IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 29, 2018, accepted December 14, 2018, date of publication January 1, 2019,
date of current version January 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2889910

Energy-Efficient Clustering Algorithm for Magnetic Induction-Based Underwater Wireless Sensor Networks

SAI WANG, THU L. N. NGUYEN, AND YOAN SHIN, (Senior Member, IEEE)
School of Electronic Engineering, Soongsil University, Seoul 06978, South Korea

Corresponding author: Yoan Shin (yashin@ssu.ac.kr)

This work was supported by the NRF of Korea through the Korean Government (MSIT) under Grant 2016R1A2B2014497.

ABSTRACT Magnetic induction (MI) communication is a promising technology for next-generation low-power underwater wireless sensor networks (UWSNs). Clustering algorithm design becomes an important and challenging issue in today's MI-based UWSNs. In contrast to the conventional approaches which suffer from continuous movement of ocean current and traffic loads in different areas of the network, we consider a clustering algorithm based on the Voronoi diagram and node density distribution to improve the energy efficiency and to prolong the network lifetime. In particular, we propose a jellyfish breathing process for cluster head selection and an automatic adjustment algorithm for sensor nodes. The simulation results show that the proposed clustering algorithm achieves a high network capacity rate and a good equalization for the remaining energy.

INDEX TERMS Underwater sensor network, magnetic induction communication, clustering algorithm, Poisson point distribution, Voronoi diagram, jellyfish breathing process.

I. INTRODUCTION

Underwater wireless sensor networks (UWSNs) have been applied in various oceanography missions such as disaster relief, surveillance, and system control of underwater environments where the human operation is impossible. A typical network is an affiliation of sensor nodes that often operate with irreplaceable batteries and limited storages. The sensors can also dynamically form a network without any underlying infrastructure support. Their responsibilities are to sense objects (e.g., moving, temperature), to process, and to send the results to the base station directly or to multi-hop transmissions. Since two-thirds of the Earth is covered by water, many countries have funded the researches and applications on this emerging area to investigate and explore the oceans. Thus, there has been considerable progress on UWSNs including microelectronics, embedded systems and telecommunication technologies. Even though the technologies are well developed, some problems related to wireless network topology, routing protocol, localization, and energy efficiency optimization, to name a few, must be investigated further [1]. Compared to terrestrial wireless sensor networks, underwater environments are more challenging for wireless communication.

Recently, magnetic induction (MI) communication is a promising technique which is not affected by multipath propagation and fading. It uses a magnetic antenna implemented as a coil to generate a magnetic field in a high-frequency band. In order to accomplish wireless communication, a semi-static near field is utilized rather than using the propagating wave. Fundamentally different from the conventional underwater communication paradigms (e.g., electromagnetic, acoustic and optical communications), the underwater MI communication relies on the time-varying magnetic field to collect information between the transmitting and receiving coils. Since the MI communication exhibits several promising features such as constant channel condition, small antenna size, and negligible multipath-fading as well as silent and stealth underwater operations, recent research provides the fundamentals of underwater MI communications including the MI channel models, MI networking protocols design, and MI-based underwater localization [1]–[5]. A comparison on major characteristics of underwater communication technologies is given in Table 1. While many research works are mainly focused on MI-based applications for underground sensor networks, only a few researches in MI-based communication have been considered in UWSNs. Thus, the

TABLE 1. Underwater communication technologies.

Characteristic	Optical	Acoustic	EM	MI
Range	10-100 m	Up to kms	1-10 m	10-100 m
Data rate	∼ Gbps	∼ kbps	∼ Mbps	∼ Mbps
Bandwidth	10-150 MHz	∼ kHz	∼ MHz	kHz
Channel dependency concern	Major	Major	Minor	Minor
Latency	Low	High	Moderate	Low
Effective range	10-100m	∼ km	∼ 10 m	∼ 10-100m
Antenna size	∼ 0.1 m	0.1 m	∼ 0.5 m	0.15 m

MI-based applications for UWSNs are still restricted in theoretical study.

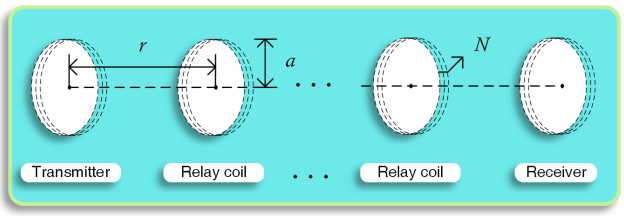


FIGURE 1. An illustration of an MI communication system.

A communication system that uses MI transmission is illustrated in Fig. 1. The magnetic antennas at a transmitter and a receiver are separated by a distance r, which depends on the coil radius a, the number of turns N, and other related factors. In order to reduce the receiving power loss, it needs to employ some relay coils between the transmitter and the receiver. Those relay coils called MI waveguides do not require any energy source and processing device. Thus, the magnetic induction from the MI transmitter can be passively relayed by them until reaching the MI receiver. However, the deployment strategy for MI waveguides to connect a large number of wireless sensor nodes is still a challenging issue. In terms of topology control, several optimal MI waveguide deployment strategies have been developed in one-dimensional (1D) [2]–[5], two-dimensional (2D) [6]–[8] and three-dimensional (3D) [9]–[13] networks. In fact, the positions of MI waveguides in 1D space are placed along a chain such as underground pipeline systems, while in 2D space they can be placed at any desired positions by following either random or regular distributions. In [9], the authors showed that the constructed network according to the Voronoi-Fermat (VF) algorithm for MI waveguides could reduce the number of relay coils and greatly enhance the network connectivity and robustness to node failures. For example, with a hexagonal tessellation topology and a random node distribution, the MI network constructed based on the VF algorithm had the optimal route because it had planar and geometrical spanner properties with lower power consumption and high energy efficiency over the network. In 3D MI-UWSNs, MI nodes can be deployed in a distributed method by adjusting their positions in the depth of underwater environments to reduce the overlapping area, thus increase the coverage of the network [10], [11].

Using the features of the Voronoi diagram, clustering techniques are often deployed for designing an energy-aware and scalable network. Thus, the simple communication process of an MI-based UWSN can be divided into three phases. (i) First, the normal sensors collect the environmental measurements and transmit them to a cluster head (CH) within their ranges. (ii) Second, the CH collects that information and sends the aggregated data to a mobile autonomous underwater vehicle (AUV). (iii) Finally, the AUV collects the data of all CHs over the network and provides information to the system, which is a collection of CHs for the next round of data aggregation. The energy resource management during network operation is an important task to prevent the creation of energy holes, thus prolongs the network lifetime. An energy hole happens when a normal node or a CH consumes its energy budget faster than other sensors. Once this phenomenon appears, no more data can be transmitted among nodes/clusters. Thus, to keep the energy conservation, we must concern how to balance load distribution among nodes during the network operations. With typical physical layers of MI technology and underwater environment characteristics, designing an appropriate clustering protocol for gathering and aggregating data in an energy efficient manner faces with different challenges and constraints such as long communication range of MI coils and dynamic aquatic environments (e.g., dynamics of ocean currents).

In this paper, we are interested in designing the clustering protocol and the dynamic CH selection. From a practical point of view, the following assumptions have been made. First, all sensors have the same sensing radius, computing capability and battery capacity. Second, in order to perform a realistic simulation of real-life applications, we adopt the Poisson probabilistic model for sensor node distribution in 3D space. In fact, in a large scale UWSN, since a neighborhood can consist of a large number of nodes, the expected number of nearest neighbors of a transmitter must grow logarithmically with the area. Also, since some nodes in the network may become invalid under the environment, the network must run a re-deployment to keep the network connectivity and coverage. Based on our previous work [10], we propose in this paper an approach to improve the high-energy node priority clustering (HENPC) protocol that can balance the cluster size during network operation. The main contributions of this paper are summarized as follows.

(i) We briefly introduce the system model and the motivation of this work.

(ii) Then, based on the established system model, we propose a clustering protocol that includes two main parts. First, in order to optimize the cluster size, we construct the jellyfish breathing algorithm based on the node contribution density. Second, to obtain an appropriate number of selected CHs, we build a node adjustment based on the Voronoi diagram.

(iii) Finally, we provide numerical results to show that the proposed scheme can achieve a higher capacity rate and longer lifetime than the conventional design [10].

The rest of the paper is organized as follows. Section II introduces a research background including the related works and the motivation of this paper. The problem is analyzed and the improved HENPC protocol is presented in Section III. In Section IV, the description of performance metrics and the detailed analysis of the simulation results are discussed. Finally, the conclusions and future work are drawn in Section V.

Notations: We use the following notations throughout the paper. A discrete random variable $X \sim Pois(\mu)$ has a 1D Poisson distribution with $\mu \geq 0$ if its probability mass function is $p(k) = e^{-\mu}\mu^k/k!$ for $k = 0, 1, \cdots$. We denote the volume of a set A in 2D space as $\mathcal{M}_A$, the number of points in set A as $\mathcal{N}_A$. A counting process $X_1, X_2, \cdots$ is called the homogeneous Poisson process [14] with intensity $\lambda \geq 0$, if it satisfies the following two properties. (a) The random variable $\mathcal{N}_A \sim Pois(\lambda\mathcal{M}_A)$. (b) For any finite set of disjoint sets $\{A_1, \cdots, A_n\}$, their numbers of points $\mathcal{N}_{A_1}, \cdots, \mathcal{N}_{A_n}$ are mutually independent.

II. RESEARCH BACKGROUND

A. RELATED WORKS

As we presented in Section I, in order to balance the overall energy consumption loads in the network (e.g., reducing the remaining energy gaps among sensor nodes), there have been a lot of research works on this issue [6]–[8]. Most of which assume that sensor nodes have the same capabilities such as communication range and initial energy, and the sensor node deployment is homogeneous in some scenarios, while we assume that the sensor nodes follow the Poisson distribution and the whole network is divided into a set of hexagonal clusters. For each cluster, the CH is located at the center of the hexagon.

In [6], the low energy adaptive clustering hierarchy (LEACH) protocol was proposed, in which the distribution and the remaining energy of sensor nodes were not considered. In this case, some clusters contain many nodes, while others contain a few ones, leading to high power consumption in large-size clusters. The LEACH–centralized (LEACH-C) protocol proposed in [7] is an extension of the LEACH. A selected CH is based on the gathered information about node location and energy status. However, although LEACH-C provides an efficient clustering, it still suffers from many drawbacks such as equal probability for CH selection and unbalanced energy loads. When the network energy is low, a node with low remaining energy still has a chance to become a CH. Once it becomes a CH, its energy will drain quickly. To prevent the early death of the CH, Dajin Wang [8] used a set of fixed hexagonal clusters. The CH is located in the center of each cluster and equipped with a powerful transmitter that can communicate with the autonomous underwater vehicle (AUV) and other CHs. This approach saves the overall power consumption of sensors in the cluster, but unfortunately increases the complexity of network deployment because the network adopts two different types of sensor nodes and needs to plan the CH positions before placing them into the cluster. In [10], we proposed a new approach called the HENPC algorithm in which a CH would be selected according to the remaining energy of sensor nodes and the geometry distance among them. Also, the ant colony optimization (ACO) was applied to find the shortest path for the data collection at the AUV. It has been shown that the HENPC protocol performs a better energy coverage and achieves a higher network capacity compared with those aforementioned conventional approaches.

Despite the fact that the HENPC achieves a good performance in some scenarios, it is still essential to develop and improve it, especially when deploying the mobile nodes to collect data in 3D-UWSNs. The mobility of free-floating nodes brings up another challenge in clustering protocol design. In order to reduce the wasted energy and to prolong the network lifetime, one needs to select the optimum CHs and cluster size for each region. As regarding to the conventional work [10], CHs are selected with a fixed radius, so the optimal number and distribution of CHs cannot be ensured. Due to the dynamic aquatic environments, the clustering process should be run periodically to adapt the nodes' locations as well as decrease communication overheads and energy consumption over the network. Moreover, adopting single AUV to collect data works well in medium-sized UWSNs, but it brings a huge end-to-end delay in larger size networks. To solve this problem, a 2D query region division algorithm based on the angle was proposed in [15], where the aggregation data are constructed in a family-set of a tree that can be used at the same time.

B. MOTIVATION

We consider a 3D-UWSN with several mobile sinks, mobile relays (i.e., AUV), and fixed nodes as illustrated in Fig. 2. In this model, the fixed nodes are randomly distributed at different depth of water on the seabed to observe or detect a phenomenon. They are usually equipped with a pressure sensor that can provide its depth information. They also have limited power and should preserve their energy. The mobile relays are AUVs, which are in charge of data collection at CHs and transmission to the mobile sink at the upper layer. The mobile sinks are deployed at the surface as buoys and being aware of their locations via a GPS system. They also have abilities to communicate with the base station and with each other on data gathering. In order to adapt to the environment dynamics and to reduce energy consumption, we adopt a routing protocol based on a reverse route search in [16] for

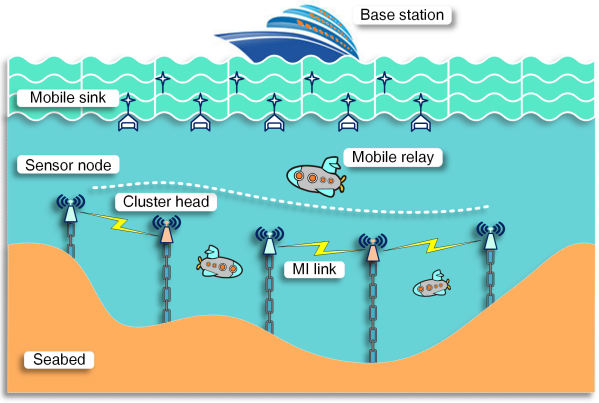


FIGURE 2. An illustration of an MI-based UWSN architecture.

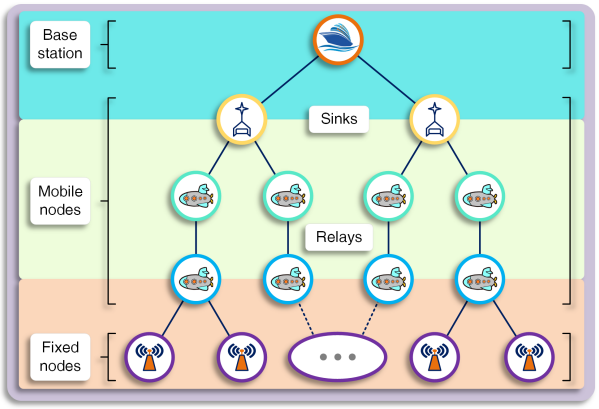


FIGURE 3. Network topology architecture in UWSNs.

data transmission, reception and acknowledgment. Following this way, by taking advantage of using several mobile nodes to collect data, we can avoid data redundancy and make the network more energy efficient and data aggregation. Fig. 3 shows the topology architecture of a three-layer network. With a clustering protocol, the network organizes the sensor nodes into different clusters in which a CH in each cluster will execute data collection and transmission tasks for the cluster members (CMs) in the cluster. Our goal is to design a clustering protocol that avoids uneven distribution problem of CHs and cluster sizes in large-scale 3D-UWSNs. Finally, these mobile relay nodes will traverse all the CHs to gather data in a collaborative monitoring mission.

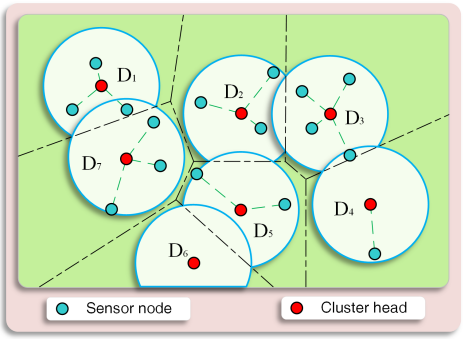
The HENPC algorithm in [10] is a dynamic clustering protocol, which consists of two phases. (i) The clustering phase selects appropriate CHs that manage their member nodes (MNs) and communicate with mobile relays; (ii) In order to ensure that every cluster has the similar number of MNs, the network reorganizes the nodes into hexagonal clusters. Before the beginning of a new round, the CHs are re-selected based on a fixed cluster radius and the remaining energy of all nodes, where the fixed cluster radius and CHs can determine which nodes are aggregated into one cluster, and these nodes cannot be selected as CHs in the current round. Unlike the aforementioned protocols in [6]–[8], the CHs are selected according to the amount of remaining energy, i.e., the sensor nodes with more energy will have a larger probability to become the CH, and they locate at the centers of the clusters in the HENPC algorithm. Thus, the HENPC algorithm can establish more efficient clusters that can balance the remaining energy of each node. However, it still has some drawbacks to be resolved. First, it assumes a uniformly distributed random network, e.g., distributed honeycomb network. Second, distribution density is not considered, so it is not suitable for randomly distributed networks in some cases (i.e., some clusters have many MNs), while others do not. Thus, those clusters with many nodes will quickly drain energy of the CH, which leads to energy hole creation. To overcome this problem, we propose in this paper an approach to improve the HENPC algorithm. We also consider the Poisson distribution to provide a better coverage for large-scale UWSNs and to adapt to the seafloor movements. Here, the sensor nodes are deployed in a randomly distributed manner at the seafloor. This difference leads to significantly different insight and the performance analysis. For a given area, the UWSN is divided into several clusters based on the jellyfish breathing process. Then, in order to balance the number of nodes among clusters, a node adjustment algorithm for sensor nodes is considered based on the Voronoi diagram after CH selection.

III. TECHNIQUE TO IMPROVE ENERGY EFFICIENCY BASED ON HENPC ALGORITHM

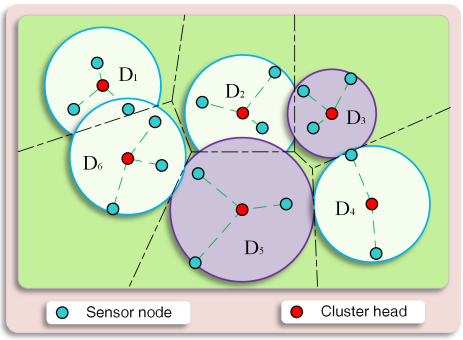
In this section, we provide a guide to go through the HENPC and how to improve it. First, the Voronoi digram is applied to find an optimal cluster to obtain the best coverage for the network in the HENPC. By this way, the network is divided into several Voronoi cells in which each cell has only one point (that is, CH) located at the center of the cell as depicted in Fig. 4(a). Here, a CH forms a circle region D_i according to the preset clustering radius. The basic idea is that they first sort the sensor list in the descending order of their residual energy. The node with the largest remaining value will be selected as the CH. The algorithm also guarantees that the distances between the CH and its CMs do not exceed the preset clustering radius. When sensor nodes are placed according to the Poisson distribution, different clusters may have different sizes due to fixed clustering radius (i.e., when node distribution is concentrated), and the number of CMs in a cluster is large. Conversely, when nodes are sparsely distributed, the number of CMs in a cluster is small (see Fig. 4(b)).

A. CLUSTER HEAD SELECTION

It is assumed that a network is divided into a predefined number of clusters n. To form clusters with the equal number of CMs in each of them, all sensor nodes communicate the information on their positions and the remaining energy levels with the CHs. Those sensor nodes with high remaining energy will have chances to become CHs in the next round. Also, clusters will perform the *Jellyfish breathing process*



(a)



(b)

FIGURE 4. An example of CH selection based on two schemes. (a) HENPC protocol. (b) Proposed protocol.

(Algorithm 2) to adaptively adjust the cluster size based on the node density. The act of breathing includes two stages; contraction and extension. Each node i calculates its own cluster range R_i, then $R = Mean(R_i)$ will be selected as the preset clustering radius. We denote $M_{\min}$ and $M_{\max}$ as the lower bound and the upper bound for the number of MNs in a circle formed by a CH. First, we select the node with the highest energy as the CH. The CH forms a disk area $\mathcal{D}_m$ according to the preset clustering radius R, and the number of MNs in $\mathcal{D}$, denoted by N_m. The contraction process is executed when $N_m \geq M_{\max}$. The clustering radius will be decreased until the number of the MNs is less than or equal to $M_{\max}$. Otherwise, the extension process will be executed, that is, if $N_m \leq M_{\min}$ the clustering radius will be increased until the number of MNs in the circle meets $M_{\max}$ or the clustering radius reaches the maximum communication distance R_{com} of the node. If $N_m \in [M_{\min}, M_{\max}]$, the sensor node will be selected as the next CH. For example, Fig. 4(b) is a result after applying Algorithm 2. We observe that the number of NMs in purple circles D_3 and D_5 is the same, even though their clustering radius are different. Thus, the aim of this algorithm is to adjust the circle radius based on the node density. Compared to Fig. 4(a), we can see that our algorithm not only reduces the number of clusters but also balances the number of NMs in each circle. Let $\mathcal{P}_m$ represents the list after the m-th cluster is selected. The details of the proposed improvement on the HENPC algorithm are described in Algorithm 1.

In the next section, we target to balance the number of MNs among the clusters according to the initial configuration

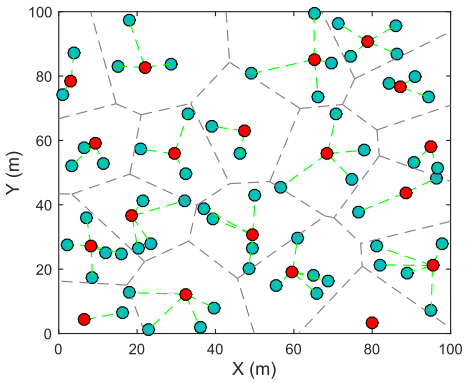
Algorithm 1 Improved HENPC Algorithm

Input: Sensor nodes $\{\mathbf{p}_1, \cdots, \mathbf{p}_N\}$ and remaining energy for the given area

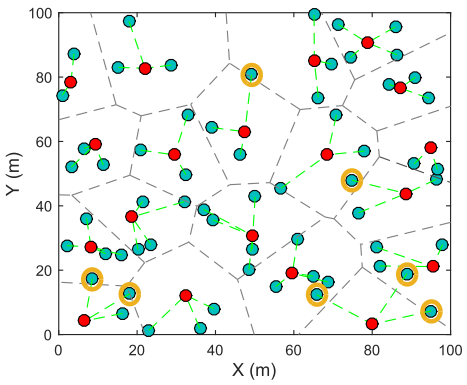
Output: CH C

Initialization: Sort amount of remaining energy of each node in the descending order

1: Set the preset cluster radius R_{com} and $m = 1$
2: **while** length($\mathcal{P}_m$) > 0 **do**
3: Take CH_m as the center and R_{com} as the radius of the circle region $\mathcal{D}_m$
4: Calculate the number of MNs N_1 in the circle
5: **if** $N_1 \in [M_{\min}, M_{\max}]$ **then**
6: Break
7: **else**
8: Jellyfish breathing process (Algorithm 2)
9: **end if**
10: **if** $(\mathbf{p}_j \in \mathcal{D}_m)$ and $(\mathbf{p}_j \in \mathcal{P}_m)$ **then**
11: Delete $\mathbf{p}_j$ from the list $\mathcal{P}_m$
12: **end if**
13: $C = CH_m$; $\mathcal{P} \leftarrow CM_m$
14: $m = m + 1$;
15: **end while**
16: **return** C



(a)



(b)

FIGURE 5. An example of node adjustment algorithm. (a) Before node adjustment. (b) After node adjustment.

settings, thus an node adjustment procedure is presented to obtain the finest clustering strategy. As shown in Fig. 5(a), we observe that the improved algorithm can reasonably divide

the network into clusters, where the red nodes are the cluster heads and the green nodes are the sensor nodes.

B. CLUSTER DIVISION ADJUSTMENT PROCEDURE

Note that by applying the jellyfish breathing process described in Algorithm 2, each normal node can indicate which CH is closest to itself. The entire network is reasonably divided into several circle regions in which CHs are located at the center and they almost have the same number of MNs. However, for accomplishing the coverage task, we must reorganize the clusters and balance the number of CMs for the overall network so that clusters have almost the same number of CMs. In order to do that, we propose an automatic node adjustment algorithm which is described in Algorithm 3.

Algorithm 2 Jellyfish Breathing Process

Input: MNs, CH_m, R_{com}, $M_{\min}$, $M_{\max}$
Output: Circle region D_m

Contraction process:

if $N_m > M_{\max}$ **then**
  while $(N_1 > M_{\max})$ **do**
    $R^{(m)} = (1-\alpha)R^{(m)}$;
    Update the circle region D_m using $R^{(m)}$
    Calculate the number of MNs N_m in the D_m
  end while
end if

Extension process:

if $N_m < M_{\min}$ **then**
  while $(N_m \leq M_{\min})$ and $(R^{(m)} \leq R_{com})$ **do**
    $R^{(m)} = (1+\alpha)R^{(m)}$
    Update the circle region D_m using $R^{(m)}$
    Calculate the number of MNs N_m in the circle
  end while
end if

return Circle region D_m

By applying the Voronoi diagram, we divided the network area into several Voronoi cells $V(CH_i)$ $(i = 1, \cdots, m)$ based on the density of sensor nodes. In particular, in this paper we assume that the distribution of sensors follows the Poisson distribution with a given node density. A set of all sensor nodes closer to a CH $CH_i \in \mathbf{C}$ than any other CH $CH_j \neq CH_i$ $(i, j = 1, \cdots, m, i \neq j, CH_j \in \mathbf{C})$ is defined as the Voronoi cell for CH_i [17]. Mathematically, for a network of n distinct nodes $\{\mathbf{p}_1, \cdots, \mathbf{p}_n\}$, we write

$$V(CH_i) = \{\mathbf{p}_k : ||CH_i - \mathbf{p}_k|| \leq ||CH_j - \mathbf{p}_k||, \forall i \neq j\}. \quad (1)$$

Although the Voronoi diagram allows MNs to find their nearest CHs, sometimes it leads to uneven allocation among MNs (see Fig. 5(a)). Thus, we have to balance the cluster size among those cells by applying an automatic node adjustment algorithm which is described in Algorithm 3. The main task of this algorithm is to check the numbers of NMs in the circles formed by given CHs, then to rearrange a number of redundant nodes according to the predefined parameters ($M_{\min}$ and $M_{\max}$). For instance, if the number of CMs N_i in a cluster is smaller than $M_{\min}$, other normal nodes nearby the CH will be assigned to the cluster until N_i reaches $M_{\min}$, and vice versa. A numerical study case is shown in Fig. 5, which compares the network division results before and after running the proposed node adjustment. From the figure, we observe that for both cases clusters are consistent with good coverage and CHs are well-distributed in the network area. After node adjustment as illustrated in Fig. 5(b), the numbers of CMs among clusters are almost equal, which yields a balanced load allocation result.

IV. PERFORMANCE EVALUATION AND ANALYSIS

A. COMPLEXITY ANALYSIS

The complexity of Algorithm 1 includes two parts. The procedure of sorting sensor nodes based on their remaining energies requires a complexity of $\mathcal{O}(n^2)$. To select a CH, it takes $\mathcal{O}(m)$ to calculate the number of MNs in each circle regions, $\mathcal{O}(n)$ to perform the jellyfish breathing process, and $\mathcal{O}(n)$ to delete the MNs in a circular region. Consequently, the total complexity of selecting m CHs is $O(mn)$, that is, the procedure of CH selection has linear complexity in the number of CHs and the number of sensor nodes, and thus can be easily implemented in real-time. The worst-case happens when m reaches n, and the corresponding complexity is $\mathcal{O}(n^2)$. Afterwards, the complexity becomes $\mathcal{O}(n^2)$. Regarding to Algorithm 3, it requires a complexity of $\mathcal{O}(m^2)$ at every adjustment scheduling to obtain the optimal performance. In particular, it requires a complexity of $\mathcal{O}(m)$ to calculate

Algorithm 3 Automatic Node Adjustment Algorithm

Input: MNs, R_{com}, and CHs $\{CH_1, \cdots, CH_m\}$
Output: Voronoi clusters with automatic node adjustment
Initialization: Cluster division by applying the Voronoi diagram $\{D_1, \cdots, D_m\}$;

for i=1:m **do**
  Calculate the number of CMs N_i in the circle D_i
  if $(N_i > M_{\max})$ **then**
    for j=1:$N_i - M_{\max}$ **do**
      Find adjacent cluster(s) D_j to the D_i with the minimum number of CMs N_j
      if $(N_j < M_{\min})$ **then**
        Find the normal node $\mathbf{p} \subset D_i$ nearest to D_j
        if $(\text{dist}(\mathbf{p}, CH_j) < R_{com})$ **then**
          The node $\mathbf{p}$ will be assigned to D_j
          Update N_i, N_j
        end if
      end if
    end for
  end if
end for
return Voronoi clusters

the number of CMs in the i-th cluster and $\mathcal{O}(m)$ to find the cluster(s) with the minimum number of CMs.

B. PERFORMANCE METRICS

In order to evaluate the performance of the proposed algorithms, the following three quantities are considered in this section: the network capacity rate, the network lifetime and the number of sensor nodes.

1) NETWORK CAPACITY RATE

First, the path loss of MI propagation underwater is given by [4]

$$L_p(r) = -10\log_{10}\frac{P_r(r)}{P_t(r)}, \tag{2}$$

where r is the transmission distance in meters, $P_r(t)$ and $P_t(r)$ are the received power and the transmitting power at the transmission range r, respectively. Therefore, the channel capacity between two MI coils is defined by

$$C = B\log_2\left(1+\frac{P_t\cdot L_P}{N_{noise}}\right)\ [\text{bit/s}], \tag{3}$$

where B is the operating bandwidth, and N_{noise} is the constant ambient noise level [1], [18]. For a network of n clusters, since we utilize mobile nodes to gather data at CHs, they have interdependent paths at each cluster. Denote t as the amount of time that the network has spent to transmit data to the base station, the network capacity is calculated by

$$C_{MI} = \sum_{i=1}^{n}\frac{T_i\cdot C_i'}{t+T_{\max}}. \tag{4}$$

Here, T_i represents the time for the i-th CH spent on collecting data and $T_{max} = \max_i T_i$ represents the longest time to collect data among clusters. In fact, we set $t = 1.7T_{max}$ and $T_{max} = T_i = 2.7$ (s) for all clusters. Thus, we have

$$C_{MI} = \sum_{i=1}^{n}\left[C_i'/2.7\right]. \tag{5}$$

Let Var_{node} be the variances among cluster sizes, the network capacity rate C_{MI} is defined as

$$\begin{aligned} C_{net} &= 1+\ln(C_{MI}/\text{Var}_{node}) \\ &\approx \ln\sum_{i=1}^{n}[C_i'/\text{Var}_{node}]\quad [\text{bit/s}]. \end{aligned} \tag{6}$$

Here, the network capacity reflects the maximum workload the network can handle per unit time, and the clustering size determines the efficiency of processing data. When there are no energy holes and the clustering sizes are similar among clusters, C_{net} becomes large, that is, the network has high throughput and high efficiency of processing data at the clusters. Conversely, it implies that the communication distance between the sensor nodes becomes larger and the number of normal nodes in each cluster is unbalanced.

2) NETWORK LIFETIME

The network lifetime is determined by the number of rounds for a network to collect data. A round is defined as the amount of time that data are aggregated from the source nodes to the CHs and collected by the mobile relays. For a sensor node, if its power is exhausted prematurely, there will be an energy hole. When many energy holes appear in the network, the connectivity of the network may become worse and the network requires more energy to collect the data, which accelerates the demise of the network. Therefore, the network lifetime is an important metric to evaluate the clustering algorithm.

C. SIMULATION RESULTS

We consider a network with N nodes that are deployed in a seafloor of 100 m$\times$100 m with the homogeneous amount of energy. The distribution of sensors follows the Poisson distribution with node density equal to 0.008. We set the maximum and the minimum number of CMs in a cluster as $M_{\max} = 4$ and $M_{\min} = 2$, respectively. The changed rate of the cluster radius is $\alpha = 10$ %. In terms of MI deployment settings, other parameters are given in Table 2 [10], [18]. Sensors are deployed in the same depth of water, which is about 30 m from the seafloor. Note that the distance between the nodes and the node position offset should meet the connectivity condition and coincide with the central concentrations [10].

TABLE 2. Simulation parameters.

Parameter	Value
Number of fixed nodes	80
Operating frequency	10 MHz
Ambient noise level	-73 dBm
Presetting clustering radius	20 m
Initial energy of the node	2 J
Percentage of invalid nodes	10 %
Minimum power of the received signal	4 pJ/s

First, in order to show how much we improve the network capacity rate from the previous work (HENPC) [10], Fig. 6 illustrates a complete statistical report on network capacity rate with different network lifetime. For a given round, we observe that the proposed scheme (''Improved HENPC'') achieves higher rates than the HENPC algorithm. For example, at the 100-th round, in the HENPC protocol about 86 % C_{net} obtain high values: 82.5 % of $C_{net} \in [16, 17.5]$ and 3.5 % of $C_{net} \geq 17.5$, while in the improved HENPC protocol about 99 % achieve high values: 34 % of $C_{net} \in [16, 17.5]$ and 65 % of $C_{net} \geq 17.5$. For longer lifetime, e.g., at 1900-th round, we can see that around 3 % of $C_{net} \leq 17.5$ in Fig. 6(a) and around 26 % of $C_{net} \geq 17.5$. Also, the death time of the first node in the improved HENPC protocol is later than the HENPC (after 2100 rounds). This is because we distribute the network nodes energy more efficiently. The adaptive clustering protocol with balanced cluster sizes over the network is able to accomplish less energy to transmit data and thus prolong the lifetime of the network.

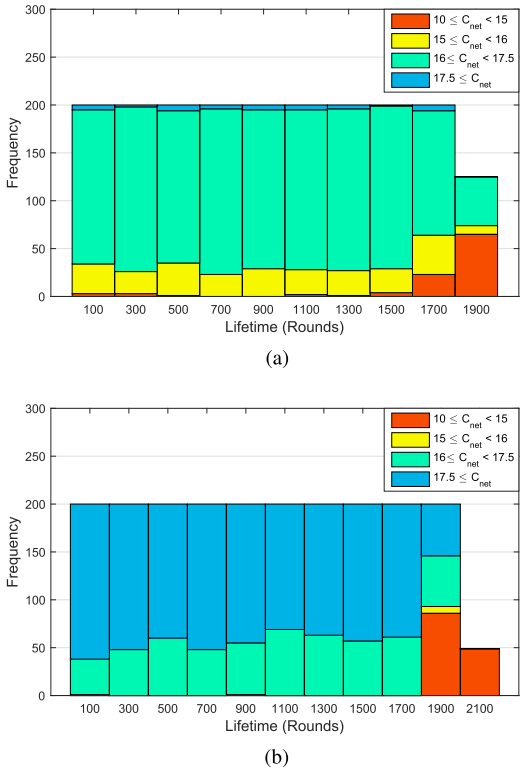


FIGURE 6. Network capacity rate versus lifetime. (a) HENPC protocol. (b) Improved HENPC protocol.

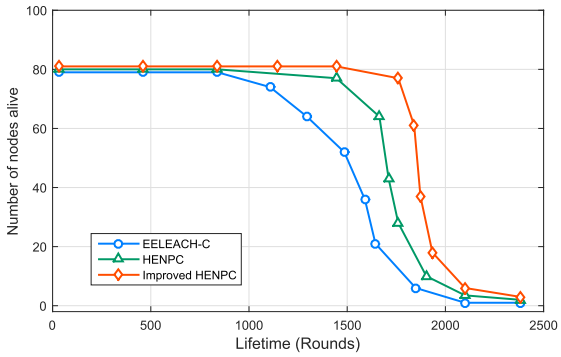


FIGURE 7. Lifetime versus number of nodes alive.

Fig. 7 shows the relationship between the number of nodes alive and the network lifetime of the three schemes: EELEACH-C [7], HENPC [10] and our improved HENPC. We observe that the number of nodes alive in the proposed approach surpasses those in the EELEACH-C and the HENPC protocols at the same round. The network using the EELEACH-C protocol stops its life at about 750 rounds and the HENPC can maintain the lifecycle to about 1400 rounds, while the improved HENPC protocol can prolong its lifetime to 1750 rounds. This is because the improved HENPC can balance the energy consumption so that it avoids premature energy holes, thus improves the network lifetime. Furthermore, we observe that the three clustering algorithms have different downfall intervals of the number of nodes alive. As the lifetime increases, the downfall interval of the proposed algorithm is the trajectory of values on the round interval [1750, 2150], while those intervals are focused on [1500, 2150] in the EELEACH-C and [1800, 2150] in the HENPC, respectively.

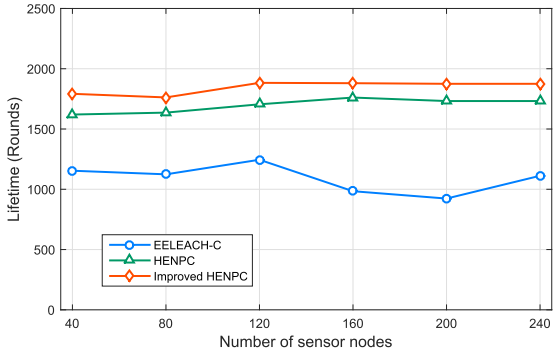


FIGURE 8. Number of sensor nodes versus lifetime.

For the network described above, simulations to determine the average lifetime are done for various number of sensor nodes as shown in Fig. 8. Under the same input conditions, the proposed scheme can keep most of the nodes alive in the network, that is, it can balance the energy consumption more effectively than the HENPC and the EELEACH-C protocols. Thus, the first death node appears later. From the obtained results, the proposed scheme extends the lifetime period by 600 rounds as compared to the EELEACH-C protocol and by 200 rounds as compared to the HENPC protocol. Moreover, it performs more stable than the EELEACH-C due to the inherited characteristics from the HENPC. This means that the proposed protocol is more energy-efficient and robust than the other two, thus it allows sensor nodes to function well for a long time.

V. CONCLUSION AND DISCUSSION

In this paper, we study the clustering strategy for MI-based UWSNs, where sensor nodes are deployed with the Poisson distribution. In order to obtain a clustering protocol with saving the energy consumption, based on the idea of the conventional HENPC, we modify the CH selection process by considering the dynamic change of sensor nodes' residual energies and the number of CMs in the clusters. The proposed scheme determines the number of CHs based on the density of sensor nodes in the given area and adaptively adjusts the numbers of CMs in the clusters. Simulation results show that our improved HENPC protocol significantly outperforms the conventional HENPC and EELEACH-C algorithms in terms of network capacity rate and network lifetime.

In summary, the proposed scheme is a dynamic protocol, i.e., the CH selection is based on the residual energy and geometry distances of sensor nodes. From the perspective of saving energy, multihop aggregation data and nodes with high residual energy preferentially selected as CHs, can effectively balance the energy consumption for the entire network. On the other hand, under the condition on full coverage of

the network, we provide an effective way for collecting data by deploying mobile nodes (e.g., AUV, mobile sink) for a large-scale UWSNs. However, there are some open issues that we need to work on. For instance, using mobile nodes may take extra-costs, so we have to balance the trade-off between energy efficiency and hardware costs. Also, how to avoid the collision at the surface as well as analyzing the complexity of operation and maintenance avoidance collision should be also investigated.

REFERENCES

[1] I. F. Akyildiz, P. Wang, and Z. Sun, "Realizing underwater communication through magnetic induction," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 42–48, Nov. 2015.

[2] J. Jiang, K. Song, G. Wei, R. Lu, Q. Zhang, and C. Zhu, "Capacity and bandwidth analysis of symmetric two-coil resonant magnetic communication using frequency divarication," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 370–373, 2015.

[3] H. Guo, Z. Sun, and P. Wang, "Multiple frequency band channel modeling and analysis for magnetic induction communication in practical underwater environments," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6619–6632, Aug. 2017.

[4] M. C. Domingo, "Magnetic induction for underwater wireless communication networks," *IEEE Trans. Antennas Propag.*, vol. 60, no. 6, pp. 2929–2939, Jun. 2012.

[5] B. Gulbahar and O. B. Akan, "A communication theoretical modeling and analysis of underwater magneto-inductive wireless channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3326–3334, Sep. 2012.

[6] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. IEEE HICSS*, Jan. 2000, pp. 1–10.

[7] M. Tripathi, M. S. Gaur, V. Laxmi, and R. B. Battula, "Energy efficient LEACH-C protocol for wireless sensor network," in *Proc. IEEE CIIT*, Oct. 2013, pp. 402–405.

[8] D. Wang, "Cluster subdivision towards power savings for randomly deployed WSNs—An analysis using 2-D spatial Poisson process," in *Proc. IEEE ICICS*, Irbid, Jordan, May 2016, pp. 156–161.

[9] Z. Sun and I. F. Akyildiz, "Optimal deployment for magnetic induction-based wireless networks in challenged environments," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 996–1005, Mar. 2013.

[10] S. Wang, T. L. N. Nguyen, and Y. Shin, "Data collection strategy for magnetic induction based monitoring in underwater sensor networks," *IEEE Access*, vol. 6, no. 1, pp. 43644–43653, Dec. 2018.

[11] K. Akkaya and A. Newell, "Self-deployment of sensors for maximized coverage in underwater acoustic sensor networks," *Comput. Commun.*, vol. 32, nos. 7–10, pp. 1233–1244, May 2009.

[12] H. M. Ammari and S. Das, "A study of k-coverage and measures of connectivity in 3D wireless sensor networks," *IEEE Trans. Comput.*, vol. 59, no. 2, pp. 243–257, Feb. 2010.

[13] T. E. Abrudan, Z. Xiao, A. Markham, and N. Trigoni, "Underground incrementally deployed magneto-inductive 3-D positioning network," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4376–4391, Aug. 2016.

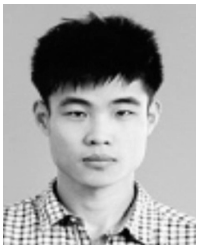
[14] R. Pasupathy, "Generating homogeneous Poisson process," in *Wiley Encyclopedia of Operations Research and Management Science*. Wiley, 2010, pp. 1–11.

[15] S. Wan, Y. Zhang, and J. Chen, "On the construction of data aggregation tree with maximizing lifetime in large-scale wireless sensor networks," *IEEE Sensors J.*, vol. 16, no. 20, pp. 7433–7440, Oct. 2016.

[16] J. Wang, W. Shi, L. Xu, L. Zhou, Q. Niu, and J. Liu, "Design of optical-acoustic hybrid underwater wireless sensor network," *J. Netw. Comput. Appl.*, vol. 92, pp. 59–67, Aug. 2017.

[17] S. Wan, Y. Zhao, T. Wang, Z. Gu, Q. H. Abbasi, and K. R. Choo, "Multi-dimensional data indexing and range query processing via Voronoi diagram for Internet of Things," *Future Gener. Comput. Syst.*, vol. 91, pp. 382–391, Feb. 2019.

[18] Z. Sun, I. F. Akyildiz, S. Kisseleff, and W. Gerstacker, "Increasing the capacity of magnetic induction communications in RF-challenged environments," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3943–3952, Sep. 2013.



SAI WANG received the B.S. degree in communication engineering from the Shandong University of Science and Technology, Qingdao, China, in 2016. He is currently pursuing the M.S. degree in information and telecommunication engineering with Soongsil University, Seoul, South Korea. His research interests include magnetic induction communication and wireless underwater sensor networks.

THU L. N. NGUYEN received the B.S. degree in mathematics and computer science from the Ho Chi Minh City University of Science, Ho Chi Minh, Vietnam, in 2011, and the M.S. degree from Soongsil University, Seoul, South Korea, where she is currently pursuing the Ph.D. degree in information and telecommunication engineering. Her research interests include wireless localization, RF-based energy harvesting, and magnetic induction communication.

YOAN SHIN received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, South Korea, in 1987 and 1989, respectively, and the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, in 1992. From 1992 to 1994, he was with Microelectronics and Computer Technology Corporation, Austin, TX, USA, as a Member of Technical Staff. Since 1994, he has been with the School of Electronic Engineering, Soongsil University, Seoul, where he is currently a Professor. From 2009 to 2010, he was a Visiting Professor with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada. His research interests include magnetic induction communication, localization, and compressive sensing. He has been serving as an Organizing/Technical Committee Member for various prominent international conferences, including VTC 2003-Spring, ISITA 2006, ISPLC 2008, APCC 2008, ISIT 2009, APWCS 2009, APWCS 2010, APCC 2010, ICTC 2010, APWCS 2011, ICTC 2011, ISAP 2011, APWCS 2012, and APCC 2012. In particular, he was a Technical Program Committee Chair of APWCS 2013 and a General Co-Chair of APWCS 2014 and APWCS 2015.