RESEARCH ARTICLE

Optimization of an Underwater Wireless Sensor Network Architecture with Wave Glider as a Mobile Gateway

Tatiana A. Fedorova[1], Vladimir A. Ryzhov[1], Nikolay N. Semenov[2] and Shaharin A. Sulaiman[3]

Abstract
This paper presents an original probabilistic model of a hybrid underwater wireless sensor network (UWSN), which includes a network of stationary sensors placed on the seabed and a mobile gateway. The mobile gateway is a wave glider that collects data from the underwater network segment and retransmits it to the processing center. The authors consider the joint problem of optimal localization of stationary network nodes and the corresponding model for bypassing reference nodes by a wave glider. The optimality of the network is evaluated according to the criteria of energy efficiency and reliability. The influence of various physical and technical parameters of the network on its energy efficiency and on the lifespan of sensor nodes is analyzed. The analysis is carried out for networks of various scales, depending on the localization of stationary nodes and the model of bypassing the network with a wave glider. As a model example, the simulation of the functional characteristics of the network for a given size of the water area is carried out. It is shown that in the case of a medium-sized water area, the model of "bypassing the perimeter" by a wave glider is practically feasible, energy efficient and reliable for hourly data measurements. In the case of a large water area, the cluster bypass model becomes more efficient.

Article Highlights

- A probabilistic model is presented for a hybrid underwater wireless sensor network (UWSN) with stationary sensors placed on the seabed and a wave glider as a mobile gateway;
- Analytical expressions for the criteria of energy efficiency and reliability of the network are evaluated;
- The functional characteristics of various scales networks are obtained depending on the placement of stationary nodes and the network bypass model using the wave glider.

✉ Shaharin A. Sulaiman
  shaharin@utp.edu.my

[1] Department of Applied Mathematics and Mathematical Modeling, Saint-Petersburg State Marine Technical University, 3 Lotsmanskaya, Saint Petersburg 190121, Russia

[2] Department of Marine Information Systems and Technologies, Saint-Petersburg State Marine Technical University, 3 Lotsmanskaya, Saint Petersburg 190121, Russia

[3] Department of Mechanical Engineering, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Malaysia

1 Introduction

The Internet of Underwater Things (IoUT) has great promise of becoming a strong foundation for "smart ocean" (Domingo, 2012; Toma et al., 2011) under the United Nations Decade of Ocean Science (2021-2030) for Sustainable Development announced by UNESCO (Ryabinin et al., 2019). The development of underwater networks, which are part of IoUT, is triggered by active development of ocean resources. This involves comprehensive scientific researches, including the ones related to natural phenomena forecasting and natural disaster prevention, climate change, processes of marine technical objects robotization, automation of marine transportation, navigation and safety assurance. The current IoUT solutions integrate marine technical objects, such as buoys, surface and underwater vehicles, ocean observing systems, and onshore data processing centers, into heterogeneous networks of various scales and purposes (Crout et al., 2006; Ho et al., 2018; Riser et al., 2016).

Long-term and non-continuous ocean observing cable

networks have been actively developed in the last two decades. Examples of key players are Australian Integrated Marine Observing System, Ocean Works Canada, Dense Ocean Floor Network System for Earthquakes and Tsunamis (Barnes and Team, 2007; Kaneda, 2009; Oke and Sakov, 2012; Wallace et al., 2014). However, due to the fact that the deployment of cable networks in marine conditions is very expensive, the concept of wireless networks has been actively developed along with the introduction of the cable networks. An example of the practical implementation of such networks is the Persistent Littoral Undersea Surveillance Network-PLUSNet (Grund et al., 2006), which uses a few underwater vehicles to communicate using semi-autonomous sensors in the absence of human direction.

The scale of practical use and the dynamics of research work on the creation of underwater systems, primarily wireless networks, indicates their relevance for a wide range of applications within the IoUT concept. It is worth mentioning that it was the development of wireless technology that initiated the transition of the underwater applications to a qualitatively new level. The research of radio-frequency, optical, magneto-inductive and acoustic technologies of wireless data transmission under water has showed the advantages and disadvantages of each of them. Radio-frequency and magneto-inductive communication can provide high data transmission rate, but only over short distances due to signal absorption and scattering by water. Acoustic communication can provide relatively longer transmission ranges, but its use is highly dependent on properties of the marine environment.

Despite the positive and negative qualities of each of the specified technologies of underwater communication, it can be argued that hydroacoustic technologies for underwater wireless sensor networks (UWSNs) are the most applicable and successful in marine practice at present (Rossi et al., 2014; Felemban et al., 2015; Yan et al., 2019; Fattah et al., 2020). When designing hydroacoustic UWSNs of different scale for specific tasks, the problem of creating the architecture and network components that are able to provide high reliability of delivery of the collected information to the end user and long-term autonomous network operation in complex conditions of the marine environment, comes to the fore.

A number of the underwater environment features, such as rapid variability of its characteristics, long delays of the hydroacoustic signal propagation, significantly limited frequency band that can be used for signal transmission, great intensity and duration reverberation, extended shadow zones, signal freezing due to its multipath propagation, large Doppler distortions determine the conditions' complexity (Brekhovskikh and Lysanov, 2003; Urik, 1967). Taking into account these features in designing hydroacoustic UWSNs would be a complex scientific and technical problem. From the research and theoretical points of view, this problem combines a great number of interrelated subproblems, which can be solved within various mathematical models.

There are a large number of works devoted to UWSN modeling, network energy balancing, security etc. These works were devoted to various aspects of the functioning of UWSN: development of node deployment algorithms (Han et al., 2013; Choudhary and Goyal, 2021), routing protocols (Khan et al., 2018; Mahmood et al., 2019; Menon et al., 2022; Samy et al., 2022), localization techniques (Su X. et al., 2020; Nain et al., 2022), mobility network sensor nodes (Choudhary and Goyal, 2022), network energy balancing (Xing et al., 2019; Rizvi et al., 2022), etc. The problems were solved in 2D and 3D cases for stationary, mobile and hybrid UWSNs. Since the time they started to be actively developed, marine robotics platforms, which use the principles of underwater gliders and wave gliders (WG), have been considered as promising platforms for mobile UWSN agents (Lan et al., 2021; Li et al., 2019; Su Y et al., 2020).

Following the recent trend, this the paper proposes to use wave glider (WG) (Nikushchenko et al., 2019; Ovchinnikov et al., 2020) as a UWSN mobile agent, which is a two-module structure consisting of a surface and an underwater module connected by a cable-rope. The surface module, which resembles the body of a ship, is equipped with solar-to-electric energy conversion panels to power up the WG systems and its propulsion. The underwater module is propelled by converting wave energy into thrust by a system of oscillating wings located on the underwater module. Due to the use of renewable energy from the sea, the WG has a high autonomy and its position and course can be controlled (Wang et al., 2018). The underwater module is equipped with a hydroacoustic modem, which provides communication with the sensor network. On the other hand, the surface module has radio-frequency and satellite communication facilities, which support communication with the data processing centers. These features make the WG a convenient platform for mobile UWSN agents. In addition, the WG is relatively low cost, which is also important in the UWSN creation.

The paper proposes a simple universal mathematical model for the functioning of a hydroacoustic hybrid UWSN and the results of the modeling in terms of energy efficiency of the network, depending on its configuration. The UWSN functional characteristics are assessed using a probabilistic approach.

Taking into account the aforementioned technical peculiarities of the WG, the models of its use as a mobile gateway of hybrid UWSN are considered. Three models of the WG for servicing the underwater network segment are analyzed: the clustering model, the water area perimeter bypass model and the coastline bypass model. The problem of selecting the WG usage model is considered together with the problem of optimal localization of ordinary and

reference sensor nodes (agents) depending on the design parameters of the developed network.

The proposed probabilistic model is used to analyze the influence of design parameters such as network dimensions, message delivery probability, distance between the network nodes, absorption and scattering by the water and time interval of data acquisition by sensors on the network total power consumption and nodes' lifespan. The simulation results allow determining the optimal parameters of the UWSN in terms of energy efficiency and reliability of the entire network. The WG applicability as a mobile gateway is assessed based on the functional characteristics correlation of the considered stationary nodes' localizations with the WG capabilities, which is limited by its speed in the water area.

This article is organized as follows. Section 2 proposes a communication architecture for an underwater wireless sensor network (UWSN), which includes a group of stationary agents, ordinary and reference ones and a mobile gateway-a wave glider moving over the surface of a water area. Three different variants of reference node localizations are considered with the aim to improve network bandwidth, and consequently increase network reliability and efficiency-partitioning the net into a finite number of clusters, reference nodes' placement on the perimeter of the water area and reference nodes' placement along the coastline.

In Section 3, the physical formulation of the problem is given, determined by the parameters of the environment, the technical characteristics of the simulated network and the routing protocols. A mathematical formulation of the problem is presented in Section 4 with base calculation formulas. In Section 5, the optimization of the UWSN for water areas of an arbitrary scale is carried out. Section 6 presents the results of calculations that make it possible to analyze the dependence of the efficiency and reliability of the network on the distance between nodes for a given network scale, to determine the optimal localization in a given water area, the optimal distance between nodes and the probability of message delivery. At the end of this article, the conclusions are formulated and possible directions for further work are outlined.

2 Communication architecture of underwater wireless sensor network

This paper considers the communication architecture of the underwater wireless sensor network. As shown in Figure 1, the network includes a group of stationary nodes, both ordinary and reference ones, placed at the sea bottom, and a mobile gateway - wave glider, moving on the surface of a certain water area. The messages between the network nodes are transmitted via a hydroacoustic channel.

Two variants of sensor placement can be considered, which

(a) Experimental model of the wave glider
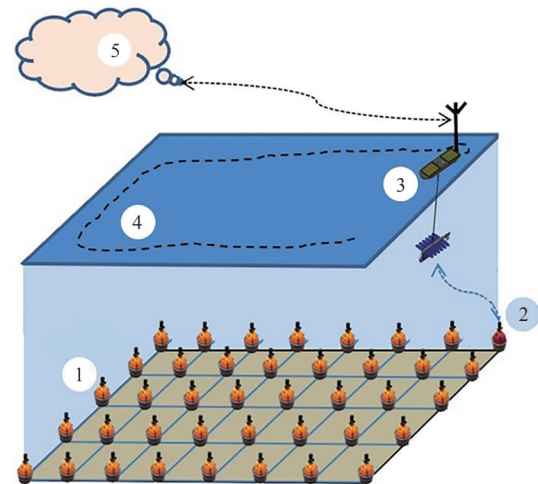
1-ordinary sensor nodes, 2-reference sensor node, 3-mobile gateway - wave glider, 4-WG trajectory, 5-data processing center

(b) UWSN architecture with one mobile gateway (wave glider)

Figure 1　Model of the wave glider and UWSN architecture

are determined by the topography of the seabed.

In the first case, the stationary orthogonal grid of sensors is placed on a flat bottom (the presence of various obstacles is not taken into account). This corresponds to the topology of the bottom surface at depths of 1 000 m, which are considered further in the paper. The probability of message delivery is determined only by the distance between the sensors and physical characteristics of the environment, such as absorption and scattering by the water. This probability is the same in all directions (isotropic environment).

In the second, which is a more general case, the same grid is located on the bottom with a complex 3D topology. There may be various obstacles at the bottom - reefs, sunken objects, accumulations of algae, which make it difficult to transfer the message. The type of obstacle determines the probability of message delivery from one sensor to another. Thus, in different zones of the water area, the probability of message delivery will be different (anisotropic environment). The second case is more complex and less representative for analyzing the efficiency of the routing pro-

tocols of data forwarding to the gateway with the required reliability.

Therefore, this article considers the first variant of the sensor placement, when the probability of message delivery depends only on the distance and physical characteristics of the medium.

It is assumed that there are one or more dedicated reference nodes to which data from the ordinary nodes in the network is forwarded. Thus, a significantly larger traffic volume passes through the ordinary nodes near the reference nodes and through the reference nodes themselves. Apparently, the more data passes through the underwater wireless network node, the greater is its power consumption. Consequently, there is a problem of imbalanced power consumption in the network. This problem leads to reference nodes failing earlier than others due to battery drain, resulting in a shortened lifespan of the sensor network as a whole. Various energy balancing methods are used to equalize the energy consumption of all nodes in the network:

· Individual selection of the battery capacity depending on the position of nodes in the network and the functions they perform. In this case, the reference nodes can be equipped with high-capacity batteries.

· Different density of network nodes in the water area, depending on the expected traffic intensity. This solution is aimed at ensuring redundancy in the network structure and duplication of the individual nodes' functions. If one network node fails, its functions will be shifted to an available redundant neighboring node.

· Using routing protocols based on such principles as alternating long and short distance transmission and assessing the residual energy value of the nodes on the way to the gateway. In this case, the route in which the nodes have more residual energy is selected from the set of alternative routes.

· Using clustering as division of all network nodes of the UWSN into groups (clusters) with internal routing and internal reference nodes.

· Using the individual network components mobility (mobile gateways). Autonomous underwater vehicles (AUV), underwater gliders (UG) and wave gliders (WG) can be used as mobile gateways.

In this paper a wave glider is used for energy balancing of the network. Let the UWSN be located in a physically homogeneous medium and have the form of a square orthogonal grid with dimension $n \times n$, where n is the size of the grid in terms of the number of nodes. The sensors are located at their nodes. Localization of sensor nodes on an orthogonal grid is convenient because its geometric characteristics are easy to compute. If the network has a single reference node, which is located in the lower right corner of the grid, then summarizing the total number of transmissions of all messages from ordinary nodes to the reference node through the nearest neighbors, we get the following

expression:

$$L(n) = n^2(n - 1) \qquad (1)$$

This is apparently the minimum number of transmissions needed. It will increase if it is necessary to resend messages in the event of losses. It is assumed that an ordinary node is at the k-th range level from the reference node, if k hops are required to deliver the message. Thus, it is convenient to characterize an ordinary node by its range level from the reference node.

Taking into account the message delivery probability, p, and the possibility of sending the message for an unlimited number of times, the mathematical expectation of the number of transmission attempts between neighboring nodes, required for successful delivery is $1/p$. The total expected number of transmissions in the network is then defined by the following expression:

$$L_p(n) = \frac{L}{p} = \frac{n^2(n - 1)}{p} \qquad (2)$$

Since it is obvious that the bandwidth of a single reference node is limited, and a large value of n will slow down message reception, leading to a bottleneck. In this paper, three different variants of reference node localizations are considered with the aim to improve network bandwidth, and consequently increase network reliability and efficiency:

1) Partitioning the net into a finite number of clusters: in each cluster, ordinary nodes send messages to the reference node, which in turn sends the information to the mobile gateway. In this case, the mobile gateway - the wave glider (WG) - needs to bypass the entire water area, sending notification to the reference nodes to establish a connection and receive messages from them (Figure 2a).

2) Reference nodes' location on the perimeter of the water area. The WG should bypass only the perimeter of the water area (Figure 2b).

3) Reference nodes' location along the coastline (one of the sides of the net). The WG should bypass only the coastline (Figure 2c).

3 Physical problem descriptions

This paper deals with the problem in the following model formulation descriptions. Let it be required to measure some characteristics and transmit data (information messages or network packets) from each ordinary node through the other nodes to the reference ones, which are in their turn pass packets to the gateway (Figures 1 and 2). The probability of message delivery between two nearest nodes depends only on the distance between them (r) and is equal to $p=p(r)$. Transmitted packet amount generated by each node is the

(a) Cluster model

(b) "Perimeter bypass" model

(c) "Coastline bypass" model
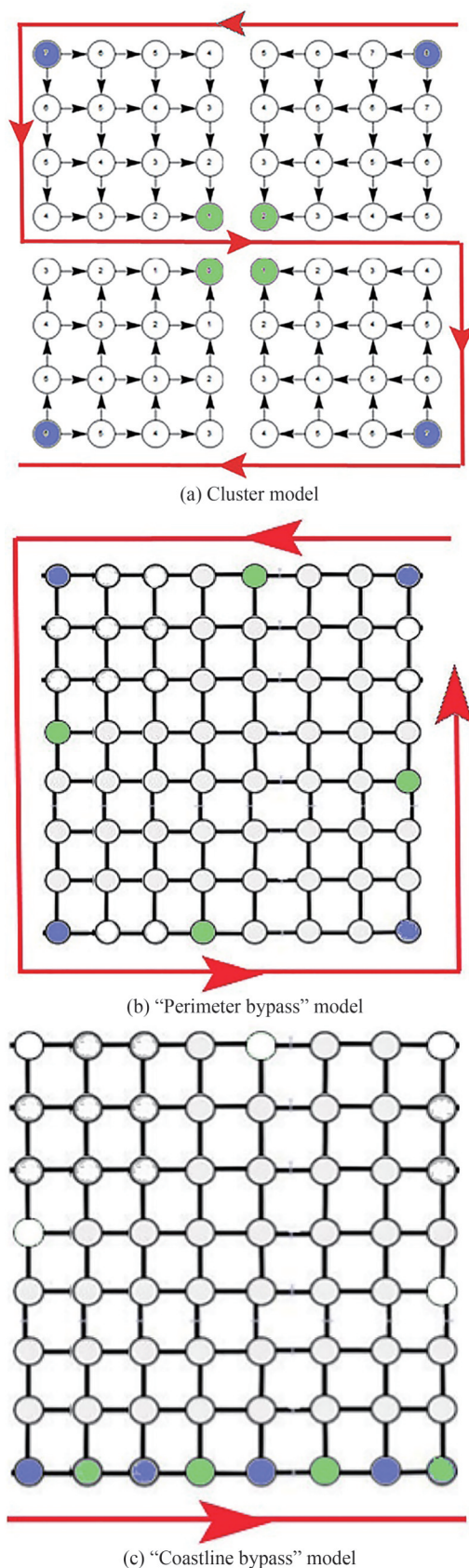
Figure 2 The model of the FSGS of 64 nodes with a mobile gateway (wave glider). Red arrow line - trajectory of WG movement

same. In addition to information on the measured characteristics, the packet contains the necessary service information.

The following physical parameters are used to determine the network characteristics. The total running time of the network between two measurements of the characteristics (cycle) is assumed to be limited and is represented by T. All ordinary nodes must pass the results of measurements (their own and all others') through the reference nodes to the gateway and proceed to the next measurement during this time. In this paper three time-ranges of measurements are considered - once every few minutes, once every hour, and twice a day. The choice of time range is determined according to the network application purposes.

The routing protocols used are neighbor-to-neighbor routing protocols, which means that the ordinary nodes have to pass messages only to their nearest neighbors. If after a fixed number of transmissions to all neighbors the message is not transmitted, it is considered as lost and the node begins transmitting the next message in the queue. The routing table determines which neighbor has a higher, lower or zero priority for transmitting messages for each node. The priorities are determined during the agents' learning process, based on the shortest distance to the reference node and on the number of successful transmissions along the given direction on the previous cycles. Thus, the routing table can continuously be rebuilt during learning. The agent learning algorithms is not discussed since it is not the primary interest of this paper.

The medium access control protocol is non-synchronized hydroacoustic channel access based on competition. The well-known classical ALOHA protocol is an example of such a protocol (Ahn et al. 2011). For energy conservation reasons, the number of reference nodes m is set to 5% of the total number of nodes in the network (Lindsey and Raghavendra 2012):

$$m = 0.05n^2 \qquad (3)$$

In addition, the total expected number of transmissions in the network is determined by Equation (2), in which the probability of message delivery between the two nearest nodes is $p=p(r)$.

If $L_p(k)$ is the total expected number of transmissions through one ordinary node at the k-th range level from the reference node, then $T_k = T/L_p(k)$ is the average running time of transmitting of one message and $\lambda_k = L_p(k)/T$ is the flow of messages through the node during one cycle. The running time of the ordinary node is divided into the following intervals: T_{inf} is the time duration for data collecting, T_s is the time duration for packet transmission and receipt of delivery acknowledgement. In this work it is assumed that all messages have the same length, and the time of their single transmission is $T_s = 0.2$ s. The average waiting time between any two consecutive messages is defined

as represented by T_w. The energy characteristics of the sensor are:

· P_s = 25 W is the power when sending a message,

· P_w = 0.3 W is the power when waiting for and receiving a message, and

· E_0 = 864 kJ is the capacity of the 12V battery.

4 Mathematical model of the problem

4.1 Lifespan estimates for ordinary and reference nodes

Knowing the initial battery energy E_0 and the power P_k consumed by the node at the k-th range level, the network node lifespan t_k can be estimated as follows:

$$t_k = \frac{E_0}{P_k} \qquad (4)$$

The node spends energy E_k^c for data collection, energy E_k^s for $L_p(k)$ message transmission, and energy E_k^w for waiting during one cycle, so the total energy spent per cycle E_k is defined as follows:

$$E_k = E_k^c + E_k^s + E_k^w \qquad (5)$$

Energy E_k^s, which is spent on sending the message, is far higher than the energy E_k^c needed to collect data, so neglecting the energy E_k^c and considering that $E_k = P_k T$, the expression for the total energy consumed by the sensor per work cycle T will look as follows:

$$P_k T = L_p(k) P_s T_s + P_w T_w \qquad (6)$$

The waiting time T_w can be calculated as the time not taken up by message transmission:

$$T_w = T - L_p(k) T_s \qquad (7)$$

It is not difficult to obtain the sensor lifespan at the k-th range level from Equations (4) and (6):

$$t_p(k) = \frac{E_0 T}{L_p(k) P_s T_s + P_w \big(T - L_p(k) T_s \big)} \qquad (8)$$

Substituting the known values of the parameters into Equation (8) and converting them into hours, the following result is obtained:

$$t_p(k) = \frac{240}{4.94 L_p(k)/T + 0.3} \qquad (9)$$

Equation (9) shows that at $\frac{4.94 L_p(k)}{T} \ll 0.3$, the lifespan of the node at this range level ceases to depend on the number of packets transmitted, and tends to 800 hours. Let us call such a mode, (at which this condition is met) the low-loading mode, and the mode at which $4.94 L_p(k)/T \approx 0.3$ is met, the high-loading mode. If the reference node operates in the low-loading mode, its lifespan will depend very little on localization, and optimization becomes problematic according to this criterion. For this reason, only high loading of the network is considered in the first theoretical part of the paper.

Thus, knowing the node workload and the cycle duration, it is possible to determine their lifespans, and estimate the lifespan of the whole network (by analogy with reasoning (Lindsey and Raghavendra 2012), that the system remains viable, provided that at least 25% of the network nodes are functional).

4.2 Energy costs estimate

The power is assumed to be negligibly small $P_w \approx 0$ in message waiting and reception modes, and thus their costs are negligible. This assumption allows isolation of "pure" energy costs of message transmission and to compare different localizations more effectively in terms of the number of retransmissions. The total energy costs of such an "ideal" network are related only to message transmission, which is expressed as:

$$E_p(n) = L_p(n) P_s T_s \qquad (10)$$

It should be noted that the higher the network loading, the closer is the considered network to an "ideal" one, since most of the time is actually spent transmitting messages instead of waiting. This again leads us to the necessity to consider only high network loading in the first part of the paper. Dividing the last expression by the total number of nodes in the network, it is possible to obtain an expression for the specific energy cost (energy per node) given by:

$$e_p(n) = \frac{L_p(n)}{n^2} P_s T_s \qquad (11)$$

Network optimization criteria.

The three models of localization shown in Figure 2 - clustering, perimeter bypass, and coastline bypass – are evaluated in terms of two criteria:

· total and specific energy costs of transmitting all messages to the gateway in a single cycle. The network would be regarded as more efficient at lower energy costs;

· lifespan of reference nodes t_{ref} and of the network as a whole. It is assumed that the longer is the node lifespan, the higher would be the network reliability.

Consider two options for setting the problem:

· Optimization of the operation of the UWSN in water areas for arbitrary scale (construction of a network in which the distance between nodes and the number of network nodes is determined by the probability of message delivery);

· Optimization of UWSN operation in water areas for a specified size (construction of a network for a specified water area size and number of nodes).

5 Optimizing UWSN in water areas of arbitrary scale

The water area is assumed to be unlimited in size and here it is defined as the "sensor network scale". Any number of nodes can be placed at any distance from each other, and therefore with any probability of delivery. It is assumed that the given probability lies in a wide range of $0.5 < p < 1.0$. This allows us to consider the entire range of physical conditions - from ideal conditions $p=1.0$ to extremely adverse ones $p=0.5$. It makes no sense to consider lower delivery probabilities because they either result in a high percentage of lost messages, or require so many retransmissions per message that the sensor network cannot work properly.

A short cycle of two minutes, i.e., $T=120$ s, is considered. The calculation results of the dependencies of network efficiency and reliability on sensor localization are given below to determine the optimal one at different scales.

5.1 Clustering usage model

The study of possible reference nodes localizations begins with the variant of cluster localization. The network of n^2 dimension is divided into m clusters. Each cluster is a network of smaller dimension with side $l = n/\sqrt{m}$, there is a reference node marked with a blue circle in one of corner nodes in Figure 3a. Ordinary nodes transmit messages to a reference node in each cluster, which then sends data to a mobile gateway. It is assumed that the same amount of power $P_s = 25$W is consumed in transmitting each message from the reference node to the glider.

The dynamic protocol indicates to transmit messages only to the nearest neighboring nodes down and to the right.

It is convenient to calculate the number of transmissions in a cluster L^m from Equation (1) and, multiplying it by the number of clusters, the number of transmissions in the whole system would be:

$$L^{cl} = mL^m = n^2(l - 1) \qquad (12)$$

Based on Equation (3) the number of nodes on the cluster side is independent of the network scale

$$l = \frac{n}{\sqrt{m}} = \frac{1}{\sqrt{0.05}} \approx 4.54 \qquad (13)$$

and can be chosen to be either 4 or 5.

The total number of transmissions in the network, for example, for clusters with side $l = 4$, i.e., consisting of 16 nodes (Figure 2a) at a given probability of message delivery p can be obtained from the Equations (2), (12) and (13) as follows:

$$L_p^{cl}(n) = \frac{3n^2}{p} \qquad (14)$$

Substitution of this expression into Equation (10) makes it possible to determine the energy cost of message transmission in the case of clustering.

It is easy to see that the specific costs in the case of clustering do not depend on dimension n^2, but they are proportional to the probability multiplier $3/p$. The lower the delivery probability, the higher the unit cost becomes due to the increased number of retransmissions. The total network energy consumption grows depending on its size at different probabilities of the p message delivery is shown in Figure 3. With a high probability of message delivery ($p=1$), no retransmissions are required and the power consumption is minimal. The lower the probability of delivery, the greater the energy consumption.
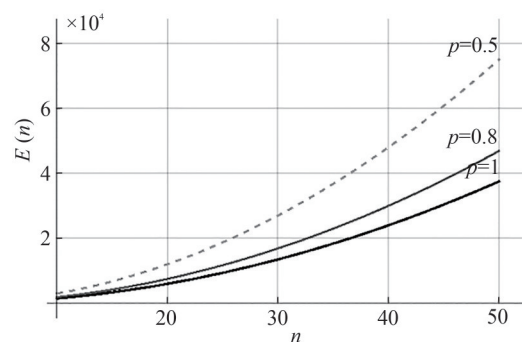
Figure 3   Total energy costs versus linear size n of the network for clustering model at different probabilities p of message delivery

Ordinary nodes on the k-th range level must pass their message and all messages from farther levels to the net corner; i.e., the reference node. Obviously, energy imbalance occurs in such a system. To balance the energy, a method is adopted by changing the reference node at each time interval, passing these rights in the cluster alternately to the upper left and lower right nodes of the net, as the least and the most loaded. The blue nodes in Figure 2, a represents the reference nodes in the first cycle, and the green nodes in the second cycle. By performing a simple transformation, it is possible to derive an equation for the number of transmissions from the k-th range level instead of equation for the total number of transmissions in the cluster.

The average node load at the k-th range level in the system for several time cycles at $l=4$ is determined by the following expression:

$$L_p^{cl}(k) = \frac{64}{(8-k)kp} - \frac{1.5}{p}, k = 1,2,3,4 \quad (15)$$

Thus, if $k=1$, the average load of the reference nodes themselves and the flow of messages (traffic) through the reference node per cycle can be determined by:

$$\lambda_{ref}^{cl} \approx \frac{7.6}{120\,p} \quad (16)$$

Knowing the nodes' load, their lifespan can be found, and consequently, the lifespan of the entire network can be estimated by substituting Equation (15) into Equation (9).

Consider a network of 2 500 nodes (linear size $n = 50$) for an example. Figure 4 shows the nodes lifespan depending on their distance from reference nodes with the cycle time of 2 minutes. The grey shaded area in the graph corresponds to the lifespan of the network or, otherwise, nodes "death time" of levels 1 to 3, i.e., 75% of all nodes in the network.

The network lifespan with delivery probability $p=1$ and cycle time of 2 minutes is approximately 580 hours and does not depend on the size of the water area, but only on the size of the cluster in the considered model.
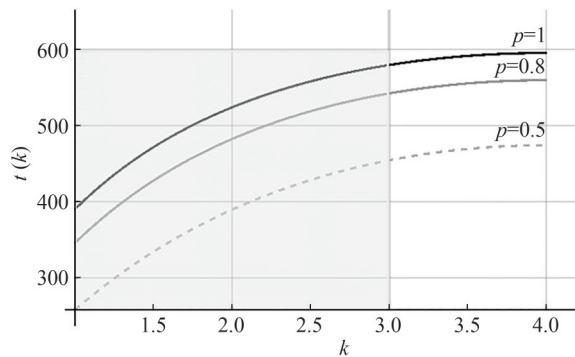
Figure 4   Nodes' lifespan at different range levels k for a clustering model for a network of 2 500 nodes with a cycle time of 2 minutes

5.2 "Perimeter bypass" usage model

The next model of FSGS, known as "perimeter bypass" as illustrated in Figure 2b, requires assumption that the reference nodes are the ones located along the perimeter of the water area. Their number can be $m = 4(n-1)$ or less, otherwise, all nodes on the perimeter are given rights of reference nodes. The range levels from the reference nodes are the perimeters of the squares nested within the "main" grid square for this localization. Summing up all the transmissions to the reference nodes, the following result is obtained:

$$L^{per} = \frac{1}{6}(n^3 - 3n^2 + 2n) + \frac{16(n-1)^2 - m^2}{4m} \quad (17)$$

If all nodes on the outer perimeter are reference, their load is defined by the following expression:

$$L_{ref}^{per} = \frac{n^2}{4(n-1)} \quad (18)$$

If only m nodes are reference nodes on each side, the same number as in the case of clustering, namely n^2/m can be obtained. The load on the reference nodes increases in this case.

The result is an unbalanced network in terms of energy costs. Balancing is then attained by letting the reference nodes on the perimeter to be selected on the assumption that their total number m is about 5% of the total number of nodes. The method of assigning reference nodes among all perimeter nodes can be close to the LEACH protocol (Mansouri and Ioualalen 2016), which sets a threshold value of sufficient energy to assign a node as a reference one. In this case each node on the perimeter will send to the gateway only a fraction of the messages coming to the perimeter, which is equal to $\frac{m}{4(n-1)}$ for several operation cycles. Multiplying the reference node load by this fraction returns the process to Equation (18).

We use the neighbor-to-neighbor routing protocol considered earlier (in the clustering problem). The entire network is divided into four triangles, each with its apex in the center of the water area and its base on one of the sides of the water area outer perimeter. The nodes within a particular triangle have information about their location and transmit messages only to the right side. This allows messages to be evenly distributed to the four sides of the perimeter. Consider a dynamic protocol, which indicates to transmit messages first to the nearest neighbor in the direction of the selected side of the perimeter, and, if unsuccessful, to both sides of that direction. The total number of transmissions on the network can be obtained by substituting Equation (17) into Equation (2), taking into account Equation (3)

$$L_p^{per}(n) = \frac{1}{6p}(n^3 - 3n^2 + 2n) + \frac{80(n-1)^2}{pn^2} - 0.012\,5\frac{n^2}{p} \quad (19)$$

Substitution of Equation (19) into Equation (10) gives the total energy cost of message transmission in the case of perimeter bypass.

The dependence of the total energy costs for one cycle on the number of nodes (from 100 to 2 500) at different probabilities of message delivery is shown in Figure 5a. When the delivery probability decreases from 1.0 to 0.5, the energy costs are doubled, similar to the clustering case, but the absolute energy costs become much higher, 100 kJ and

(a) Total energy costs

(b) Specific energy costs
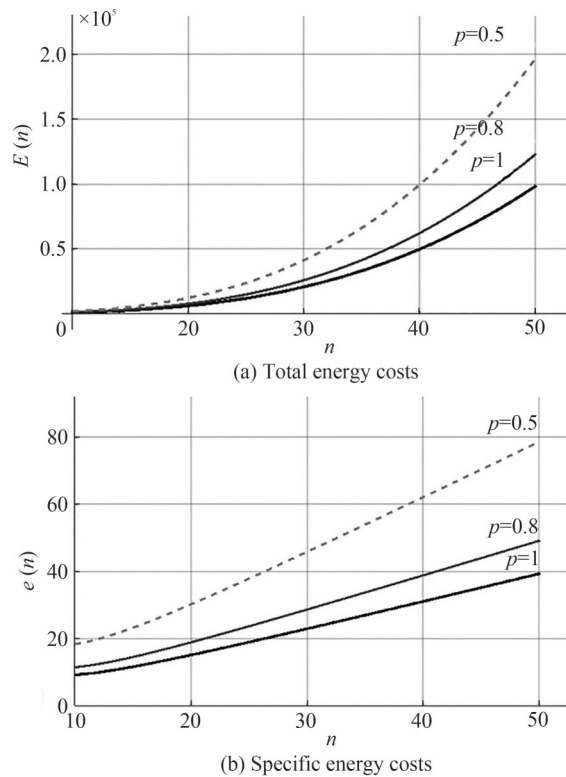
Figure 5 Variation of energy costs with linear size n of the network ($10 < n < 50$) for the "perimeter bypass" model with different message delivery probabilities

Figure 6 shows how the node lifespan grows depending on their distance from the reference nodes with a cycle time of 2 minutes for a network of 2 500 nodes. The grey highlighted area of the graph corresponding to the lifespan of the entire network.

Figure 6 Nodes lifespan at different range levels k for the "perimeter bypass" model for a network of 2 500 nodes with a cycle time of 2 minutes

5.3 "Coastline bypass" usage model

The "coastline bypass" model of FSGS shown in Figure 2c assumes that the reference nodes are located on the side of the water area that is parallel to the coastline. If the number of reference nodes is still m, then their load is equal to n^2/m, as in the previous case. The network is balanced similarly to the perimeter bypass variant, assigning 5% of the total number of nodes as reference. In this case, each node on the coastline will transmit to the gateway only a part of incoming messages equal to $\frac{m}{n}$ during several cycles. By multiplying the reference node's load by fraction $\frac{m}{n}$, the following relationship is obtained:

$$L_{\text{ref}}^{\text{coast}} = n \tag{22}$$

The total number of transmissions on the network is defined as:

$$L^{\text{coast}} = \frac{1}{2}\left(n^3 - n^2\right) + \frac{n^2 - m^2}{4m} \tag{23}$$

Consider a dynamic protocol, which indicates to transmit messages first to the nearest neighbor in the direction to the shore, and in case of failure - in both directions from it. The total number of transmissions per network can be calculated by substituting Equation (23) into Equation (2), taking into account of Equation (3):

$$L_p^{\text{coast}}(n) = \frac{1}{2p}\left(n^3 - n^2\right) + \frac{5 - 0.012\,5n^2}{p} \tag{24}$$

Substituting Equation (24) into the general Equation (10) gives the total energy cost of message transmission in the "coastline bypassing" case.

200 kJ respectively, compared to 37.5 kJ and 75 kJ for the cluster use model.

Substituting Equation (19) into the general Equation (11) determines the specific energy costs of message transmission for "perimeter bypass" model. Figure 5b shows how the specific costs increase as a function of the number of nodes.

By performing a simple transformation, the formula for the number of transmissions from the k-th range level instead of the formula for the total number of transmissions can be obtained:
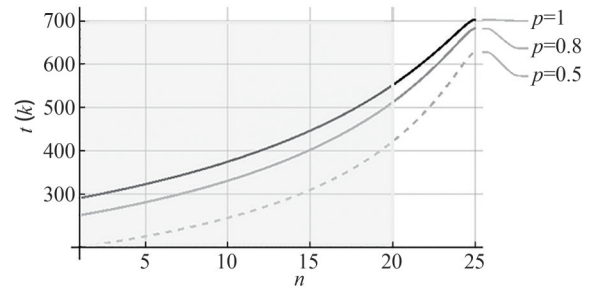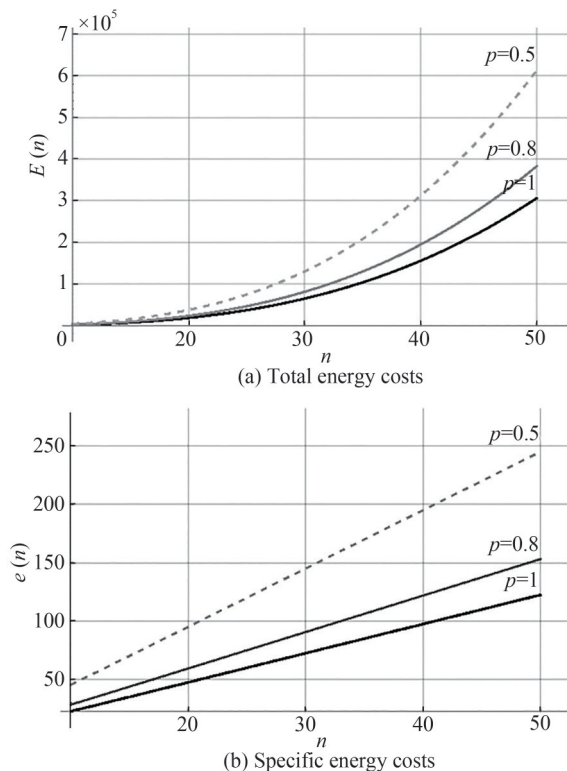
$$L_p^{\text{per}}(k) = \frac{\left(\frac{n}{2} - (k-1)\right)^2}{(n - 2k + 1)\,p}, k = 1, \cdots \frac{n}{2} \tag{20}$$

If $k=1$ the average load of reference nodes and the flow of messages through the reference node in one cycle can be determined

$$\lambda_{\text{ref}}^{\text{per}} = \frac{n^2}{480(n-1)\,p} \tag{21}$$

Knowing the nodes load, their lifespan can be determined and the lifespan of the entire network can be estimated by substituting Equation (20) into Equation (9).

Figure 7(a) shows how energy costs increase depending on the number of nodes. As the probability decreases from 1.0 to 0.5, the energy cost doubles as in the previous cases. However, the absolute values of energy costs become even higher at 300 kJ and 600 kJ, respectively, as compared to 100 kJ and 200 kJ in the "perimeter bypass" model, and 37.5 kJ and 75 kJ in the "clustering" model.

(a) Total energy costs

(b) Specific energy costs

Figure 7   Variation of energy costs with linear size n of the network ($10 <n< 50$) for "coastline bypass" model with different message delivery probabilities

Substituting Equation (24) into the general Equation (11) gives specific energy costs of message forwarding in case of "coastline bypass" model, which is depicted in Figure 7b. By performing a simple transformation, the formula for the number of transmissions from the k-th range level instead of the formula for the total number of transmissions can be obtained:

$$L_p^{coast}(k) = \frac{n - (k - 1)}{p}, \ k = 1, \cdots n \qquad (25)$$

Thus, if $k=1$, the average load of reference nodes and the flow of messages through the reference node per cycle (assuming that measurements are taken every 2 minutes) can be found:

$$\lambda_{ref}^{coast} = \frac{n}{120 \, p} \qquad (26)$$

Knowing the nodes load, their lifespan can be determined and the lifespan of the entire network can be estimated by substituting Equation (25) into Equation (9). Figure 8 shows how the nodes lifespan grows depending on their distance from the reference nodes for a network of 2 500 nodes. The grey highlighted area of the graph corresponds to the lifespan of the entire network.

Figures 4, 6 and 8 show that the lifespan of agents close to the reference nodes is very short (about 100 hours for "coastline bypass", 300 hours for "perimeter bypass" and 400 hours for "cluster"); the lifespan of agents farthest from the reference nodes is much longer–about 700 hours.

When "coastline bypass" the lifespan of various network agents has the greatest difference than for clusters and the "perimeter bypass". Clustering model is the most energetically balanced network structure, and the "coastline bypass" is the least balanced.
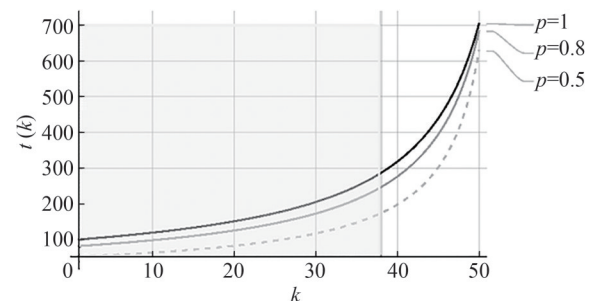

Figure 8   Nodes lifespan at different range levels k for the "coastline bypass" model for a network of 2 500 nodes with a cycle time of 2 minutes

5.4 Comparison of the three usage models in arbitrary scale water area

This section compares the proposed variants of localization at different network scales. Consider the energy costs for all localizations separately for small linear sizes (scales) $2<n<14$ (the number of nodes in the water area will thus change from 4 to 196), and for large linear sizes (scales) $20<n<50$ (the number of nodes from 400 to 2 500). In addition, the different values of probability of message delivery in the range $0.5 <p<1.0$ is considered. The first case is the "ideal" delivery probability $p=1$, which, although it cannot be implemented in practice, gives a good idea of the theoretical capabilities of the proposed localization in ideal conditions. And the second case - very low delivery probability $p=0.5$, which, of course, should be avoided in modeling, but it gives an idea about behavior of chosen network localization in extremely unfavorable conditions.

For the study of the energy efficiency of small-scale networks the results are shown in Figure 9. It is shown in Figure 9 that regardless of the probability p, the "perimeter bypass" model is the most energy effective on small networks. The "clustering" model is close enough to it. The
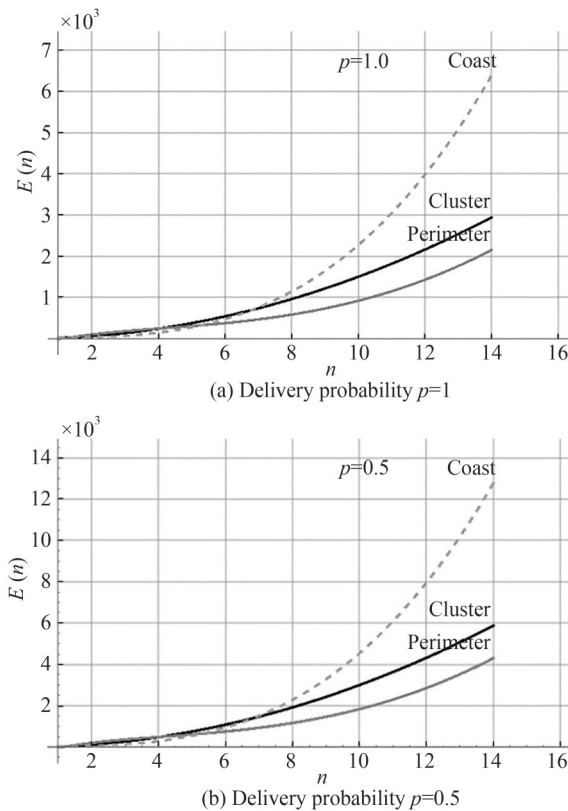
Figure 9 Total energy costs $E(n)$ versus linear size (scale) of the network in the range $2<n<14$ under three different usage models
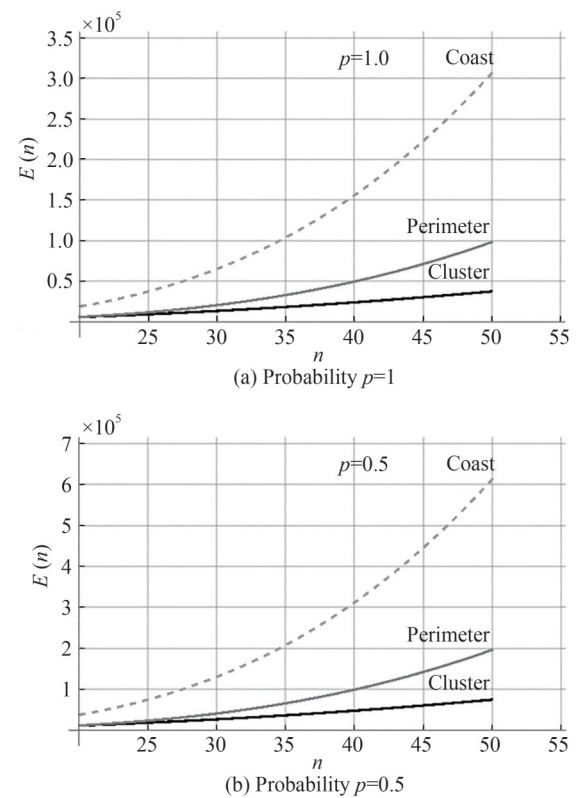

Figure 10 Total energy costs $E(n)$ versus linear size of the network ($20<n<50$) for three different usage models

energy characteristics of the "coastal bypass" model are much worse.

As for large networks, it is shown in Figure 10 that on large networks the "clustering" model is the most energy efficient. The "perimeter bypass" model is somewhat worse, but close enough to it. The "coast bypass" model has much worse energy characteristics. Of special importance is fact that the two models - "perimeter bypass" and "clustering" - are almost identical in terms of energy efficiency on an average scale of the order of $n=20$.

With regard to comparison of node lifespan for different models, it must be noted that the fundamentally different structure of the network in different localizations does not allow direct comparison of the nodes' characteristics at the k range levels. The only universal value, independent of the localization and characterizing the reliability of the system, is the lifespan of the reference nodes.

Shown in Figure 11 are the comparisons involving small-scale networks reliability. It is depicted in the figure that the "perimeter bypass" model is the most reliable on small networks in terms of lifespan. Coast bypass is reliable on very small networks (less than $n \times n=64$ nodes). Obviously, it is also shown in the figure that the reliability of the "clustering" model does not depend on the network size at all.

Comparison of Figure 11a and Figure 11b shows that,

first, the qualitative character of the dependencies does not change when the delivery probability changes. Secondly, the lifespan of the network is less sensitive to the delivery probability than the energy efficiency.

The results for the large networks are presented in Figure 12, which shows that in terms of lifespan the "perimeter bypass" usage model is still the most reliable on large networks with $n<30$. The situation changes at $n>30$. On very large networks, "clustering" usage model is the most reliable; the "coastline bypass" usage model is less reliable.

A quantitative comparison of the localizations/usage models discussed above is presented in Table 1. The best characteristics are highlighted in bolded texts in the table. The table shows that "clustering" usage model for an arbitrary unlimited water area has the best characteristics for large networks. On the other hand, the "perimeter bypass" usage model has the best characteristics for small networks.

It is important to note here that when solving real problems in a given finite water area, the simplifying assumptions made above are not quite correct. Firstly, changing the number of nodes will change the delivery probability in a finite water area. Secondly, it is necessary to match the cycle duration with the frequency of the wave glider bypassing the water area when solving a real problem. This means that the waiting time T_w between successive transmissions cannot be neglected. In addition, it is necessary to consider the various intervals of data collected by the sensors for a
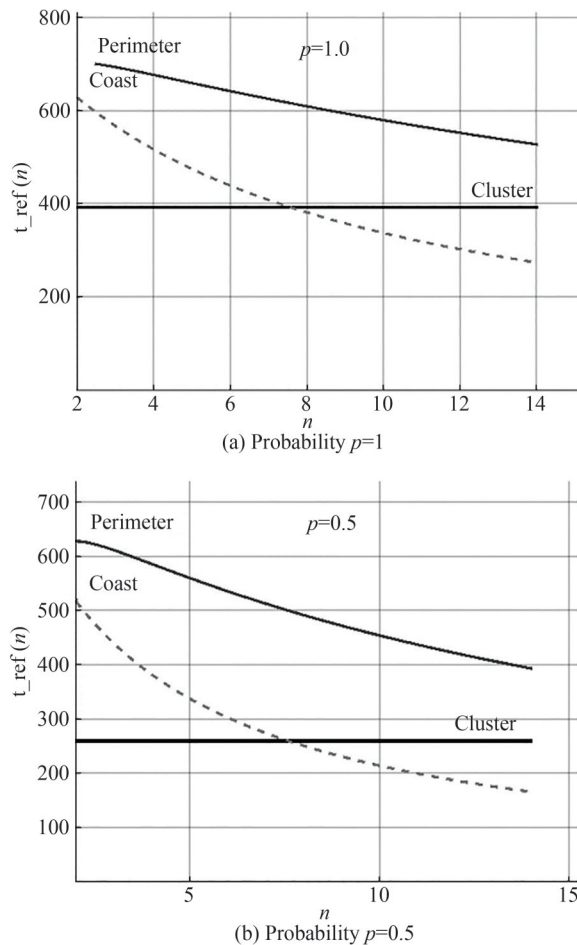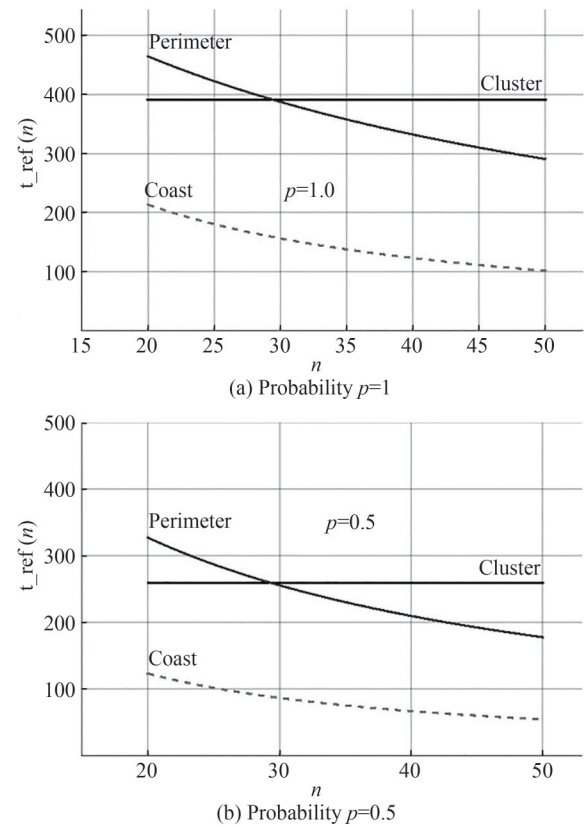
Figure 11 Reference node's lifespan $t_{\text{ref}}(n)$) versus linear size of the network ($2<n<14$) for three different usage models for a cycle time of 2 minutes

Figure 12 Reference nodes lifespan ref (n) versus linear size of the network ($20<n<50$) for three different usage models

complete picture of the possible states of the network. This is performed in the next section.

6 Optimization of UWSN within limited water area

If the size of water area, where the sensor network should be located, is specified in advance, the number of sensors would be proportional to the distance between them. The more sensors in the water area, the higher the probability of message delivery. Thus, the probability p cannot be an input parameter, but is a function of distance $p(r)$. In this work, a wide range of probability is considered: $0.5 < p\,(r) <1.0$.

The results of the computational simulation to analyze dependences of the network efficiency and reliability on the distance between nodes at the given scale of a network are presented here. They allow determination of optimum localization in the given water area, i.e., the optimum distance

Table 1 Comparison of the main characteristics for different UWSN locations in small (n=13, total 169 nodes) and large (n=50, total 2 500 nodes) networks under high network load (data captured every 2 minutes) with delivery probabilities p=1 and p=0.5

	Clustering (16 nodes)	Perimeter bypass	Coastline bypass
	Energy costs $E_p(n)(kJ)$		
$n = 13,\ p = 1.0$	2.5	1.5	6
$n = 13,\ p = 0.5$	5	3.0	12
$n = 50,\ p = 1.0$	37.5	98.2	306.1
$n = 50,\ p = 0.5$	75	196.4	612.2
	Reference node lifespan t_{ref} (h)		
$n = 13,\ p = 1.0$	391	536	282
$n = 13,\ p = 0.5$	259	398	173
$n = 50,\ p = 1.0$	391	291	102
$n = 50,\ p = 0.5$	259	178	54.3

between nodes and probability of messages delivery.

The data in the network is collected by a wave glider, which moves along a specified route, from all nodes that at least were once the reference. Let the velocity of the wave glider be $v = 0.7$ m/s for the model estimation. Then the waiting time between message transmissions begins to play a significant role, because it determines the frequency of data

collection by the glider, which is related to its speed. There-fore, it is reasonable to consider the three time ranges of data collection by the sensors (cycle time T) specified above and compare them with the time spent by the glider to by-pass all reference nodes.

6.1 Comparison of characteristics

The probability of message reception, depending on the distance between the receiver and transmitter, is defined by the Neumann-Pearson theorem according to the following relationship:

$$p_c(r) = 1 - \exp\left(-\text{SNR}^2(r)\right) \qquad (27)$$

where $\text{SNR}(r)$ is the signal-to-noise ratio at a distance be-tween the nodes:

$$\text{SNR}(r) = \frac{\text{SNR}_0}{r} 10^{-0.05\beta r} \qquad (28)$$

Let the signal-to-noise ratio at distance $r_0 = 1$ m be equal to $\text{SNR}_0 = 2\,000$ in a given water area, frequency, and hy-drometeorological conditions.

The losses associated with acoustic energy attenuation are usually characterized by the coefficient of energy atten-uation due to signal absorption and scattering by the medi-um (water absorption coefficient β), expressed in decibels per km (Sheehy and Halley 1957). For low absorption and scattering by the medium (LM) - the absorption coefficient is equal to $\beta=0.036\times10^{-3}$, for middle (M) $- \beta=0.90\times10^{-3}$ and, for high (HM) $- \beta=0.90\times10^{-3}$.

Graphically, the probability of message delivery depend-ing on the distance has the profile shown in Figure 13. The graph on Figure 13 marks the area of change of the mes-sage delivery probability in the range $0.5 <p<1.0$. The graph shows that to hit the range $0.5 <p<1.0$. in the water with low and middle β, it is necessary to place sensors at distances $500 <r<2\,500$ m, and in the water with high β - at distances $500 <r <1\,500$ m. Placing the sensors at short-er distances will not give any gain in terms of improved re-liability, as the 100% delivery probability has already been reached. It does not make sense to place sensors at longer distances, because the delivery probability at those distances is extremely low, $p<0.5$. These preliminary estimates are used later in analyzing the simulation results.

It is necessary to proceed from the given reliability re-quirements when determining the optimal network config-uration. Let us take three variants of sensors placement, providing: (i) absolute reliability (800 m from each other), (ii) average reliability (1 100 m) and (iii) extremely low re-liability (1 700 m). The data on delivery probabilities for these placements is obtained from the graph in Figure 13 and is summarized in Table 2.

Now we determine the optimal configuration of the net-
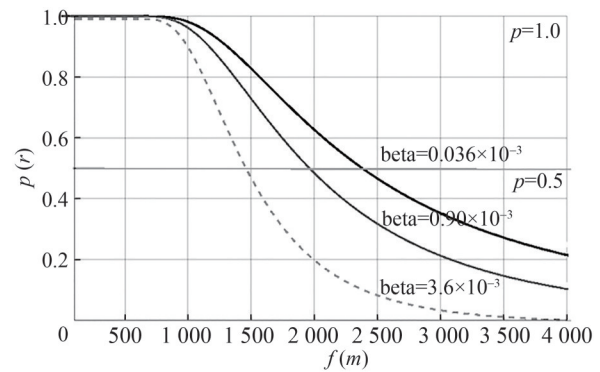
Figure 13 Probability of message delivery versus distance for different absorption coefficient β

Table 2 Probabilities of single message delivery $p_c(r)$ at fixed distances r in water areas of different muddiness

Distance between sensors, r (m)	Probability of single message delivery		
	In clear water $p_{\text{LM}}(r)$	In muddy water $p_{\text{M}}(r)$	In highly muddy water $p_{\text{HM}}(r)$
800	1.00	0.99	0.97
1 100	0.97	0.92	0.82
1 700	0.73	0.60	0.35

work in the water area of given dimensions. It is assumed here that a number of stationary nodes are placed at the bottom of the water area of size $R\times R$ for data collection, which is sent to an onshore processing center using a wave glider (mobile gateway). The distance between the nodes is equal to r. The water area is $S=144$ km² (with side $R=12$ km) for an example. The number of nodes is related to the dis-tance between them by:

$$n(r) = \frac{R}{r} + 1 \qquad (29)$$

The wave glider sends a command to begin message trans-mission during movement along the specified route and up-on entering the "visibility zone" of a particular reference node. It is assumed that the considered water area has low depths of up to 1 000 m, so the message transmission in the verti-cal direction from the reference node to the wave glider is carried out with high enough probability. In this case, due to the fact that the wave glider moves along a given route at a relatively low speed ($0.2-1.0$ m/s), the packet can be transferred by the reference node in the direction of the gateway so many times as to ensure guaranteed delivery of the message. The glider needs to move between the refer-ence nodes along some "optimal" path, which is generally a separate problem similar to the traveling salesman's prob-lem. The simplest route, assuming that the glider starts from the upper left corner and sequentially bypasses all the ref-erence nodes of clusters, is proposed in case of clustering.

The total time of data collection by the glider from all nodes (the time of single bypass of the water area) is the sum of the time spent for $\left(\sqrt{m}+1\right)$ movement horizontally and the time spent for movement downwards (see Figure 2a). Taking into account that the side length is equal to $(n-1)r$, and taking into account of Equation (3), it is possible to obtain the expression for the total time of data collection from reference nodes:

$$T_{gl}^{cl}=\left(\sqrt{m}+2\right)(n-1)\frac{r}{v} \qquad (30)$$

If the wave glider bypasses the perimeter of the water area, the nodes send messages from the center to the perimeter, Figure 3b. Then the total time for data collection from all nodes (time of a single bypass by the wave glider of the water area perimeter) can be calculated by:

$$T_{gl}^{per}=4(n-1)\frac{r}{v} \qquad (31)$$

However, if the glider bypasses the coastline of the water area (Figure 3c), then the total time for data collection from all nodes during a single bypass is determined by:

$$T_{gl}^{coast}=(n-1)\frac{r}{v} \qquad (32)$$

Substituting the numerical values into Equations (30), (31) and (32) at glider speed v=0.7 m/s, it can be concluded that the glider needs about 4.8 hours for the "coastline bypassing" model, 9.2 hours for the "perimeter bypassing" model, and finally, the reference node bypass time (in hours) for the clustering model will depend on the number of clusters m in the water area as $4.8\left(\sqrt{m}+2\right)$.

The network configuration parameters are summarized in Table 3, where (1) is clustering, (2) is perimeter bypass and (3) is coastline bypass. The table shows that the high loading of the network, which is presented in detail in the previous section, must now be abandoned as it doesn't correlate with the time spent for the glider to bypass the water area. Modeling an "ideal network" without taking into account the speed of the glider turned out to be unsuitable for the real case. It can be expected that the most suitable case

for study would be medium UWSN load, namely at once per hour.

Let us take integral characteristics, which now depend on distance between sensors, as main criteria of network optimization. The total energy cost of the network can be obtained from Equation (10):

$$E(r)=L(r)P_L T_L \qquad (33)$$

The lifespan of the reference node - from Equation (9) is:

$$t_{ref}(r,T)=\frac{240}{4.94\,L_{ref}(r)/T+0.3} \qquad (34)$$

Equation (33) represents energy consumption of the given size network at different sensor localizations, Equation (34) shows the lifespan of reference nodes (as most overloaded nodes in the network) at different sensor localizations. Using the resulting Equations, compare the three localization models in the next section.

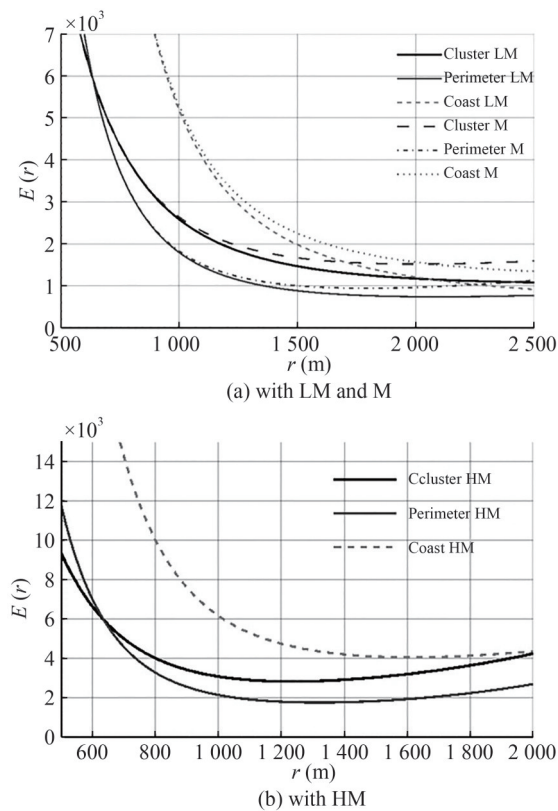## 6.2 Energy costs in three UWSN usage models for a given water area

The dependence of the total energy consumption of the network $E(r)$ for message transmission per cycle, represented by Equation (33), in the water area of S=144 m² with low LM and middle M absorption coefficient β of medium for three considered localizations (cluster model, "perimeter bypass" model and "coastline bypass" model) on the distance r between nodes are shown in Figure 14(a).

First of all, it should be noted that no matter which localization of nodes is considered, the total energy cost per cycle has a minimum. This is due to the fact that at small distances r between the nodes, their number is very large and a significant number of retransmissions are required in order for the message to be delivered to the gateway. Therefore, the energy costs are quite high in the region of small r. As the distance r increases, the number of retransmissions decreases and the energy cost drops, but at larger distances the energy cost will increase again, because the probability of successful message delivery decreases and a large number of the retransmissions from the node to its neighbor is required.

This property can be more clearly seen in the graph of

Table 3 Network configuration parameters for three variants of distance between sensors, providing high, medium and low reliability in the water area S=144 km²

r (m)	$n^2(r)$	Total number of reference nodes / number worked per cycle			Time (h) of a single bypass of the reference nodes by the wave glider		
		Clustering	Perimeter bypass	Coastline bypass	Clustering	Perimeter bypass	Coastline bypass
800	256	32/16	60/12	16/12	28.8	19.2	4.8
1 100	144	18/9	44/8	12/8	24.0	19.2	4.8
1 700	64	8/4	28/4	8/4	19.2	19.2	4.8

Figure 14 Total energy costs of the network in the water area of $S=144$ km^2

consider two possible ranges: once in 1 hour, and once in 12 hours. Equation (34) is used to analyze the lifespan of the reference node for the three considered localizations. For example, data is taken every hour, and the results are depicted in Figure 15.
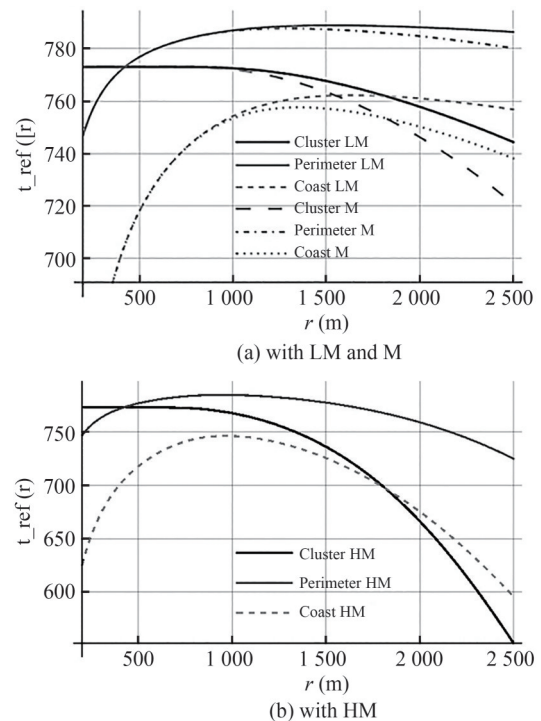
Figure 15 Lifespan of the reference node in the water area of $S=144$ km^2 when collecting data every hour

the total energy costs of the network to send messages in the water area with HM (Figure 14b). In this case, the minimum is clearly seen in the considered range of r. Indeed, in the case of clustering model, the energy costs turn out to be equal and are about 4 000 J at distances of 800 m and 1 900 m. This is quite a large cost. However, the reasons for this are different – at small distances they are due to the fact that there are a large number of sensors placed in the water area and the path to the reference node consists of a large number of hops. Each hop requires an energy expenditure. At large distances, however, the situation is the opposite – there is a small number of sensors at large distances from each other in the water area. To successfully transmit a message, it must be sent many times from sensor to sensor. This again requires energy expenditure. The minimum lies in the middle between these extreme cases. In this example it is achieved at distances of 1 200 m and the required energy is about 3 000 J. The perimeter bypass model is energetically optimal in the considered water area.

6.3 Time dependencies in three UWSN usage models for a given water area

To analyze reference node lifespan for three variants of localization it is necessary to analyze the influence of frequency of data collection by the sensor network, i. e. to

It is revealed from Figure 15 that the lifespan of the reference node has a maximum. The reasons for this are the same as for the minimum of energy costs. The clustering model has some peculiarity. The curve of reference node in clustering model has an asymptote, to which the graph tends at small distances. The presence of such an asymptote is due to the fact that the loading of the reference node at a fixed cluster size does not depend on the number of nodes in the water area and stops changing when the 100% delivery probability is reached. This always occurs at 800 m. For other localizations, reference node loading increases as the number of nodes in the water area increases even when the delivery probability is one hundred percent.

The maximum is more pronounced and is achieved at about the same distance between nodes - about 1 000 m for the "perimeter bypass" model and "coastline bypass" model, and at 800 m for the clustering model in water with high absorption coefficient β. The perimeter bypass model provides maximum reliability in the given water area.

If measurements are made once every 12 hours (twice a day), the maximum lifespan of the reference node becomes much less pronounced. This is due to the fact that the lifespan is now significantly affected not only by the number

of transmissions, but also by the waiting time between transmissions, which increases for any localization. The network becomes more robust and less sensitive to the choice of localization type.

Generally speaking, Equation (34) shows that from some T, optimization with respect to the time criterion becomes impossible, since the lifespan of the reference node tends to 800 hours, regardless of the usage model. Dependence of the reference node lifespan on the interval between measurements for the optimal distance between nodes (chosen as 1 700 m) for water area with low absorption coefficient β is shown in Figure 16. Figure 16 shows that the reference node lifespan does not exceed 800 hours for all localization models. This corresponds to the standby time without packet forwarding.
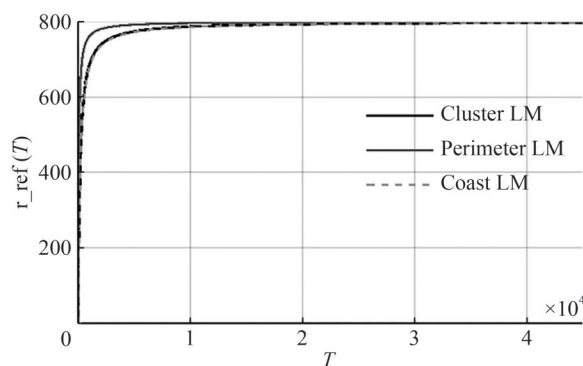
Figure 16 Reference node lifespan versus time between measurements in the water area of S=144 km^2 with low absorption coefficient $β$ when the nodes are placed at a distance of 1 700 m from each other

Consider the solution of a real problem of designing an optimal hybrid UWSN with a mobile gateway, the role of which is performed by the WG, for given design parameters as an example. As highlighted before, the UWSN, measuring the physical characteristics of the medium by bottom sensors with a frequency of once per hour, in the water area of S=144 km^2 and depth of 1 000 m with known (low) muddiness (packet delivery probability equal to p=0.97) is considered.

It follows from the results obtained above that the sensors should be located in the nodes of the orthogonal grid at distances r=1 100 m from each other in order to provide the required probability. Under this condition the total number of sensors in the water area will be equal to $n^2(r) = 144$ km^2. For energy optimization and increase in network lifespan it is necessary to choose the "perimeter bypass" model. In this case, the total number of sensors on the perimeter, alternately performing the role of reference nodes, will be equal to 44. The load of reference node can be now calculated, which turns out to be on the average 3.27 messages per cycle for given parameters. Then, if data is collected hourly by sensors, then the last reference node will accumulate 62.7 messages during a single wave-

glider bypass of the perimeter (19.2 hours).

Taking into account the technical features of the modem, let us assume that it can confidently receive/send messages within a cone with an angle of $α = 60°$. From the location of sensors, it is easy to determine that the glider at the design speed can continuously take data from the sensors moving along the perimeter of the water area. In this case, the WG have an average of about 25 s to receive one message. If, for any reason, several retries are required to receive a message from the reference node, they must be performed within the specified time interval.

This example of the hybrid UWSN architecture shows that a wave glider for relatively small water areas is a promising mobile platform that effectively acts as a gateway for retransmitting data collected in the underwater segment of the network to data centers.

7 Conclusions

This paper analyzes the main functional characteristics of three different localizations of sensor nodes for the model of using the UWSN with a mobile gateway, the role of which is performed by a wave glider. The mathematical apparatus, based on the probabilistic approach, allows to estimate the energy characteristics of the considered UWSN communication architectures, namely to determine the total energy costs of the network for message transmission and the network nodes lifespan. In this case, a simple dynamic "neighbor-to-neighbor" protocol based on competitive access to the environment is considered as an MAC protocol in model problems.

Due to the fact that the applied problems solved by the UWSN have different requirements for the network design parameters, a wide enough range of changes in such parameters as: water area size (network scale); required number of sensor nodes located in a specific water area; frequency of data collecting from sensors (cycle time); physical characteristics of the environment are considered to find the optimal solution. The paper considers three models of using the WG for servicing underwater network segment: clustering model, water area "perimeter bypass" model and "coastline bypass" model. The indicators of power efficiency and reliability depending on the above-mentioned design parameters are studied for the nodes of the underwater segment of the network within these models.

Analysis of the simulation results leads to the conclusion that the clustering model of nodes localization has the best performance and reliability characteristics for large-scale networks. On the contrary, for small-scale networks, the "perimeter bypass" model of the node localization has the best characteristics.

It is shown that for all considered localizations there are extremums (minimum) of dependences of the network to-

tal energy costs on the distance between nodes. And there are also extremums (maximum) of dependences of the reference nodes lifespan on the distance between nodes for all considered locations, except clustering model.

These results allow determination of the optimal placement of nodes for water areas with known physical characteristics of the medium to obtain the best reliability and efficiency. The presented model example of designing an optimal UWSN with a mobile gateway for a water area of a specific size shows that the "perimeter bypass" use model is practically feasible, energy-efficient and reliable for the case of data collecting by a sensor once per hour on a relatively small water area of 144 km^2. Under the accepted assumptions, the wave glider is able to reliably serve the underwater segment of the network and transmit data to the processing center. The WG has enough time to receive the messages accumulated by reference nodes (on average, about 25 s for every message) in the mode of continuous motion for depths of about 1 000 m. These conclusions confirm the fact that the WG can be used in practice as a mobile gateway of the UWSN with stationary bottom nodes.

The results obtained in the work show the need for further research based on the proposed probabilistic model of the hybrid UWSN taking into account: the spatial and uneven placement of sensor nodes in the water area, the spatially variable probability field of signal transmission.

References

Ahn J, Syed A, Krishnamachari B, Heidemann J (2011) Design and analysis of a propagation delay tolerant ALOHA protocol for underwater networks. Ad Hoc Networks, 9: 752-766. https://doi.org/10.1016/j.adhoc.2010.09.007

Barnes CR, Team NC (2007) Building the world's first regional cabled ocean observatory (NEPTUNE): realities, challenges and opportunities. In: OCEANS. IEEE, 1-8. https://doi.org/10.1109/OCEANS.2007.4449319

Brekhovskikh LM, Lysanov YuP (2003) Fundamentals of Ocean Acoustics. 3rd Edition. Springer, 278

Choudhary M, Goyal N (2021) Node deployment strategies in underwater wireless sensor network. In Proceedings 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 773-779. https://doi.org/10.1109/ICACITE51222.2021.9404617

Choudhary M, Goyal N (2022) Dynamic Topology Control Algorithm for Node Deployment in Mobile Underwater Wireless Sensor Networks. Willey Online Library. https://doi.org/10.1002/cpe.6942

Crout RL, Conlee DT, and Bernard LJ (2006) National Data Buoy Center (NDBC) National Backbone Contributions to the Integrated Ocean Observation System (IOOS). In: OCEANS 2006, 1-3. https://doi.org/10.1109/OCEANS.2006.30.307073

Domingo MC (2012) An overview of the internet of underwater things. Journal of Network and Computer Applications, 35 (6): 1879-1890. https://doi.org/10.1016/j.jnca.2012.07.012

Fattah S, Gani A, Ahmedy I, Idris MYI, Hashem IAT (2020) A survey on underwater wireless sensor networks: requirements, taxonomy, recent advances, and open research challenges. Sensors, 20, 5393, 30. https://doi.org/10.3390/s20185393

Felemban E, Shaikh FK, Qureshi UM, Sheikh AA, Bin Qaisar S (2015) Underwater sensor network applications: a comprehensive survey. International Journal of Distributed Sensor Networks. 11(11): 14. https://doi.org/10.1155/2015/896832

Grund M, Freitag L, Preisig J, Ball K (2006) The PLUSNet underwater communications system: acoustic telemetry for undersea surveillance. In: IEEE OCEANS 2006, 1-5. https://doi.org/10.1109/OCEANS.2006.307036

Han G, Zhang C, Shu L, Sun N, Li Q (2013) A survey on deployment algorithms in underwater acoustic sensor networks. International Journal of Distributed Sensor Networks, 9(12). https://doi.org/10.1155/2013/314049

Ho T, Hagaseth M, Rialland A, Ornulf J, Rodseth O, Criado R, Ziaragkas G (2018) Internet of Things at Sea: Using AIS and VHF over Satellite in Remote Areas. In: Proceedings of 7th Transport Research Arena TRA-2018, April 16-19, Vienna, 1-10. https://doi.org/10.5281/zenodo.1473565.

Kaneda Y, Kawaguchi K, Araki E, Matsumoto H, Nakamura T, Kamiya S, Ariyoshi K, Hori T (2009) Dense ocean floor network for earthquakes and tsunamis (DONET) around the Nankai trough mega thrust earthquake seismogenic zone in southwestern Japan—Part 2: Real time monitoring of the seismogenic zone. In: International Conference on Offshore Mechanics and Arctic Engineering, Paper No: OMAE2009-79599, 715-720. https://doi.org/10.1115/OMAE2009-79599

Khan A, Ali I, Ghani A, Khan N, Alsaqer M, Ur Rahman A, Mahmood H (2018) Routing Protocols for Underwater Wireless Sensor Networks: Taxonomy, Research Challenges, Routing Strategies and Future Directions. Sensors, 18, 1619, 30p. https://doi.org/10.3390/s18051619

Lan H, Lv Y, Jin J, Li J, Sun D, and Yang Z (2021) Acoustical observation with multiple wave gliders for Internet of Underwater Things. IEEE Internet of Things Journal, 8 (4): 2814-2825. https://doi.org/10.1109/JIOT.2020.3020862

Li X, Xu X, Yan L, Zhao H, and Zhang T (2019) Energy-efficient data collection using autonomous underwater glider: a reinforcement learning formulation. Sensors, 20 (13): 3758. https://doi.org/10.3390/s20133758

Lindsey S and Raghavendra CS (2012) PEGASIS: Power-efficient gathering in sensor information systems. In: Proceedings IEEE Aerospace Conference, Vol. 3, Big Sky, Montana, USA, 9-16 March, 2002. https://doi.org/10.1109/AERO.2002.1035242

Mahmood T, Akhtar F, Ur Rehman K, Ali S, Mokbal FM, and Daudpota S (2019) A comprehensive survey on the performance analysis of underwater wireless sensor networks (UWSN) routing protocols. International Journal of Advanced Computer Science and Applications, 10 (5): 590-600. https://doi.org/10.14569/IJACSA.2019.0100576

Mansouri D, Ioualalen M (2016) Adapting LEACH algorithm for underwater wireless sensor networks. In: Proceedings ICCGI 2016: The Eleventh International Multi-Conference on Computing

in the Global Information Technology, 36-40

Menon VG, Midhunchakkaravarthy D, Sujith A, John S, Li X, Khosravi MR (2022) Towards energy-efficient and delay-optimized opportunistic routing in underwater acoustic sensor networks for IoUT platforms: an overview and new suggestions. Comput Intell Neurosci, 7061617. https://doi.org/10.1155/2022/7061617

Nain M, Goyal N, Awasthi LK, Malik A (2022) A range based node localization scheme with hybrid optimization for underwater wireless sensor network. International Journal of Communication Systems. March 2022, e51470. https://doi.org/10.1002/dac.5147

Nikushchenko DV, Ryzhov VA, Tryaskin NV (2019) Modeling hydrodynamic characteristics of wave glider. XII All-Russian Congress on Fundamental Problems of Theoretical and Applied Mechanics, Ufa, August 19-24. Proceedings, 71-73

Oke P and Sakov P (2012) Assessing the footprint of a regional ocean observing system. Journal of Marine Systems, vol. 105, pp. 30–51. https://doi.org/10.1016/j.jmarsys.2012.05.009

Ovchinnikov KD, Ryzhov VA, Sinishin AA, Kozhemyakin IV (2020) Experimental study of running characteristics of wave glider. Proceedings of XV All-Russian Scientific and Practical Conference "Advanced Systems and Control Problems", Southern Federal University Press, 91-97

Rossi PS, Ciuonzo D, Ekman T, Dong H (2014) Energy detection for MIMO decision fusion in underwater sensor networks. IEEE Sensors Journal, 15 (3): 1630-1640. https://doi.org/10.1109/JSEN.2014.2364856

Riser SC, Freeland HJ, Roemmich D, Wijffels S, Troisi A, Belbeoch M, Gilbert D, Xu J, Pouliquen S, Thresher A et al (2016) Fifteen years of ocean observations with the global Argo array. Nature Climate Change, 6 (2): 145-153. https://doi.org/10.1038/NCLIMATE2872

Rizvi HH, Khan SA, Enam RN, Nisar K, Haque MR (2022) Analytical model for underwater wireless sensor network energy consumption reduction. Computers, Materials & Continua, 72 (1): 1611-1626. https://doi.org/10.32604/cmc.2022.023081

Ryabinin V, Barbière J, Haugan P (2019) The UN decade of ocean science for sustainable development. In Oceanobs'19: An Ocean of Opportunity. https://doi.org/10.3389/fmars.2019.00470

Samy A, Wang X, Hawbani A, Alsamhi S, Aziz SA (2022) Routing protocols classification for underwater wireless sensor networks based on localization and mobility. Wireless Networks, 28 (2): 797-826 https://doi.org/10.1007/s11276-021-02880-z

Sheehy MT, Halley R (1957) Measurement of the attenuation of low-frequency sound. J Acoust Soc Amer, 29 (4): 464-469. https://doi.org/10.1121/1.1908930

Su X, Ullah I, Liu X, and Choi D (2020) A review of underwater localization techniques, algorithms, and challenges. Journal of Sensors, Article ID 6403161, 24. https://doi. org/10.1155/2020/6403161

Su Y, Zhang L, Li Y, Yao X (2020) A glider-assist routing protocol for underwater acoustic networks with trajectory prediction methods. IEEE Access, 8, 154560-154572. https://doi. org/10.1109/ACCESS.2020.3015856

Toma DM, O'Reilly T, del Rio T, Headley K, Manuel A, Broring A, Edgington D (2011) Smart sensors for interoperable smart ocean environment. In: OCEANS 2011 IEEE-Spain, 1-4

Urik RJ (1967) Principles of Underwater Sound for Engineers. N.Y.-London-Toronto-Sydney, 342

Wallace DW, de Young B, Iverson S, LaRoche J, Whoriskey F, Lewis M, Archambault P, Davidson F, Gilbert D, Greenan B (2014) A Canadian Contribution to an Integrated Atlantic Ocean Observing System (IAOOS). In: Oceans-St. John's IEEE, 1-10. https://doi.org/10.1109/OCEANS.2014.7003244

Wang L, Li Y, Liao Y, Pan K, Zhang W (2018) Course control of unmanned wave glider with heading information fusion. IEEE Transactions on Industrial Electronics, 66 (10): 7997-8007. https://doi.org/10.1109/TIE.2018.2884237

Xing G, Chen Y, He L, Su W, Hou R, Li W, Zhang C, Chen X (2019) Energy consumption in relay underwater acoustic sensor networks for NDN. IEEE Access, 7: 42694-42702. https://doi.org/10.1109/ACCESS.2019.2907693

Yan J, Xu Z, Luo X (2019) Feedback-based target localization in underwater sensor networks: a multisensor fusion approach. IEEE Trans Signal Inf Process Over Networks, 5 (1): 168-180. https://doi.org/10.1109/TSIPN.2018.2866335