Relay Node Placement Considering Age of Information for Seafloor Optical Wireless Networks

Hayato Taniguchi*, Nagisa Iwami*, Tomotaka Kimura*, Takahiro Kodama†, and Jun Cheng*
*Graduate School of Science and Engineering, Doshisha University, Kyoto, Japan
Email: tomkimur@mail.doshisha.ac.jp; jcheng@ieee.org
†Faculty of Engineering and Design, Kagawa University, Kagawa, Japan,
Email: kodama.takahiro@kagawa-u.ac.jp

***Abstract*—In this paper, we consider the optimal placement problem in seafloor optical wireless networks. In a seafloor optical wireless network, data collected by sensors are forwarded to relay nodes and delivered to a sink node via multiple relay nodes. The performance of data delivery to the sink node depends greatly on the placement of relay nodes. Therefore, this paper formulates an optimization problem in which the objective function is to minimize information freshness, i.e., the age of information. Since this optimization problem is a mixed integer linear problem once the relay node locations are determined, we propose a relay node placement method using a genetic algorithm. In the proposed method, the placement position of the relay node is assumed to be a gene, and the fitness value is obtained by solving the mixed integer programming problem. Through the simulation experiments, we show that efficient data transfer can be achieved by using the proposed method for relay node placement.**

***Index Terms*—underwater optical wireless communication, genetic algorithm, mixed integer programming problem, visible light communication, optimization problem**

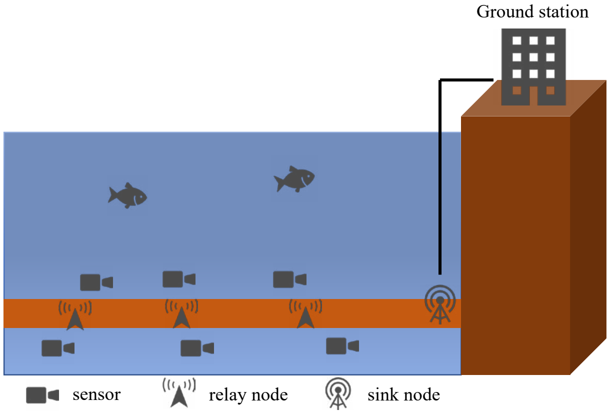


Fig. 1: Seafloor optical wireless network.

I. INTRODUCTION

In recent years, underwater monitoring systems have been studied actively [1], [2]. In some underwater monitoring applications, sensors have been deployed on the seafloor to monitor the state of certain aspects of the ocean (e.g., ocean trenches and underwater volcanoes). In many previous underwater monitoring studies, the sensor data were transmitted using sound waves [3]–[5]. This approach was used because sound waves offer good underwater communication characteristics. In other words, the attenuation of sound waves in water is low and they are more tolerant of turbidity than radio waves. However, the transmission rate of sound waves is extremely slow, which makes it difficult to transmit large volumes of data, e.g., the quantities required for photographs and video images, even though visual techniques are very useful for ocean monitoring. To overcome this problem, use of visible light for data transmission has been researched extensively in recent years [6], [7].

Although use of visible light can increase the data transmission rate, light also has the disadvantage of short propagation distances in the ocean. To overcome this disadvantage, a seafloor optical wireless network that uses multiple relay nodes for data delivery has been considered [7]. In this seafloor optical wireless network, data are transmitted via visible light communication, and a connected cable is used for the communication process from the sink node to the ground station. The communication performance of such underwater optical wireless networks is greatly dependent on the locations of the relay nodes. Therefore, in [7], the relay node deployment problem was considered. In the paper, it was assumed that the relay nodes and the sink node were deployed along a one-dimensional straight line. Under this assumption, the optimal placement of the relay nodes was analyzed mathematically. Although this study was mathematically solid, the occurrence of the information was given by a probability distribution, and it did not take the data transmission schedules at the relay nodes into account.

In this paper, we consider the problems involved in placement of the relay nodes and the transmission schedules of the data packets for underwater optical networks. In the placement problem, the relay node placements that can deliver the packets to the sink node most efficiently are determined. In contrast, in the packet transmission schedule problem, we determine the time at which each packet is transmitted over a link between the relay nodes. In this work, we formulate these problems as an optimization problem and demonstrate that the transmission scheduling problem becomes a mixed integer programming problem when the locations of the relay nodes have been determined. Furthermore, based on this approach, we propose a relay node placement method that uses a genetic algorithm. In this genetic algorithm, each gene represents the coordinates

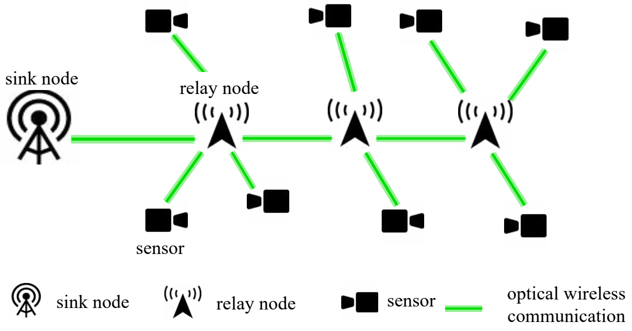


Fig. 2: System model.

of a relay node. Then, because each gene represents the location of a relay node, the fitness of each gene can be calculated by solving the transmission scheduling problem. As the fitness metric in this case, we use the age of information (AoI), which is a quantitative indicator of the freshness of the sensor information. When the AoI increases, the sink node receives information from each sensor node more frequently. The genetic algorithm then leaves the best fitting genes to obtain the relay node placement with the highest possible AoI. Through simulation experiments, we confirm that the proposed method can be used to obtain relay node arrangements with high AoI.

II. PROPOSED METHOD

A. System Model

The system comprises a sink node, N relay nodes, and S sensors deployed on the seafloor, as shown in Fig. 2. Let $\mathcal{N}$ and $\mathcal{S}$ denote the sets of the relay nodes and sensors, respectively. The sink node is deployed at 0 and all relay nodes are deployed on a half line with the sink node acting as the base point. Let x_n be the position of relay node $n \in \mathcal{N}$, and $x_n \in \mathbb{R}^+$. $(x_n, 0)$ represents the coordinates of relay node n. When the sensors are placed around the half-line, the coordinates of sensor $s \in \mathcal{S}$ are denoted by (x_s, y_s).

In this system, the sensing data are delivered to the sink node using a multi-hop approach via relay nodes. When it senses the state of the seafloor, sensor node s generates a packet that contains the required sensing information and then forwards the packet to the nearest relay node n_s. This relay node n_s then forwards the packet to an adjacent relay node that lies closer to the sink node than node n_s. This procedure is then repeated until the packet reaches the sink node. Here, let $\mathcal{Q}_n$ denote the set of relay nodes that the packet passes through from relay node n ($n \in \mathcal{N}$) to reach the sink node. In addition, $c_{s,m}$ denotes the transmission capacity of the link through which the packets from sensor s ($s \in \mathcal{S}$) are sent to the relay node m ($m \in \mathcal{Q}_n$).

$$c_{s,m} = W \log(1 + \mathrm{SNR}(r_m)),$$

where W is the bandwidth, and $\mathrm{SNR}(r_m)$ is the signal-to-noise ratio in underwater optical communications represented by the relay node m and is given by the following equation.

$$\mathrm{SNR}(r_m) = \frac{P_t D^2 \cos\phi}{4(\tan^2\theta) P_n} \frac{e^{-K r_m}}{(\epsilon + r_m)^2},$$

where the transmitted power is P_t, the noise power is P_n, the receiver aperture diameter is D, the half angle of the transmitter beam width is θ, the beam attenuation factor is K, and ϵ is a constant.

Let $\mathcal{S}_n$ be the set of sensors located nearest to the relay node n. Packet collisions occur when the sensors in $\mathcal{S}_n$ forward two or more packets simultaneously. To avoid these collisions, the proposed system considers the scheduling problem involved in determining the time that the data transmission from sensor s uses each link, where $t_{s,m}$ is the time at which a packet from sensor s starts processing for link m, and $\tau_{s,m}$ is the time required to transmit the packet via link m. Furthermore, the size B of each packet generated by each sensor is uniform, and thus $\tau_{s,m} = B/c_{s,m}$. Each sensor s senses every time T and begins forwarding new packets to the sink node. Let $t_s^{(k)}$ be the time of generation of the kth ($k = 1, 2, \ldots$) packet from sensor s.

$$t_s^{(k)} = \xi_s + (k-1)T,$$

where ξ_s is the time at which the first packet is generated. Furthermore, the delay time between generation of the packet and its arrival at the sink node is denoted by d_s, where d_s represents the delay time between packet generation and that packet's arrival at the sink node.

In terms of the sensing data, it is important to keep up with the current information, and thus in this work, we minimize the average AoI, which represents a measure of the freshness of the information. The AoI is expressed as the difference between the current time t and the time of generation of the latest packet held by the sink node. The AoI $A_s(t)$ for the sensor s is represented as follows:

$$A_s(t) = t - \eta_s(t),$$

where $\eta_s(t)$ is the time of generation of the latest packet that the sink node has among all packets generated by sensor s. Additionally, let A_s denote the time-averaged AoI for the sensor s.

$$A_s = \lim_{T_s \to \infty} \frac{1}{T_s} \int_0^{T_s} A_s(t) dt.$$

As shown in Fig. 3, because each sensor s generates a packet every unit of time T and the delay is represented by d_s, the time-averaged AoI A_s is given by

$$A_s = \frac{1}{T}\left(\frac{T^2}{2} + d_s T\right) = \frac{T}{2} + d_s.$$

Therefore, the average AoI A for all the sensors is given by

$$A = \frac{1}{S}\sum_{s=1}^{S} A_s = \frac{1}{S}\left(\sum_{s=1}^{S} \frac{T}{2} + d_s\right) = \frac{T}{2} + \frac{1}{S}\sum_{s=1}^{S} d_s.$$

TABLE I: Symbols used in formulation of the optimal problem

Symbol	Meaning
A	Average AoI
T	Time at which information from all sensors reaches the sink node (s)
d_n	Delay of sensor n (s)
$t_{n,i}$	Transmission start time for sensor n at link i (s)
$s_{n,i}$	Transmission time for sensor n at link i (s)
α_i	Next hop link for link i
$\mathcal{Q}_n$	Set of links that packets from sensor n must pass through
P_t	Transmitting power (W)
P_n	Noise power (W)
D	Receiver opening diameter (m)
ϕ	Misalignment
θ	Half-power beamwidth of transmitter
W	Bandwidth (Hz)
K	Beam attenuation coefficient (1/m)

In this paper, we consider minimization of the average AoI A via adjustment of both the packet transmission scheduling and the relay node placements.

B. Problem Formulation

The problem addressed in this paper is formalized as an optimization problem of a combination of packet transmission scheduling and relay node placement as follows.

$$\text{Minimize} \quad A \tag{1}$$

$$\text{Subject to}$$

$$A \geq \frac{T}{2} + \frac{1}{N}\sum_{n=1}^{N} d_n \tag{2}$$

$$\forall n \in \mathcal{N}, \forall i \in \mathcal{Q}_n,$$

$$d_n \geq t_{n,1} + s_{n,1} - t_{n,i}, \tag{3}$$

$$s_{n,i} = B/c_{n,i} \tag{4}$$

$$\forall n \in \mathcal{N}, \forall i \in \mathcal{Q}_n,$$

$$c_{n,i} = W \log(1 + \mathrm{SNR}(r_i)) \tag{5}$$

$$\mathrm{SNR}(r_i) = \frac{P_t D^2 \cos\phi}{4(\tan^2\theta)P_n}\frac{e^{-Kr_i}}{(\epsilon + r_i)^2} \tag{6}$$

$$\forall n \in \mathcal{N},$$

$$T \geq t_{n,1} + s_{n,1}, \tag{7}$$

$$\forall n \in \mathcal{N}, \forall i \in \mathcal{Q}n,$$

$$t_{n,i} \geq 0, \tag{8}$$

$$\forall n \in \mathcal{N}, \forall i \in \mathcal{Q}_n,$$

$$t_{n,a(i)} \geq t_{n,i} + s_{n,i}, \tag{9}$$

$$\forall n, m \in \mathcal{N}, n \neq m, \forall i \in \mathcal{Q}_n, \forall j \in \mathcal{Q}_m,$$

$$t_{n,i} \geq t_{m,j} + s_{m,j} - M(1 - x_{n,i,m,j}) \tag{10}$$

$$\forall n, m \in \mathcal{N}, n \neq m, \forall i \in \mathcal{Q}_n, \forall j \in \mathcal{Q}_m,$$

$$t_{m,j} \geq t_{n,i} + p_{\alpha,i} - M(x_{n,i,m,j}), \tag{11}$$

Equation (1) is the objective function and it minimizes the average AoI A. As discussed in the previous section, the average AoI is given by (2). The packet delay d_n for each sensor is given by (3). The time $s_{n,i}$ for which each sensor uses the link is given by (4), (5), and (6). Equation (7) constrains the sensing interval T to be longer than the total of the time taken from the starting point at which the last link was sent to the sink node plus the transmission time. Equation (8) indicates that the time must be nonnegative, (9) indicates that compliance with the order of use of the links is required, and (10) and (11) indicates that more than two packets cannot be transmitted on the same link simultaneously.

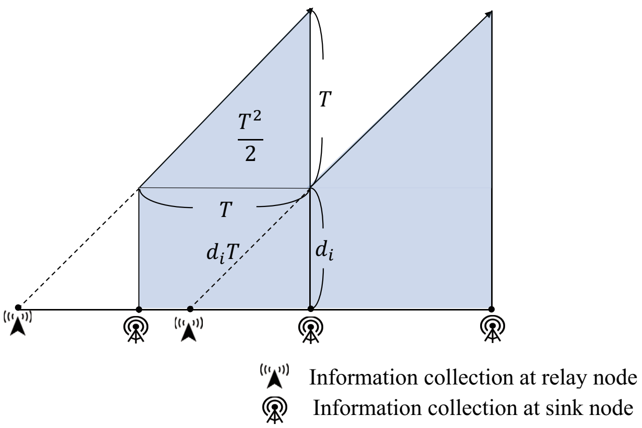


Fig. 3: Age of information.

Because the SNR given by (6) is nonlinear, it is difficult to solve this problem directly. Therefore, we solve the optimization problem by dividing it into two stages and then repeating these stages. Specifically, in the first stage, we select the relay node placements, and during the second stage, we solve the optimization problem related to the packet transmission schedule to determine the value of the objective function, i.e., AoI. When the appropriate relay node placement plan has been determined, the optimization problem with regard to the packet transmission schedule then becomes a mixed integer programming problem. The genetic algorithm is used to perform the first stage, i.e., relay node placement selection. Furthermore, the AoI value obtained during the second stage is used as the fitness value.

C. Procedures used in the proposed method

The proposed method determines the position of relay node x_n ($n \in \mathcal{N}$) and the transmission schedule by repeating two stages of the proposed process. In the first stage, to determine the most appropriate placements for the relay nodes, we use a genetic algorithm. The vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$, which represents a set of N relay node locations, is used as a gene.

First, we explain the generation of the initial population $\mathcal{X}_1$. The initial number of individuals is set to be γ. Let $\boldsymbol{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, \ldots, x_N^{(i)})$ be the genes of the ith individual. Then, the j-th element $x_j^{(i)}$ of $\boldsymbol{x}^{(i)}$ is determined using a uniform distribution between $(j/N)L - \delta$ and $(j/N)L + \delta$. δ is set to be $\delta = M/(2N)$. $\mathcal{X}_1$ is set to be $\mathcal{X}_1 = \{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(\gamma)}\}$. The location of the Nth relay node is then determined based on a uniform distribution between $L - \delta$ and L.

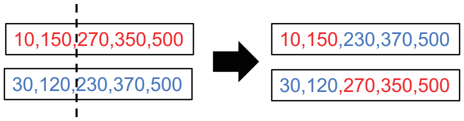


Fig. 4: Illustration of the crossover process.

Genes are improved by crossover, mutation, and selection. Let $\mathcal{X}_k$ represent the set of genes in the kth generation. We describe the procedure used to update the genes from the kth generation $\mathcal{X}_k$ to the genes of the $k+1$th generation $\mathcal{X}_{k+1}$ here. For the crossover process, our proposed method uses a single-point crossover approach. Two genes $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$ are selected at random from $\mathcal{X}_k$. For these genes, an integer n from $[1, N]$ is selected randomly, and as illustrated in Fig. 4, the elements that are less than n in $\boldsymbol{x}^{(i)}$ and greater than n in $\boldsymbol{x}^{(j)}$ are combined to form the new gene $\boldsymbol{x}'$.

$$\boldsymbol{x}' = (\boldsymbol{x}_1^{(i)}, \ldots, \boldsymbol{x}_n^{(i)}, \boldsymbol{x}_{n+1}^{(j)}, \ldots, \boldsymbol{x}_N^{(j)})$$

Each time that a new gene $\boldsymbol{x}'$ is generated, this gene is added to $\mathcal{X}_{\text{cross}}$ ($\mathcal{X}_{\text{cross}} := \mathcal{X}_{\text{cross}} \cup \{\boldsymbol{x}'\}$). At the beginning of the crossover in the $k+1$th generation, $\mathcal{X}_{\text{cross}} = \phi$. The crossover continues until γ genes are created, i.e., the crossover is repeated until $|\mathcal{X}_{\text{cross}}| = \gamma$.

Next, for the mutation process, mutations occur in each gene in $\mathcal{X}_k$ with a probability p_{muta}. When the gene that causes the mutation is $\boldsymbol{x}^{(i)}$, an integer n is then selected from $[1, N]$ and the nth element $x_n^{(i)}$ is deleted. Then, l is selected from $[1, N]$. For the elements of $x^{(i)}$, with the exception of $x_n^{(i)}$ and l, the values are arranged in ascending order, and the vector obtained with the rearranged values is then used as a new gene $\boldsymbol{x}''$. At the beginning of the mutation, $\mathcal{X}_{\text{muta}} = \phi$. Each time that a new gene is generated by mutation, $\mathcal{X}$ is updated to become $\mathcal{X}_{\text{muta}} := \mathcal{X}_{\text{muta}} \cup \{\boldsymbol{x}''\}$. The mutation process also generates γ genes ($|\mathcal{X}_{\text{muta}}| = \gamma$).

To serve as the genes of the next generation $\mathcal{X}_{k+1}$, γ genes are selected from among $\mathcal{X}_k, \mathcal{X}_{\text{cross}}$, and $\mathcal{X}_{\text{muta}}$, and the remaining genes are then discarded. In the proposed method, the fitness $f(\boldsymbol{x}^{(i)})$ of each individual is set as the reciprocal of the average AoI $1/A$. When the relay node placements have been determined, the average AoI A can then be determined by solving the mixed integer programming problem that was described in the previous section. The process used to select the genes to be retained in the next generation uses a combination of elite selection and roulette selection methods. The top N_E genes are selected from among the genes included in $\mathcal{X}_k \cup \mathcal{X}_{\text{cross}}$ using the elite selection approach, and these genes are used as the genes for the next generation. Next, selection of the remaining $\gamma - N_E$ genes is determined using a roulette selection method. In roulette selection, the gene $\boldsymbol{x}^{(i)}$ is selected with probability p_i, where:

$$p_i = \frac{f(\boldsymbol{x}^{(i)})}{\sum_{n=1}^{N} f(\boldsymbol{x}^{(n)})}.$$

This process is repeated for N_G generations, with the crossover, mutation, and selection stages being performed repeatedly. The gene with the highest fitness in the N_Gth generation of $\mathcal{X}_{N_G}$ is then used as the location for the relay node in the proposed method.

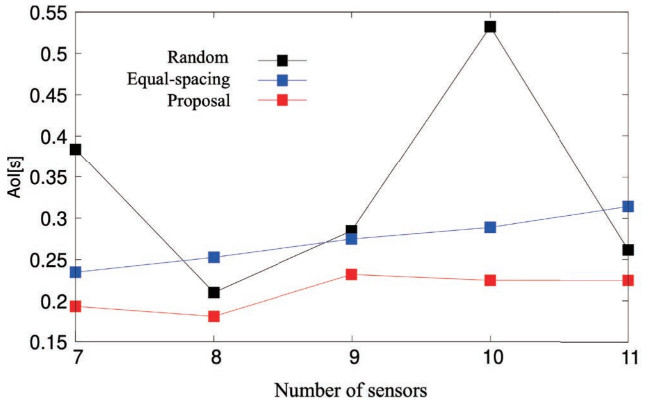


Fig. 5: Average AoI characteristics obtained using different methods.

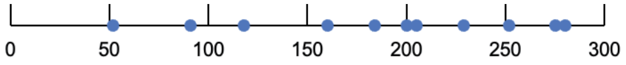


Fig. 6: Sensor positioning example.

III. Performance Evaluation

A. Overview

We demonstrate the effectiveness of the proposed method through a series of simulation experiments. In these experiments, the domain length is $L = 300$, there is only one sink node, the number of relay nodes is $N \in \{3, 4, 5\}$, and the number of sensors is $S \in \{7, 8, 9, 10, 11\}$. The sensor placement is determined randomly within $[0, L]$. Fig. 6 shows an example of the sensor placement results obtained for $S = 11$.

In addition, the parameters in the genetic algorithm used in the proposed method were set to be the values given in Table II. The number of individuals selected in the elite selection procedure was $N_E = 2$, the number of generations was $N_G = 30$, the number of individuals in each generation was set at 50, and the mutation probability was set at $p_{\text{muta}} = 0.02$.

The results of the initial random placement process, in which the initial placement was randomized, and the equally spaced placement are also shown for performance comparison purposes. In the equally spaced placement, L is divided into equal intervals, and the nth relay node is positioned at $x_n = (nL/N)$. In the initial random placement, a genetic algorithm is used in the same manner as in the proposed method, but the initial genes in this case are set at random from the entire domain. In other words, the elements x_n ($n = 0, 1, \ldots, N$) for each initial gene $\boldsymbol{x}$ are selected at random from $[0, L]$.

TABLE II: Simulation parameters

Parameter	Value
Number of sink nodes	1
N_E	2
Number of generations (N_G)	30
Number of individuals (γ)	50
Mutation probability p_{muta}	0.02
Transmission power P_t (W)	0.5
Noise power P_n (W)	2.0×10^{-6}
Receiver aperture diameter D (m)	0.2
Misalignment ϕ	10
Half-value angle of transmitter beam θ	10
Bandwidth W (Hz)	5.0×10^{8}
Beam attenuation coefficient K (1/m)	7.0×10^{-2}
ϵ (m)	1

TABLE III: Average AoI for S

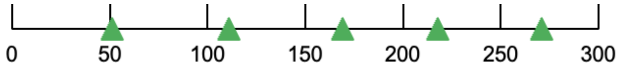
S	Proposed	Equal-spacing	Random
3	2.17	8.22	8.44
4	0.653	0.993	0.936
5	0.250	0.288	0.532

B. Results

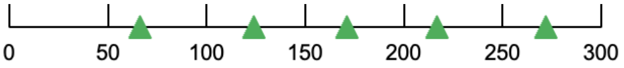
First, we present the results obtained for $L = 300$ and $N = 5$. Fig. 5 shows the average AoI as a function of the number of sensors S. For any S, the average AoI obtained using the proposed method is smaller than that obtained from the equally spaced method, and this indicates that the sensor information is being collected efficiently. Moreover, Fig. 5 shows that the average AoI obtained from the random placement method is greater than that obtained from the equally spaced method. This occurs because the performance of the genetic algorithm is heavily dependent on the initial sensor placements. In the random placement method, there is a possibility that initial placements with large distances between the relay nodes are selected. In such a placement scenario, the transmission speed of the packets to the sink node will be extremely low, and it will take a very long time for the process to converge to realize the optimal solution.

Next, Table III shows the average AoI obtained for $S = 10$. From the results in Table III, we see that the average AoI obtained using the proposed method is low for any N. Furthermore, when the number of relay nodes is small (e.g., $N = 3$), the difference between the average AoI obtained using the proposed method and that obtained via the equal-spacing method is larger, and we can see that the proposed method is more effective when N is small, rather than in the case when a sufficiently large N can be used.

Finally, Fig. 7 shows the relay node placement results obtained for the case where $S = 8$. In the relay node placements in Fig. 7 (a), the relay nodes are deployed closer to the sink node, whereas in the results from the random method shown in Fig. 7 (b), the nodes are generally arranged to be further away from the sink node. These results indicate that all packets are gathered at the relay nodes close to the sink node, which means that the average AoI obtained using the proposed method can be reduced by simply placing the relay nodes close to the sink node.



(a) Proposed method.



(b) Random method.

Fig. 7: Relay node results comparison ($S = 5$).

IV. CONCLUSION

In this paper, we propose a method for arrangement of the relay nodes in a seafloor optical wireless network that takes the freshness of the information provided into account. We considered the problems of packet transmission scheduling and relay node placement and formulated a combined solution. Furthermore, we proposed a relay node placement method that uses a genetic algorithm. Through simulation experiments, we demonstrated that the proposed method improves the average AoI by placing the relay nodes to be slightly closer to the sink node than the corresponding placements obtained via the equal spacing method.

REFERENCES

[1] E. Felemban, F. K. Shaikh, U. M. Qureshi, A. A. Sheikh, and S. B. Qaisar, “Underwater sensor network applications: A comprehensive survey,” *International Journal of Distributed Sensor Networks*, vol. 11, no. 11, p. 896832, 2015.

[2] H. Kaushal and G. Kaddoum, “Underwater optical wireless communication,” *IEEE Access*, vol. 4, pp. 1518–1547, 2016.

[3] C. Kam, S. Kompella, G. D. Nguyen, A. Ephremides, and Z. Jiang, “Frequency selection and relay placement for energy efficiency in underwater acoustic networks,” *IEEE Journal of Oceanic Engineering*, vol. 39, no. 2, pp. 331–342, 2013.

[4] F. A. de Souza, B. S. Chang, G. Brante, R. D. Souza, M. E. Pellenz, and F. Rosas, “Optimizing the number of hops and retransmissions for energy efficient multi-hop underwater acoustic communications,” *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3927–3938, 2016.

[5] L. Liu, M. Ma, C. Liu, and Y. Shu, “Optimal relay node placement and flow allocation in underwater acoustic sensor networks,” *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2141–2152, 2017.

[6] N. Saeed, A. Celik, M.-S. Alouini, and T. Y. Al-Naffouri, “Performance analysis of connectivity and localization in multi-hop underwater optical wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 18, no. 11, pp. 2604–2615, 2018.

[7] Y. Inoue, T. Kodama, and T. Kimura, “Global optimization of relay placement for seafloor optical wireless networks,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 1801–1815, 2020.