

QL-EEBDG: QLearning based energy balanced routing in underwater sensor networks

Obaida Abdul Karim¹, Nadeem Javaid^{2*}, Arshad Sher², Zahid Wadud^{3,4}, and Sheeraz Ahmed^{5,6}

1 International Islamic University, Islamabad 44000, Pakistan

2 COMSATS Institute of Information Technology, Islamabad 44000, Pakistan

3 University of Engineering & Technology, Peshawar 25000, Pakistan

4 Capital University of Science and Technology, Islamabad 44000, Pakistan

5 Faculty of Engineering and Technology, Dera Ismail Khan 29050, Pakistan

6 Career Dynamics Research Center, Peshawar 25000, Pakistan * Correspondence: nadeemjavaidqau@gmail.com,
www.njavid.com

Abstract

In this paper, we propose a Q-Learning based efficient and balanced energy consumption data gathering routing protocol (QLEEBDG) for underwater sensor networks (USNs). We set an optimal next hop forwarder for each node to transmit its sensed data. This helps to reduce distance between sender and receiver. The energy consumption is minimum. Furthermore, a node is considered an eligible forwarder node only if its next hop neighbour exists. We incorporate this mechanism to avoid void hole problem. Our technique minimizes energy consumption in the network, hence, lifespan increases. The performance of our proposed technique is validated through extensive simulations.

Keywords: Energy tax, network stability period, throughput

Received on 12 November 2017, accepted on 10 March 2018, published on 10 April 2018

Copyright © 2018 Obaida Abdul Karim *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.10-4-2018.154459

1. Introduction

Underwater sensor networks (USNs) are getting popular these days due to their variety of applications like oceanographic data collection, maritime rescue, scientific ocean sampling, pollution and environmental monitoring. However, the deployment of USNs faces certain challenges like high bit error rate, limited bandwidth availability and large propagation delay. We prefer to use acoustic communication in USNs as the radio signals get absorbed in aquatic environment. However, acoustic signals speed in aqueous environment is 1500 m/s which is five orders of magnitude less than the radio signals propagation speed. One of the major challenges for sensor networks is limited energy supply of sensor nodes. The sensor nodes transmit their sensed data towards some base station [1] - [2]. Therefore, energy efficient and void hole

avoidance routing mechanisms are needed to prolong the lifespan of USNs. Energy efficient routing has been well investigated in terrestrial sensor networks, however, USNs have unique features. Therefore, novel routing mechanisms are needed for USNs. In void hole problem, a node is selected as a forwarded node to send the packet to the base station. The forwarded node neither has a neighbour node nor does it lie in the range of sink. If this problem remains unsolved, then packets are continuously dropped in the network and a large amount of energy [3] is wasted. Basically, the void hole is created due to energy utilization, mobility and random deployment of nodes in the network. Therefore, these challenges of creating the avoidance techniques [4]. A hybrid technique is used for transmission of data packet in balancing energy consumption to maximize network lifetime in data gathering sensor networks (EBDG) [5]. This technique achieves balanced energy consumption for all nodes in the network. However, this protocol is limited to small scale network (in terms of network radius) due to the hybrid

technique. This problem is tackled by the enhanced and efficient balance energy consumption in data gathering protocol (EEBDG). According to EEBDG, a hybrid transmission is restricted up to the optimum transmission range (R_{opt}). After that particular range, the nodes in the protocol follows the multihop transmission. In this way, all the nodes in the network consumes balanced energy and network is extended to the large-scale network. The USNs routing protocols need to be adaptive, robust and energy efficient, which demands a priori information about the network and restrictions on network architecture [6]. These requirements demands the development of machine learning based routing protocols for USNs. Therefore, we introduce the QL based routing protocol which is named as QL based efficient and balanced energy consumption data gathering routing protocol (QL-EEBDG) for USNs. The performance of our proposed protocol is satisfied in terms of dynamic nature of underwater. We perform the optimal procedure for the decision of one hop neighbour node. This decision is based on the reward parameter. It is the important parameter that helps to decide the optimal selection policy for the immediate successor node. In many other protocols, this parameter is achieved on different condition like, initial energy, residual energy, energy among the neighbour, distance and waiting time of a mobile sink(MS), node density, etc. [7], [9] and [8]. Thus, in our proposed protocol the reward is set to the minimum distance to the static sink. To avoid energy hole, those nodes are selected as neighbour node which have further neighbour node exist in range, as compared to the neighbour node which have no more neighbour in the range. To enhance the performance parameters of the EBDG, EEBDG and QL-EEBDG a MS is used in the network. When MS is introduced in the aforementioned protocols the resultant protocols are named as EBDG-MS, EEBDG-MS and QL-EEBDG-MS. The movement of MS in these protocols is clockwise. The decision parameter for MS is the minimum transmission range. If the MS is in node's range then nodes send the data packet to MS, otherwise it follows QL-EEBDG procedure. The objective of our proposed protocol is to maximize the network lifetime, network stability period and throughput. Moreover, it minimizes the energy utilization throughout the network and also overcomes the variation of energy consumption in the network.

2 Related work and motivation

In this section, we provide state of the art routing protocols for energy efficiency and void hole avoidance in USNs. An adaptive and efficient energy using QL based delay tolerant network routing protocol (QDTR) for USNs[8]. Due to the water current sensor node position varies, which results in void hole problem. QDTR employs a QL learning scheme to perform online learning

and handle node mobility using contact history of successor node. Hu. et al. proposed a reinforcement learning based routing protocol (QELAR). This protocol employs a fitness factor for selection of neighbour node. Thus QELAR achieves balanced energy consumption in the network [9]. Forster et al. [10] proposed feedback routing for optimizing multiple sinks in wireless sensor networks with reinforcement (FROMS). This protocol avoids the overhead of neighbour nodes with the help of multiple MSs. Also provides the recovery mechanism for node failure due to node mobility. As a result FROMS achieves low network cost as compared to its counterpart protocols. A well-known protocol, weighting depth and forwarding area division depth based routing for USNs is used to avoid the void hole problem. In this protocol, if during transmission a node does not find any forwarded node to send data packet, then the selected data packet is sent to available candidate node to further forward it to base station [11]. Moreover, in underwater environment nodes change their position due to the water current. Although, mobicast routing protocol (MR), handled this limitation [12]. In MR, a MS is used that moves on the predefined routes. This MS collects the data from all the sensor nodes and in this way it covers the whole network. Authors also avoid the direct transmission over the long transmission range[13]. Youngta et al. [14] propose a hydraulic based routing protocol (HRP). HRP addresses the low bandwidth, high energy consumption and mobility of the nodes in USNs. Like DBR, HRP has adopted the energy hole prevention mechanism. Thus, its performance improves in terms of maximum packet delivery, minimum delay in the network. It also helps to minimize the energy utilization.

3 QL-EEBDG

In this section, we provide a quick overview about the QL technique. Then, we discussed that how our routing protocol adopts this technique to make optimal decision in USNs

3.1 An overview of QL technique

QL is a reward based learning technique, in which agents make decisions based upon their optimal cumulative reward in order to reach the destination. Features of QL are finite set of actions, finite set of states, state action transition probability, expected reward, discount parameter γ (its limit lies between 0 and 1. Its main purpose is to discount the future rewards [9]). The evaluation function in QL is the $Q(\cdot)$ function. It works in pair form of state and action. This function is defined as, when an agent takes an action on the current state it received the future information from the next state. Stored this information in the form of a single number on the current state to the corresponding action. Therefore, Q table is maintained for all the states with associated action

and next state. Overall, QL chooses one terminal state as a target and then starts effort for achieving it. Thus, all states and actions chooses and work iteratively at least to reach the target state. The main equation in Q function is

$$Q(s_t, a_t) = R(s_t, a_t) + \gamma[\text{Max}_{a'}(Q(s', a'))]$$

This equation depends on two inputs from the next state. One is the immediate reward $R(s_t, a_t)$, and the second is the maximum value of Q function of the next state of any action. This input values generate the Q value for a given state s_t . Moreover, this function works recursively, because the reward and Q value of the next state is used as input for current state. Then, updates the current value according to the input value [15]. Next, we discussed that, how this QL technique works in our routing protocol.

3.2 The routing protocol

In USNs, routing protocols face difficulty in achieving the key objectives due to crucial nature of underwater. From recent routing protocols i.e. [7], [8], [9], [10], [17] and [18] it is known that QL is an energy efficient artificial intelligent algorithm. It enables an agent to performs efficiently in harsh underwater environment to achieve the protocol objectives. Moreover, QL works in the network which does not know the full architecture of the network. Nodes in the network act as agents. Sender node acts as a source agent, the neighbour node acts as a receiver agent. Each node generates a control packet (CP) in the network and sends CP to all the nodes. CP is received by node(s) which is in its range. In return, the receiver agent(s) sends back an acknowledgement packet to it. By this exchange of packets, source agent known that receiver agent(s) exists in its transmission range and declared it as a neighbor node(s). When source agent confirm the connection with receiver agent, then network generate the Q value for source agent. This value computes with the help of reward from neighbor node and Q value from that neighbor node as explain in equation (1). The Q value of a neighbor is come from its own neighbor node(s) which is more nearer to the sink or at less distance to the sink. Notice that, this Q value of a node(s) is high when, node is near to the sink. Moreover, in QL the parameter which is highly prominent and sensitive is the reward. In our case, the reward parameter is shortest distance towards the sink. Thus, we use different kinds of reward as, when a node finds a sink as a next node, then node gets the sink reward (Reward-sink). On the other hand, if node find out the neighbor, then, receives the positive reward (Reward-pos) from neighbor. At last, when a node gains negative reward (Reward-neg) when finds nor neighbor neither sink.

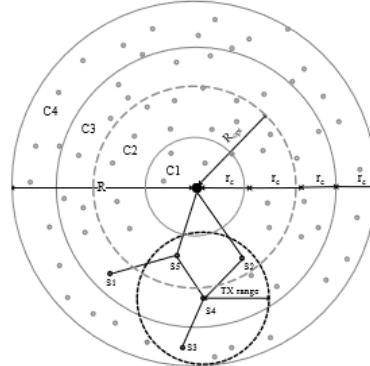


Fig. 1: Network topology

Similarly, the network maintained the Q table, in this table each node stores the Q value of its neighbor node(s). At the beginning, the table for all nodes are filled with zeros. Therefore, when node start knowing about neighbor through QL process. This Q table start updating. However, for each node the Q value of that node(s) in Q table is updated which is exist in its transmission range. We assume that our field is circular. This field having a radius R . This circular region is further divided into n concentric circles i.e. C_1, C_2, \dots, C_n with radius r_c . Moreover, a static sink (dark hexa pentagon) is deployed at the center of network field. 0.04 nodes per square meter are uniform randomly deployed in that particular field and it is indicated by the small circle shaped as shown in Fig. 1. All nodes follow the direct and multihop transmissions in the network. However, we restrict the mixed transmission at some extent in the network by the help of R_{opt} and this range is shown by the large dashed circle in figure. The nodes with in R_{opt} , performs the mixed transmission successfully. While, nodes out of this range perform only the multihop transmission, even if data distribution ratio π_i (explained later in transmission phase) of their particular concentric circle is satisfied for direct transmission. In addition, the transmission in our protocol occurs from outer concentric circles to inner concentric circles towards the sink. According to proposed QL algorithm, we initialize the Q_{matrix} with zeros while R_{matrix} fills with different rewards (as explained in the routing protocol section). The selected node checks all corresponding actions to its row in R_{matrix} . If the source node finds an action(s) that has reward instead of Reward-neg, then the action pointed towards the next node, then this neighbor node selected and store in the temporary matrix. After that, next node(s) is extracted from the temporary matrix. Then, with help of equation (1) generates the Q value for source node and the next node which is point as next node, for that particular source node. This Q value shows that how much the next node is reliable to be selected as a neighbor node for the source node. Source node calculates and stores Q value for the next node(s) in Q_{Matrix} . This training performs repeatedly up to the Q value for all nodes are converged. Due to converged values in the Q_{matrix} each node in the network easily find out the next node that leads to the destination

state. The QL algorithm helps in training of nodes to learn from the environment on the base of the Q value. Next we check the detail about the structure of a packet which we follow while implementing this QL technique.

3.3 Format of a packet

We inserted the new fields in the header of network layer due to which the size of packet header is increased from 20 to 28 bytes [19]. In addition, sender ID presents the ID of a sender node and the receiver ID indicates the ID of the receiver node. The Q value and reward provides by the receiver node and each it size is 16 bits. At last, the sender node's residual energy and the terminal state ID indicates the address of the static sink in the network. After that, we discussed the transmission phase in our routing protocol.

3.4 Data transmission

To start the transmission of data packet(s) in the network. First, we compute the p_i [5] and assigned to each concentric circle. Then, this ratio decides for concentric circle about the transmission (all nodes in this particular circle follow the decision). While R_{opt} is to find out to restrict the decision of direct transmission on the larger radii. We defined R_{opt} as,

$$R_{opt} = -\ln\left\{\frac{\ln(Pr) - \ln(Po)}{\ln(a)} \times \frac{1}{\ln(k-1) \times \ln(\ln(a)) + k}\right\} \quad (2)$$

The input power to the receiver is po and value of po is $1 \times 10^{-3} \text{J/bit}$ and the output power of the receiver is $P_u = 0.2 \times 10^{-3} \text{J/bit}$ which depends on receiving devices. The k is the spreading factor (1 for cylindrical, 1.5 for practical and 2 for spherical) and a is the absorption coefficient which depends of frequency [16].

$$a = 10^\alpha(f)/10 \quad (3)$$

The particular frequency f is calculated by the Thorps expression as,

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

If a node is within the R_{opt} then, directly communicates with sink, else forward the packet to the neighbor node. For example, The result of R_{opt} is 300 meter after calculation. Up to 300m all nodes in network perform mixed transmission, after that performs the multi hop transmission. No matter, if the p_i generate the direct ratio for concentric circle. For multihop transmission, our routing protocol taking the best route decision on the base of the maximum Q value. Sender node checks its own store information about the neighbor node(s) stored in the Q table. And then, sender node selects those neighbor node from Q table, which has maximum Q value for transmission. In a case, when two nodes have equivalent Q value, then, it takes the decision about the node which has maximum residual energy in the network.

3.5 QL-EEBDG-MS

We use a MS in our proposed scheme. A MS moves clockwise along the circular radius from lower concentric circle to the upper concentric circle to covers the whole network field. In this proposed protocol, when node(s) wants to send data packet(s) to the sink. Then, node checks the transmission range of the MS. If MS in less transmission range, then forward the data packet(s) to it, otherwise, follows the procedure of QL-EEBDG. By the addition of the MS, all the performance parameter is enhanced. While the results of our QL-EEBDGMS is more prominent in terms of MS.

4 Performance analysis

Here, we examine the behavior of QL-EEBDG and QL-EEBDG-MS through simulations. We compare both of the schemes with EBDG and EEBDG by considering static and MSs. We divide the network field into circular region, radius of field various from 100m to 1000m. Amount of energy assigned to each of the nodes is 1 Joule.

4.1 Simulation analysis of proposed protocol with static sink

Fig. 2 illustrates the energy tax of three different protocols at different network radii. On smaller radii less energy is used by all protocols as compared to larger radii because of transmission radius. Proposed protocol performs better than EBDG, because the existing protocol follows the mixed transmission procedure throughout the network. While, we follow the same procedure for transmission as in EEBDG, and then we use the QL technique for neighbour node. Hence, our protocol better than EBDG like EEBDG. On contrary, QL-EEBDG with the EEBDG, both have same performance at smaller radii. Because, the basic technique for transmission of a data packet(s) at this radii is same. While the difference occur in accessing the neighbor nodes. Because, to find optimal neighbor node we use the QL. Therefore, twice of QL-EEBDG energy is used, first to makes the decision and then, act on this decision. Due to which our energy consumption is more than existing protocol. We have simulated the network stability period of three protocols at different network radii, is shown in Fig. 3. The network stability period for all protocols is decreasing with the increase of network radius, because energy tax (Fig. 2) has indirect effect on network stability period. Our proposed protocol is better than existing protocols. Because we generates the optimal path for multihop transmission which is not find out in EBDG and neither in EEBDG. The proposed protocol stores optimal path information in the Q table. When node communicates in the network, then, simply accessed this Q table, extracts the information about neighbor node, forwards the data packet(s) to the particular preceding state. Due to our proposed protocol use less energy when

communicates with neighbour node(s). Therefore, node alive for more period of time,

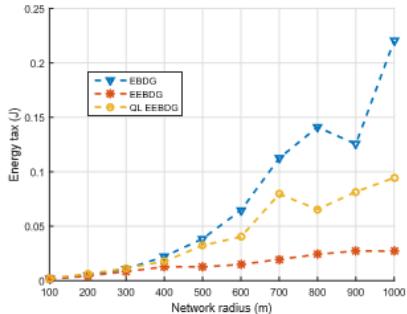


Fig. 2: Pattern of energy tax versus network radius

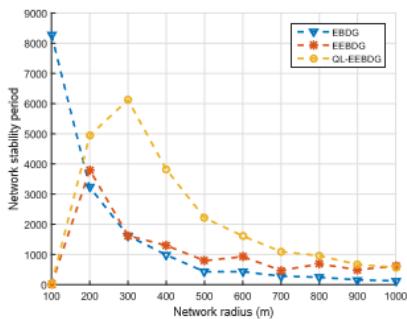


Fig. 3: Network stability period versus network radius

thus, our network stable for more time than EBDG and EEBDG protocols. Moreover, network stability period of EBDG is much smaller than existing and our proposed protocols. Because its energy consumption is more due to mixed transmission.

4.2 Simulation analysis of our proposed protocol a MS

Energy tax for EBDG-MS, EEBDG-MS and QL-EEBDG-MS under different network radii is illustrates in Fig. 4. At lower radii, all of the protocols have same performance in terms of energy tax. This transmission consumes less energy due to less transmission range. However, mixed transmission on larger radii rapidly depletes energy as in EBDG-MS protocol. While, the energy tax of QLEEBDG-MS and EEBDG-MS are lower at this radius because both avoid the hybrid transmission and perform the multihop transmission. Notice that, our proposed protocol consumes more energy than EEBDG-MS. Because of optimal decision for neighbour node(s) by the help of QL. While the EEBDG-MS has no such kind of decision. However, the energy spent due to MS in all protocols is more than that energy consumption due to static sink technique. To check the reliability of the proposed protocol compared it with the existing protocols and is shown in Fig. 5. EBDG-MS network stability

period is low as compared to EEBDG-MS and QL-EEBDG-MS because it energy is soon depletes due to the direct transmission on the larger radii. While, in EEBDG-MS, this period is significantly better as compared to EBDG-MS. Because it restricts the mixed transmission up to R_{opt} . As QL-EEBDG-MS, performing better than the EBDG-MS and EEBDG-MS, because we implements the QL algorithm for finding the one or more hop neighbor node(s). According to QL, those nodes are selected as a neighbor nodes which has further neighbor node exist or nearer to the base station. By this way, we avoid the energy hole problem. The performance of all protocols are enhanced with use of a MS in the network. The prominent change shows in our proposed protocol that is, at 300m, with MS stability reaches

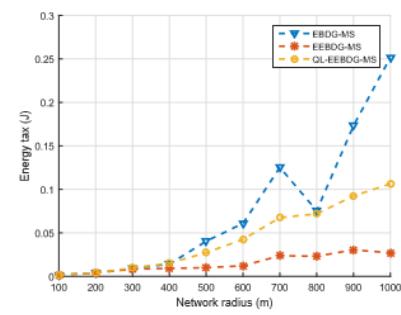
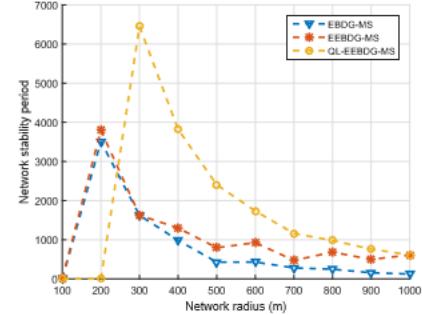


Fig. 4: Impact of network radius on energy tax with a MS



up to the 6500 rounds as compared to static sink, its reach to 6000 as shown in Fig. 4.

5 Conclusion

This paper introduces an energy efficient QL based routing protocol for USNs. This technique forms optimal routes towards sink. Also, our proposed technique defines mechanism to avoid void hole problem. The efficiency of our proposed technique has been studied using extensive simulations. Our technique achieves prolonged network lifetime and stability period.

References

1. Akyildiz, Ian F., Dario Pompili, and Tommaso Melodia. Underwater acoustic sensor networks: research challenges. *Ad hoc networks*, vol. 3, (2005): 257-279.
2. Mohammad Taghi Kheirabadi and Mohd Murtadha Mohamad, Greedy Routing in Underwater Acoustic Sensor Networks: A Survey. *International Journal of Distributed Sensor Networks*, 2013: 1-21.
3. Ghoreyshi, Seyed Mohammad, Alireza Shahrabi, and Tuleen Boutaleb. VoidHandling Techniques for Routing Protocols in Underwater Sensor Networks: Survey and Challenges. *IEEE Communications Surveys and Tutorials*, vol. 19, no. 2, 2017: 800-827.
4. J. Jiang, G. Han, H. Guo, L. Shu, and J. J. Rodrigues, Geographic multipath routing based on geospatial division in duty-cycled underwater wireless sensor networks. *Journal of Network and Computer Applications*, vol. 59, 2016: 413.
5. Zhang, Haibo, and Hong Shen. Balancing energy consumption to maximize network lifetime in data-gathering sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, 2009: 1526-1539.
6. Xu, Xin, Lei Zuo, and Zhenhua Huang. Reinforcement learning algorithms with function approximation: Recent advances and applications. *Information Sciences*, vol. 261, 2014: 1-31.
- 10 Obaida Abdul Karim et al.
7. Plate, Randall, and Cherry Wakayama. Utilizing kinematics and selective sweeping in reinforcement learning-based routing algorithms for underwater networks. *Ad Hoc Networks*, vol. 34, 2015: 105-120.
8. Hu, Tiansi, and Yunsi Fei. An adaptive and energy-efficient routing protocol based on machine learning for underwater delay tolerant networks. In *Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS)*, 2010 IEEE International Symposium on, 2010: 381-384.
9. Hu, Tiansi, and Yunsi Fei. QELAR: a machine-learning-based adaptive routing protocol for energy-efficient and lifetime-extended underwater sensor networks. *IEEE Transactions on Mobile Computing*, vol. 9, no. 6, 2010: 796-809.
10. Forster, Anna, and Amy L. Murphy. FROMS: Feedback routing for optimizing multiple sinks in WSN with reinforcement learning. In *Intelligent Sensors, Sensor Networks and Information (ISSNIP)*, 3rd International Conference on, 2007: 371376.
11. H. Yu, N. Yao, T. Wang, G. Li, Z. Gao, and G. Tan, Wdfad-dbr: Weighting depth and forwarding area division dbr routing protocol for uasns. *Ad Hoc Networks*, vol. 37, 2016: 256-282.
12. Chen, Yuh-Shyan, and Yun-Wei Lin. Mobicast routing protocol for underwater sensor networks. *IEEE Sensors journal*, vol. 13, no. 2, 2013: 737-749.
13. Javaid, Nadeem, Mehreen Shah, Ashfaq Ahmad, Muhammad Imran, Majid Iqbal Khan, and Athanasios V. Vasilakos. An Enhanced Energy Balanced Data Transmission Protocol for Underwater Acoustic Sensor Networks. *Sensors* 16, vol. 4, 2016: 487-508.
14. Noh, Youngtae, Uichin Lee, Saewoom Lee, Paul Wang, Luiz FM Vieira, Jun-Hong Cui, Mario Gerla, and Kiseon Kim. Hydrocast: pressure routing for underwater sensor networks. *IEEE Transactions on Vehicular Technology* 65, no. 1, 2016: 333347.
15. <https://web.stanford.edu/group/pdplab/pdphandbook/handbookch10.html>
16. Dou, Jinfeng, Guangxu Zhang, Zhongwen Guo, and Jiabao Cao. PAS: probability and suboptimal distance based lifetime prolonging strategy for underwater acoustic sensor networks. *Wireless Communications and Mobile Computing*, vol. 8, 2008: 1061-1073.
17. Hu, Tiansi, and Yunsi Fei. MURAO: A multi-level routing protocol for acoustical optical hybrid underwater wireless sensor networks. In *Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 9th Annual IEEE Communications Society Conference on, 2012: 218-226.
18. Zou, Lin. Machine Learning Based Delay Tolerant Protocols for Underwater Acoustic Wireless Sensor Networks. (2014).
19. DeCusatis, Casimer, ed. *Handbook of fiber optic data communication: a practical guide to optical networking*. Academic Press, 2013.