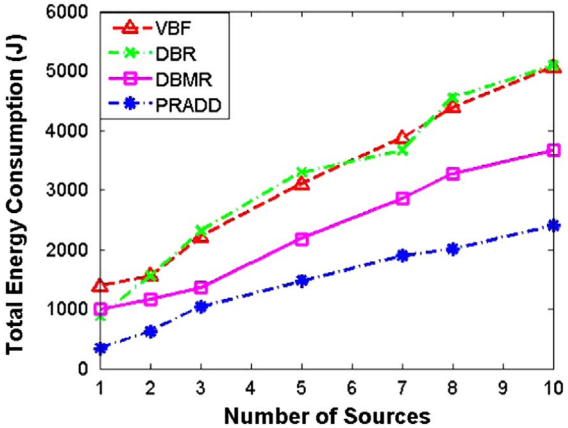

Fig. 9. Energy consumption vs number of sources.

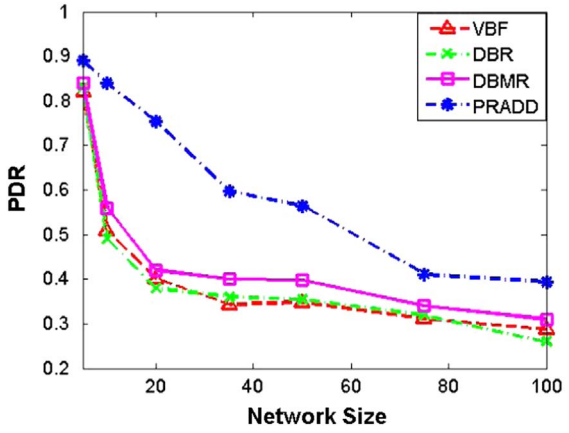

Fig. 10. PDR vs network size.

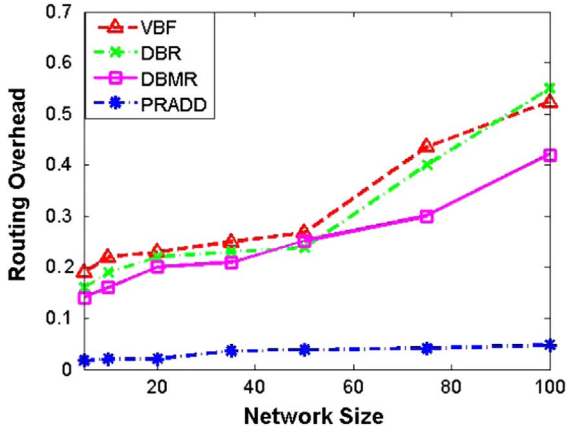

Fig. 11. Routing overhead vs network size.

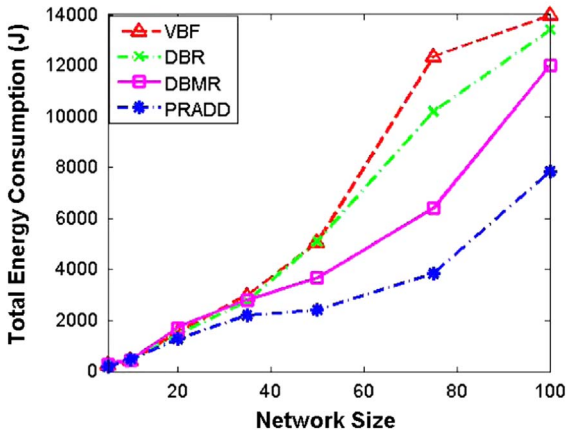

Fig. 12. Energy consumption vs network size.

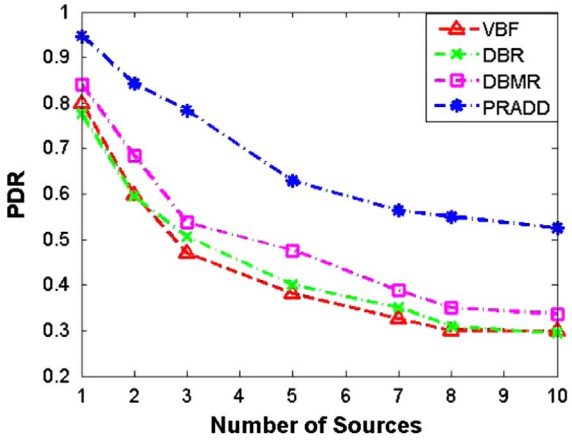

Fig. 13. PDR vs number of sources.

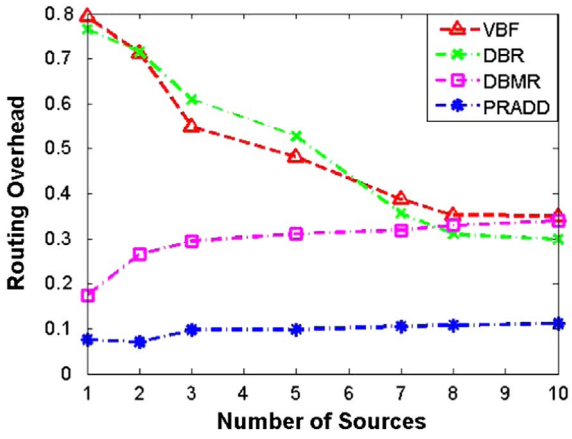

Fig. 14. Routing overhead vs number of sources.

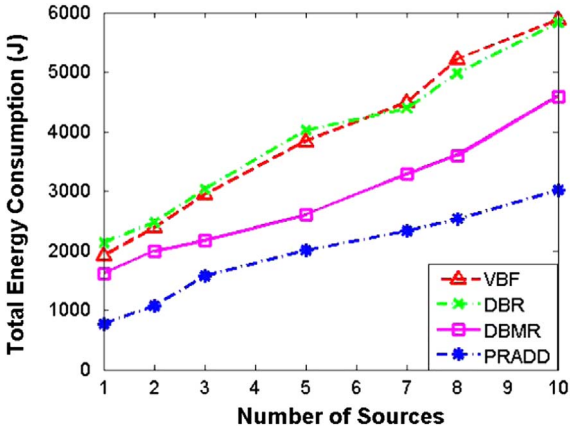

Fig. 15. Energy consumption vs number of sources.

explained in Fig. 7. However, PDR is always considerably higher in PRADD than the others. This is due to the next hop selection criteria (i.e., link quality, reachability, node movement consideration) that enable the protocol to select routes that are more reliable and robust than those selected by other protocols, yielding more successful data delivery.

Fig. 11 shows that the routing overhead increases as the network size increases for all protocols. This is due to the increased size of the collision domain i.e., as the network size increases, more nodes are likely to be far from the gateway and data packets need to travel through more hops to reach the gateway which increases its chance of collision. However, the rate of increase in PRADD is very small compared to others. The reason is that PRADD employs back to back data forwarding with less number of control packets exchange, and therefore, overhead is reduced. Fig. 11 also indicates that our protocol is more scalable than the others. In contrast to the overhead graph in Fig. 8, the overheads of VBF and DBR increase in Fig. 11. This is because in Fig. 11 at larger network size, collisions mainly occur when packets converge near the gateways. This means that in case of VBF and DBR, data packets are replicated multiple times at each intermediate node and whether or not they reach a gateway, these extraneous transmissions increase the routing overhead. On the other hand, in Fig. 8 at higher source density, for VBF and DBR, collisions mainly occur near the source nodes because the source nodes are

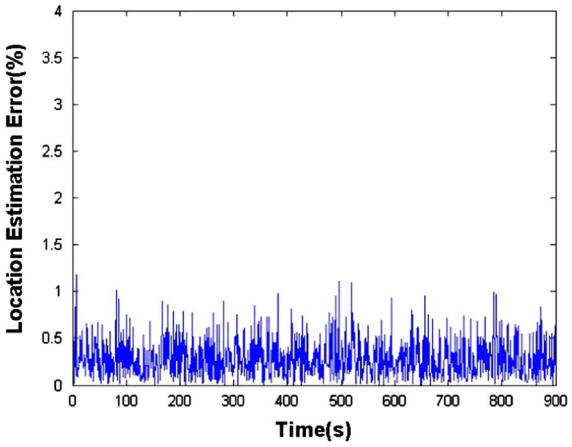

Fig. 16. Relative location estimation error vs simulation time.

situated much closer to each other geographically comparing with that for Fig. 11. This implies that packets collide at an early stage of their traversal, have less chance to traverse much farther towards gateway, and thereby, are less likely to be hugely replicated, which eventually reduces the overhead. The effect of network size on the performance of energy consumption is shown in Fig. 12. As the network size increases, there are more traffic in the network and additionally data delivery suffers from increasing overhead; the combined effect results in greater dissipation of energy as the network becomes bigger. However, the energy increase in VBF and DBR is higher than that of others while PRADD consumes the least amount of energy.

We performed paired *t*-test comparing PRADD with other protocols for PDR, routing overhead and energy consumption. In all cases, the performance improvement by PRADD is statistically significant at 95% confidence level (p is less than 3.2×10^{-3}, 1.1×10^{-3} and 0.0436 for PDR, overhead and energy consumption, respectively).

5.2.4. Simulation results using commercial acoustic modem parameters

We also tested the performance of PRADD according to the communication parameters of a commercial acoustic modem, LinkQuest UWM1000 (Linkquest) with data rate of 10 kbps, and maximal transmission range of 200 m while keeping the values of other parameters the same as provided in Table 1.

Fig. 13 shows the PDR of all protocols for varied number of data sources. PDR decreases in all cases as the number of sources increases. However, the decrement in PDR is considerably higher in other three protocols than PRADD due to the same reasons explained for Fig. 7. However, since at low frequency the communication and interference ranges are higher, the chance of collision increases. As a consequence, the overall PDR decreases. The routing overhead presented in Fig. 14 decreases for VBF and DBR but increases for DBMR with the number of sources. The reason can be explained in a similar way as discussed for Fig. 8. However, PRADD shows a slight increase, but significantly less routing overhead than other protocols. This graph follows the same trend as Fig. 8 and can be explained in a similar way as the graph for Fig. 8. However, the higher transmission range increases the chance of collision in the network, which reduces the PDR and increases routing overhead for all protocols. The total energy consumption of all protocols for varied number of data sources is shown in Fig. 15. The energy consumption increases with the number of sources for all of them. However, at low data rate, they require more transmission time, yielding more energy consumption than that shown in Fig. 9.

Fig. 16 shows error in relative location estimation using our mobility prediction technique. It shows that the error is mostly below 1% for the entire simulation time. The results indicate the efficacy of our method in predicting instantaneous location which has impact on improved PDR, message overhead and overall energy consumption as illustrated in results presented earlier.

6. Conclusion

In this paper, a path reliability-aware data delivery protocol has been proposed to improve reliability in data transfer for delay tolerant applications in large-scale UASNs. The path reliability is achieved by ensuring the best next hop selection at every step. To select next hop forwarder, the protocol utilizes link transmission reliability, reachability to the destination and coverage probability of sensor nodes as routing metric. During next hop selection, PRADD considers both the current and historical information of a link to adapt to the dynamic and error-prone nature of underwater acoustic channel. To improve reliability and maximize lifetime, gateway selection strategy considers the residual energy of candidate nodes and the amount of data generated by source nodes. In PRADD, hop-by-hop reliability is considered to improve data delivery while keeping routing overhead low. Moreover, the protocol only relies on initial coarse location of sensor nodes and does not require the support of active localization technique. PRADD estimates node movement using ocean current, buoyancy and other relevant forces, which in turn is used to measure the quasi-location of sensor nodes. Estimated locations are then utilized to calculate a node's coverage probability. Simulation results in ns2 with realistic underwater scenarios confirm that PRADD outperforms other competing protocols in all the scenarios in terms of successful packet delivery, energy consumption and routing overhead. It also demonstrates better scalability with respect to network size and traffic volume. Future research aims at employing dynamic data "*hold time*" by optimizing protocol parameters at intermediate nodes for further performance improvement of our protocol.

References

Akyildiz, I.F., Pompili, D., Melodia, T., 2005. Underwater acoustic sensor networks: research challenges,. Ad Hoc Netw. 3 (3), 257–279.

Akyildiz, I.F., Pompili, D., Melodia, T., 2004. Challenges for efficient communication in underwater acoustic sensor networks. ACM SIGBED Rev. 1 (2), 3–8. http://dx.doi.org/10.1145/1121776.1121779.

Akyildiz, I.F., Pompili, D., Melodia, T., 2007. State of the art in protocol research for underwater acoustic sensor networks. ACM SIGMOBILE Mob. Comput. Commun. Rev. 11 (4), 11–22. http://dx.doi.org/10.1145/1347364.1347371.

Ayaz, M., Baig, I., Abdullah, A., Faye, I., 2011. A survey on routing techniques in underwater wireless sensor networks. J. Netw. Comput. Appl. 34 (6), 1908–1927. http://dx.doi.org/10.1016/j.jnca.2011.06.009.

Acoustic doppler current profiler. URL ⟨http://oceanexplorer.noaa.gov/technology/tools/acoust_doppler/acoust_doppler.html⟩.

Asia-pacific data-research center. URL ⟨http://apdrc.soest.hawaii.edu/data/data.php⟩.

T. Bian, R. Venkatesan, C. Li, 2009. Design and evaluation of a new localization scheme for underwater acoustic sensor networks. In: Proceedings IEEE Global Telecommunications Conference (GLOBECOM), pp. 1–5. http://dx.doi.org/10.1109/GLOCOM.2009.5425366.

Beerens, S.P., Ridderinkhof, H., Zimmerman, J.T.F., 1994. An analytical study of chaotic stirring in tidal areas. Chaos Solitons Fractals 4 (6), 1011–1029. http://dx.doi.org/10.1016/ 0960-0779(94)90136-8.

Bagtzoglou, A.C., Novikov, A., 2007. Chaotic behavior and pollution dispersion characteristics in engineered tidal embayments: a numerical investigation, JAWRA. J. Am. Water Resour. Assoc. 43 (1), 207–219. http://dx.doi.org/10.1111/j.1752-1688.2007.00017.x.

Brekhovskikh, L.M., Lysanov, Y.P., 2003. Fundamentals of Ocean Acoustics (Modern Acoustics and Signal Processing) 3rd edition. Springer.

Cui, J.-H., Kong, J., Gerla, M., Zhou, S., 2006. The challenges of building mobile underwater wireless networks for aquatic applications. IEEE Netw. 20 (3), 12–18. http://dx.doi.org/10.1109/MNET.2006.1637927.

Chitre, M., Shahabudeen, S., Stojanovic, M., 2008. Underwater acoustic communications and networking: Recent advances and future challenges. Mar. Technol. Soc. J. 42 (1), 103–116. http://dx.doi.org/10.4031/002533208786861263.

Y.-D. Chen, C.-Y. Lien, C.-H. Wang, K.-P. Shih, 2012. DARP: a depth adaptive routing protocol for large-scale underwater acoustic sensor networks. In: Proceedings of the IEEE OCEANS, pp. 1–6. doi: 10.1109/OCEANS-Yeosu.2012.6263505.

V. Chandrasekhar, W. Seah, 2006. An area localization scheme for underwater sensor networks. In: Proceedings OCEANS 2006-Asia Pacific, pp. 1–8. http://dx.doi.org/10.1109/OCEANSAP.2006.4393969.

T. Camp, J. Boleng, B. Williams, L. Wilcox, W. Navidi, 2002. Performance comparison of two location based routing protocols for ad hoc networks. In: Proceedings of the

IEEE INFOCOM, vol. 3, pp. 1678–1687. doi: 10.1109/INFCOM.2002.1019421.
N. Chatterjee, Y. Jaya Lakshmi, A. Potluri, A. Negi, 2007. Effect of adapter promiscuous mode operation on dsr performance in manets. In: Proceedings of the International Conference on Signal Processing, Communications and Networking, ICSCN, pp. 45–48. doi: 10.1109/ICSCN.2007.350693.
David, J.B., Johnson, B., Maltz, David A., 2001. DSR: The Dynamic Source Routing Protocol for Multi-hop Wireless Ad hoc Networks. Addison-Wesley, 139–172.
Domingo, M.C., 2011. A distributed energy-aware routing protocol for underwater wireless sensor networks. Wirel Pers. Commun. 57 (4), 607–627. http://dx.doi.org/10.1007/s11277-009-9864-3.
M. Erol, S. Oktug, 2008. A localization and routing framework for mobile underwater sensor networks. In: Proceedings of the IEEE INFOCOM Workshops, pp. 1–3. 10.1109/INFOCOM.2008.4544584.
H. Elwany, R. Mahr, 2003. Deep water directional wave measurements from pressure, wave velocities and a three-axis accelerometer. In: Proceedings of the IEEE/OES Seventh Working Conference on Current Measurement Technology, 2003, pp. 127. http://dx.doi.org/10.1109/CCM.2003.1194298.
Forouzan, B.A., 2003. Data Communications and Networking 3rd edition. McGraw-Hill, Inc., New York, NY, USA.
J. Heidemann, W. Ye, J. Wills, A. Syed, Y. Li, 2006. Research challenges and applications for underwater sensor networking. In: Proceedings of the IEEE Wireless Communications and Networking Conference, WCNC, pp. 228–235. http://dx.doi.org/10.1109/%20MCNC.2006.1683469.
Harris, A.F., III, Stojanovic, M., Zorzi, M., 2009. Idle-time energy savings through wake-up modes in underwater acoustic networks. Ad Hoc Netw. 7 (4), 770–777. http://dx.doi.org/10.1016/j.adhoc.2008.07.014.
Han, G., Jiang, J., Shu, L., Xu, Y., Wang, F., 2012. Localization algorithms of underwater wireless sensor networks:a survey. Sensors 12 (2), 2026–2061. http://dx.doi.org/10.3390/s120202026.
A.F. Harris III, M. Zorzi, 2007. Modeling the underwater acoustic channel in ns2. In: Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools, ValueTools, pp. 18:1–18:8.
J. Jaffe, C. Schurgers, 2006. Sensor networks of freely drifting autonomous underwater explorers. In: Proceedings of the 1st ACM International Workshop on Underwater Networks, WUWNet, pp. 93–96. doi:10.1145/1161039.1161058.
J. M. Jornet, M. Stojanovic, M. Zorzi, 2008. Focused beam routing protocol for underwater acoustic networks. In: Proceedings of the Third ACM International Workshop on Underwater Networks, WuWNeT, pp. 75–82. http://dx.doi.org/10.1145/1410107.1410121.
Lanbo, L., Shengli, Z., Jun-Hong, C., 2008. Prospects and problems of wireless communication for underwater sensor networks. Wirel. Commun. Mob. Comput. 8 (8), 977–994. http://dx.doi.org/10.1002/wcm.v8:8.
U. Lee, P. Wang, Y. Noh, F. Vieira, M. Gerla, J.-H. Cui, 2010. Pressure routing for underwater sensor networks. In: Proceedings of the IEEE INFOCOM, pp. 1–9. doi: 10.1109/INFCOM.2010.5461986.
Linkquest. URL ⟨http://www.link-quest.com/⟩.
Munson, B.R., Young, D.F., Okiishi, T.H., Huebsch, W.W., 2009. Fundamentals of Fluid Mechanics. John Wiley & Sons.
Measuring deep ocean currents: a new velocity data set (2006). URL ⟨http://iprc.soest.hawaii.edu/newsletters/newsletter_sections/iprc_climate_vol6_1/argo_deep_ocean_currents.pdf⟩.
Mohamed, N., Jawar, I., Al-Jaroodi, J., Zhang, L., 2011. Sensor network architectures for monitoring underwater pipelines. Sensors 11 (11), 10738–10764.
N. Nowsheen, G. Karmakar, J. Kamruzzaman, 2013. An opportunistic message forwarding protocol for underwater acoustic sensor networks. In: Proceedings of the 19th Asia-Pacific Conference on Communications (APCC), pp. 172–177. http://dx.doi.org/10.1109/APCC.2013.6765937.
N. Nowsheen, C. Benson, M. Frater, 2010a. A high data-rate, software-defined underwater acoustic modem. In: Proceedings of the OCEANS, pp. 1–5. http://dx.doi.org/10.1109/OCEANS.2010.5664474.
N. Nowsheen, C. Benson, M. Frater, 2010b. Design of a high frequency FPGA acoustic modem for underwater communication. In: Proceedings of the IEEE OCEANS, pp. 1–6. http://dx.doi.org/10.1109/OCEANSSYD.2010.5603819.
J. Partan, J. Kurose, B.N. Levine, 2006. A survey of practical issues in underwater networks. In: Proceedings of the 1st ACM International Workshop on Underwater Networks, WUWNet, pp. 17–24. http://dx.doi.org/10.1145/1161039.1161045.
Pompili, D., Akyildiz, I.F., 2009. Overview of networking protocols for underwater wireless communications. IEEE Commun. Mag. 47 (1), 97–102. http://dx.doi.org/10.1109/MCOM.2009.4752684.
C.E. Perkins, E.M. Royer, 1999. Ad-hoc on-demand distance vector routing. In: Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications, WMCSA'99, IEEE Computer Society, pp. 90–100.
Perkins, C.E., Bhagwat, P., 1994. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. SIGCOMM Comput. Commun. Rev. 24 (4), 234–244. http://dx.doi.org/10.1145/190809.190336.
Prasan, U.D., Mahapatra, G., 2012. Energy efficient multiple sink variation to the depth based route protocol for underwater sensor network. Int. J. Eng. Sci. Adv. Technol. 2 (4), 951–954.
I. Rish, 2001. An empirical study of the naive bayes classifier. In: Proceedings of the IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3 (22), pp. 41–46.
Stojanovic, M., 2001. Underwater Acoustic Communication. John Wiley & Sons, Inc.. http://dx.doi.org/10.1002/047134608X.W5411.
Shin, D., Hwang, D., Kim, D., 2012. DFR: an efficient directional flooding-based routing protocol in underwater sensor networks. Wirel. Commun. Mob. Comput. 12 (17), 1517–1527. http://dx.doi.org/10.1002/wcm.1079.
W.-G. Seah, H.-X. Tan, 2006. Multipath virtual sink architecture for underwater sensor networks. In: Proceedings of the IEEE OCEANS, pp. 1–6. doi: 10.1109/OCEANSAP.2006.4393933.
Stojanovic, M., Preisig, J., 2009. Underwater acoustic communication channels: propagation models and statistical characterization. IEEE Commun. Mag. 47 (1), 84–89. http://dx.doi.org/10.1109/MCOM.2009.4752682.
Urick, R., 1983. Principles of Underwater Sound for Engineers. McGraw-Hill.
Wahid, A., Kim, D., 2012. An energy efficient localization-free routing protocol for underwater wireless sensor networks. Int. J. Distrib. Sens. Netw. 307246, 1–12. http://dx.doi.org/10.1155/2012/307246.
Wu, W., Cao, J., Fan, X., 2013. Design and performance evaluation of overhearing-aided data caching in wireless ad hoc networks. IEEE Trans. Parallel Distrib. Syst. 24 (3), 450–463. http://dx.doi.org/10.1109/TPDS.2012.147.
P. Xie, J.-H. Cui, L. Lao, VBF: vector-based forwarding protocol for underwater sensor networks. In: Proceedings of the 5th International IFIP-TC6 Conference on Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems, NETWORKING'06, pp. 1216–1221. http://dx.doi.org/10.1007/11753810_111.
H. Yan, Z. J. Shi, J.-H. Cui, 2008. DBR: depth-based routing for underwater sensor networks, in: Proceedings 7th International IFIP-TC6 Networking Conference on AdHoc and Sensor Networks, Wireless Networks, Next Generation Internet, NETWORKING'08, pp. 72–86.
Z. Zhou, J.-H. Cui, 2008. Energy efficient multi-path communication for time-critical applications in underwater sensor networks. In: Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc, pp. 221–230. http://dx.doi.org/10.1145/1374618.1374649.
Zhou, Z., Peng, Z., Cui, J.-H., Shi, Z., Bagtzoglou, A., 2011. Scalable localization with mobility prediction for underwater sensor networks. IEEE Trans. Mob. Comput. 10 (3), 335–348. http://dx.doi.org/10.1109/TMC.2010.158.
Zhou, Z., Cui, J.-H., Zhou, S., 2010. Efficient localization for large-scale underwater sensor networks. Ad Hoc Netw. 8 (3), 267–279. http://dx.doi.org/10.1016/j.adhoc.2009.08.005.
Z. Zhou, J.-H. Cui, S. Zhou, 2007. Localization for large-scale underwater sensor networks. In: Proceedings of the 6th International IFIP-TC6 Conference on Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet, NETWORKING'07, pp. 108–119.