IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 28, 2020, accepted November 17, 2020, date of publication November 26, 2020,
date of current version December 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3040731

Data Collection Algorithm of a 3D Wireless Sensor Network That Weighs Node Coverage Rate and Lifetime

BANTENG LIU[1], YOURONG CHEN[1], JINHAO WAN[1], PING SUN[1], ZHANGQUAN WANG[1], TIAOJUAN REN[1], ZEGAO YIN[2], AND RENGONG ZHANG[3]

[1]College of Information Science and Technology, Zhejiang Shuren University, Hangzhou 310015, China
[2]College of Engineering, Ocean University of China, Qingdao 266003, China
[3]Zhejiang Yugong Information Technology Company Ltd., Hangzhou 310015, China

Corresponding author: Yourong Chen (jack_chenyr@163.com)

This work was supported in part by the Public Welfare Technology Application and Research Projects of Zhejiang Province of China under Grant LGG20F010009, Grant LGF19F010005, and Grant LGG19F010011; in part by the Education Department Project of Zhejiang Province of China under Grant Y202045157, and in part by the Natural Science Foundation of Zhejiang Province of China under Grant LQ18F030006.

ABSTRACT Considering the movement of a sink node in this study to solve the data transmission problem of sensor nodes in a 3D network environment, we propose a Data Collection Algorithm of a 3D wireless sensor network (DCA_3D) that weighs node coverage rate and lifetime. DCA_3D establishes a data collection optimization model that weighs node coverage rate and lifetime with the constraints of the sink node's moving path selection, data flow, energy consumption, and link transmission. Subsequently, DCA_3D calculates the fitness value of the sink node's moving path by solving the data collection optimization model with the known sink node's moving path. Then it uses a modified artificial bee colony algorithm to solve the moving path selection problem of the sink node, and finally obtains the optimal scheme. In the scheme, sink node can find the optimal moving path, whereas the sensor node can find the optimal data communication path. The simulation results show that regardless of the moving path length of the sink node, the maximum data collection hops of the sink node and the number of static sensor nodes change, DCA_3D can find the optimal moving path of the sink node. It can improve the coverage rate of the sensor nodes, the network lifetime and average data transmission amount, and reduce the average energy consumption variance and the average packet loss rate. DCA_3D outperforms the state-of-arts such as RAND, GREED, EDG_3D, and ANT.

INDEX TERMS 3D wireless sensor networks, data collection, network lifetime, packet loss rate.

I. INTRODUCTION

Wireless Sensor Networks(WSNs), including one or several kinds of sensors, such as visual, temperature, sound, infrared, radar, and seismic sensors, have been widely studied and applied in recent decades [1], [2]. Sensor nodes work together for network monitoring tasks in dangerous environments (such as volcanoes, radiation, and toxic chemicals) and can be applied in disaster rescue, military action, underwater detection, and other application fields. WSNs are static in most applications, and their nodes report data periodically and do not move. However, static WSNs have two prominent problems. First, as the nodes are distributed in harsh environments, users can hardly reach the designated monitoring areas, and manual deployment is unsuitable for node deployment. Only random deployment methods, such as airdrop and scattering, can be used. However, random deployment can easily lead to the uneven distribution of nodes. Second, owing to the numerous communication tasks that involve data transmitting and data transporting of other nodes away from the sink node, the nodes near the sink node tends to run out of energy. Non-uniform node distribution and unbalanced energy consumption easily cause energy hole problems in monitoring areas, leading to network division and lifetime declination [3]–[5]. Thus, the WSN technology should

The associate editor coordinating the review of this manuscript and approving it for publication was Tie Qiu.

be improved to balance node energy and prolong network lifetime.

Many scholars presently focus on the moving path selection, data collection of sink nodes and the issues regarding WSNs that are distributed in 2D planes. They establish and solve the network lifetime optimization models when sink nodes stay at several fixed positions [6] and [7], but do not consider the process on how to design the moving path of sink nodes. [8] proposes a sink node scheduling algorithm with optimal coverage. The algorithm improves the particle swarm optimization and genetic algorithm to solve the optimal coverage position and moving path of the sink node. [9] proposes a predictable path selection algorithm of sink nodes to reduce energy consumption between sensor nodes and improve network lifetime. [8], [9] study the moving paths of sink nodes, but the data collection method of the sink nodes only uses a single hop. Comparing with RAND algorithm which randomly selects the unpassed grids, [10] proposes a greedy scanning data collection strategy (GREED) for large-scale wireless sensor networks. In GREED, the mobile sensor node randomly selects one of the closest grid in the sets that have not passed. RAND and GREED use graph theory to select the moving path, but the selected path are not necessarily the global optimal path. Then [11] establishes the autonomous moving model of sensor nodes calculated by the bacterial foraging optimization algorithm and obtains the autonomous moving scheme of sensor nodes that covers the network area. [12] proposes a hyper-heuristic framework to schedule the movement of sink node and maximize the network lifetime. Consider moving path selection and optimization, [13] and [14] establish and solve network lifetime optimization models with mobile sink nodes. However, [5]–[14] only consider the data collection of mobile sink nodes in 2D WSNs and ignore 3D environments. Therefore, the path selection methods of mobile sensor nodes are difficult to apply to 3D environment.

Some scholars presently focus on problems of path planning and data collection of nodes in 3D environments. [15] searches for a path to decompose key positions in 3D space and proposes a 3D path planning algorithm based on the A* algorithm. [16] divides the 3D structure of obstacles into several layers according to the structure's height and the residence position of each layer, and proposes an approximate 3D Euclidean shortest-path algorithm to find the shortest path to the destination. However, [15], [16] consider 3D path planning but not the path planning and data collection of sink nodes. [17] proposes the fuzzy clustering method to find the optimal cluster to minimize the energy consumption of nodes, and uses the particle swarm optimization method to find the optimal cluster head, then solves the network splitting problem caused by the decrease of total network energy consumption. [18] and [19] propose data routing algorithms based on eccentricity. The algorithms divide the network into 3D subspaces and select centroid nodes to calculate data routing in an eccentric region of any node in a subspace. This algorithms reduce the energy consumption of nodes in the selection stage, the energy consumption of internal communication, and the number of data transmission hops. However, [17]–[19] only consider the data collection algorithm when a sink node is stationary, and the energy hole problem remains unresolved. [20] proposes efficient data gathering algorithm in 3D linear underwater wireless sensor networks (EDG_3D) using sink mobility. EDG_3D considers the linear movements of mobile sink nodes and courier nodes, then establishes and solves an optimization model to maximize the network lifetime and throughput, and minimize the data transmission delay. This scheme reduces the communication energy consumption of mobile sink nodes and courier nodes, but the linear movements of these nodes are too simple.

In summary, the recent studies consider the movement of sink nodes to overcome the energy hole problem and reduce the energy consumption of nodes, but the movement and data collection of sink nodes are mostly about 2D WSNs, and they disregard 3D environments. The 2D moving path selection methods of sink nodes are difficult to apply to 3D environments. The studies on 3D WSNs mainly focus on path planning to destination nodes and data collection of static sensor nodes, but give less attention to the problems of 3D moving paths, data collection of sink nodes, and lifetime optimization of networks. The moving direction, stay time, data collection range, moving path length, and other influencing factors of sink nodes can affect the coverage rate and network lifetime. Many scholars tend to only study one or more aspects, ignoring the impact of other aspects on coverage rate and network lifetime in 3D network environments. The above-mentioned algorithms are also difficult to apply to practical projects in 3D environments. Therefore, in consideration of the issues mentioned above, we propose a Data Collection Algorithm of a 3D WSN (DCA_3D) that weighs node coverage rate and lifetime. Briefly, our contributions are as follows:

1. We study on a cubic grid generation method in the 3D environment, and express the constraints of the sink node's moving path selection, data flow, energy consumption, and link transmission by mathematical formulas. Then we propose the calculation formulas of node coverage rate and lifetime, and establish the model that weighs the node coverage rate and lifetime.

2. We propose the modified artificial bee colony to solve the optimization model and obtain the optimal scheme. Then we can obtain the optimal moving path of the sink node and the optimal data communication paths of the sensor nodes. The obtained scheme can improve the node coverage rate and lifetime and balance the node energy consumption.

The rest of the paper is organized as follows. In Section II, we describe the algorithm assumptions and principle. In Section III, we establish the optimization model. In Section IV, we describe the algorithm solution scheme. The simulation results are presented in Section V. Finally, we conclude the paper and describe the future work in Section VI.

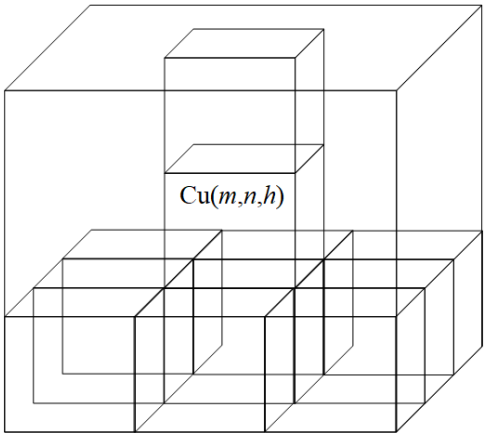


FIGURE 1. Cubic grid division diagram.

II. ALGORITHM ASSUMPTIONS AND PRINCIPLE

This study regarding the DCA_3D assumes the following:

(1) Multiple static sensor nodes and a mobile sink node exist in the 3D monitoring area, and the mobile sink node can move to any cube center in this area.

(2) Static sensor nodes have the same initial energy and communication radius.

(3) The sensor nodes and the sink node can obtain their own positions through GPS, BDS, and other satellite positioning modules or positioning algorithms.

(4) The initial energy of the sensor nodes is limited. As the initial energy of the sink node is large and the lifetime of the sink node is long, the energy of the sink node can be considered infinite in a limited network lifetime.

(5) The sink node moves and stays at the cube center of the monitoring area. If the sensor nodes in the vicinity can find the data communication paths to the sink node, then the perceived data are sent to the sink node in the multi-hop mode; otherwise, the data are stored in the cache.

Fig. 1 presents the 3D cuboid monitoring area divided into several cuboid grids of the same size, with each cuboid grid coded according to left-to-right, front-to-back, and top-to-bottom principles. $Cu\,(m_v, n_v, h_v)$ denotes the h-th cubic grid from bottom to top in the n-th row from front to back and in the m-th column from left to right. Fig. 2 shows the sensor nodes distributed in the cube's monitoring area and the mobile sink node randomly distributed in the center of a cubic grid whose positions are known. When the network is run, the mobile sink node collects the information of the sensor nodes. If the mobile sink node at a certain position can receive the routing information of sensor nodes by multi-hop routing method, those sensor nodes transmit the data to the sink node through multi-hop paths, otherwise they store the data in the cache. However, two problems in DCA_3D need to be solved. The first problem is knowing how to determine the moving path selection of the sink node and the data communication between the sink node and sensor nodes and knowing how to design the constraints of mobile selection, data flow, energy consumption, and link transmission to establish the data collection optimization model of the sink node that weighs the node coverage rate and lifetime. The second problem is knowing how to use the modified artificial bee colony algorithm to calculate the optimization model and obtain the optimal scheme. The specific solutions to the two problems are discussed subsequently.

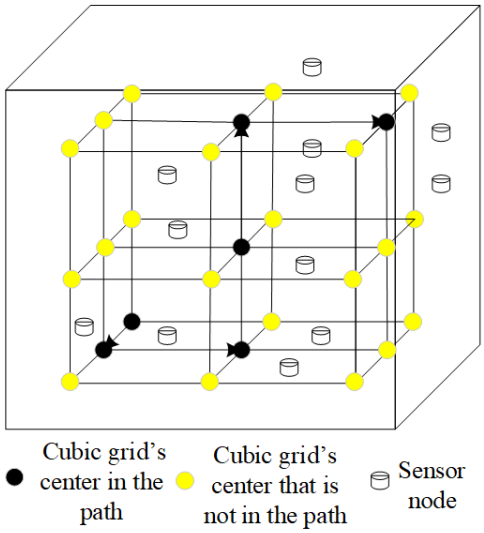


FIGURE 2. Principle diagram in DCA_3D.

III. ESTABLISHMENT OF THE OPTIMIZATION MODEL

In this section, we model a data collection optimization model that weighs node coverage rate and lifetime. Table 1 provides a complete list and definition for all symbols in this section.

A. MOVEMENT PATH SELECTION OF THE SINK NODE

The neighbor grid set of $Cu\,(m_v, n_v, h_v)$ is given by

$$L_{Cu} = \left\{ \begin{array}{ll} Cu\,(m_v - 1, n_v, h_v) & Cu\,(m_v, n_v - 1, h_v) \\ Cu\,(m_v, n_v, h_v - 1) & \\ Cu\,(m_v + 1, n_v, h_v) & Cu\,(m_v, n_v + 1, h_v) \\ Cu\,(m_v, n_v, h_v + 1) & \end{array} \right\} \quad (1)$$

Considering that part of the cubic grid is at the edge of the monitoring area, the front, back, left, right, top, or bottom part of the grid does not exist. The following processing is needed for L_{Cu}: when $m_v - 1 = 0$, delete $Cu\,(m_v - 1, n_v, h_v)$ from L_{Cu}; when $n_v - 1 = 0$, delete $Cu\,(m_v, n_v - 1, h_v)$ from L_{Cu}; when $h_v - 1 = 0$, delete $Cu\,(m_v, n_v, h_v - 1)$ from L_{Cu}; when $m_v = m_{\max}$, delete $Cu\,(m_v + 1, n_v, h_v)$ from L_{Cu}; when $n_v = n_{\max}$, delete $Cu\,(m_v, n_v + 1, h_v)$ from L_{Cu}; and when $h_v = h_{\max}$, delete $Cu\,(m_v, n_v, h_v + 1)$ from L_{Cu}. Finally, the set L_{p_v} of the neighbor grids of each cubic grid $Cu\,(m_v, n_v, h_v)$ can be obtained.

When $z_{p_v}^{p_w} = 1$, and the center of the grid $Cu\,(m_v, n_v, h_v)$ is the neighbor center of the grid $Cu\,(m_w, n_w, h_w)$, the sink node moves from the center of the grid $Cu\,(m_v, n_v, h_v)$ to the center of the grid $Cu\,(m_w, n_w, h_w)$. Then several factors need to be considered in selecting the moving path of the sink node. First, the mobile sink node only chooses the neighbor grid's center of the current grid for its next residence position. The

TABLE 1. The Notations Used in the Description of the Data Collection Model.

Notation	Definition
X_{cu}	The state indicator matrix whether all grids in the monitoring area are in the moving path of the sink node
$m_{\max}$	The maximum column number of cubic grids from left to right in the same horizontal plane
$n_{\max}$	The maximum row number of cubic grids from front to back in the same horizontal plane
$h_{\max}$	The maximum layer number of cubic grids from bottom to top
L_{Cu}	The neighbor grid set of Cu (m_v, n_v, h_v)
$z_{p_v}^{p_w}$	The movement status indicator of the grid's center
p_v	The center of cubic grid Cu (m_v, n_v, h_v)
x_{p_v}	The state value of position p_v
p_s	The initial position of the sink node
p_d	The end position of the sink node
d_{ij}^p	The distance between nodes i and j when a sink node stays at a residence position
L_{ij}^p	The neighborhood symbol of nodes i and j
$d_{\max}$	The maximum communication distance of the node
V	The set of all sensor nodes
V_s^p	The set of sensor nodes within the data collection range of the sink node when the sink node stays at position p
V_c^p	The set of sensor nodes which is not in the data collection range of the sink node when the sink node stays at position p
F_{ij}^p	The data transmission amount between nodes i and j when the sink node stays at the position p
S_i	The data sensing rate of node i
t_p	The stay time of the sink node at the position p
N_i^p	The set of neighbor sensor nodes of sensor node i when the sink node stays at the position p
E_i^{elec}	The energy consumption constant of the circuit of sensor node i
γ	The loss coefficient of data transmission
$\varepsilon_i^{\text{fs}}$	The signal amplification constant of sensor node i
N_{Path}	The number of times in which the sink node moves along the path during the network lifetime
$f_{\max}$	The maximum transmission rate of node i
C_i^p	The cached data of node i when the sink node stays at the position p
$C_{\max}$	The maximum storage capacity of the sensor node
T_i	The lifetime of sensor node i defined as the time when node i is in a working state
L_i^p	The state indicator whether sensor node i is within the data collection range of the sink node when the sink node stays at position p
T	The average lifetime of a node
N	The number of static sensor nodes in the monitoring area
$Cover$	The coverage rate of the sensor nodes

status indicator from the current grid to the non-neighbor grid is 0, that is,

$$z_{p_v}^{p_w} = 0, p_w \notin L_{p_v} \tag{2}$$

Second, except for the initial and end grids, the other grids require a sink node to move in and out by no more than once. Moreover, no duplicate centers of the grid should appear in the moving path. Let the initial position of the sink node be $x_{p_s} = 1$. The state of each grid is calculated from the initial position as follows:

$$x_{p_w} = \sum_v x_{p_v} z_{p_v}^{p_w} \tag{3}$$

Then, if the position p_v is the initial position or the end position of the sink node, then $x_{p_v} = 1$. If the position p_v is in the moving path of the sink node, then $x_{p_v} = 2$, else $x_{p_v} = 0$. Therefore, the state of each grid needs to satisfy the following conditions:

$$x_{p_v} = \begin{cases} 1, & if\ p_v \in \{p_s, p_d\} \\ 2, & if\ p_v is in the moving path \\ 0, & others \end{cases} \tag{4}$$

Third, when the sink node enters the current grid from the previous grid, it cannot return to the previous grid, that is, the sink node cannot have a self-cycling condition.

$$z_{p_v}^{p_w} + z_{p_w}^{p_v} \le 1,\ \forall v, w\ and\ v \ne w \tag{5}$$

In summary, when the initial position of the sink node is determined, and the state indicators of all grids' centers conform to *Eqs*. (2), (4), and (5), the residence position at next moment can be calculated from initial position according to *Eq*. (3). Then, we can obtain the moving path of the sink node by repeated calculation.

B. DATA COLLECTION OF THE SINK NODE

When a sink node stays at a grid's center along the moving path to collect data from the sensor nodes, the data collection problem becomes a data collection problem of the multiple sink nodes staying at different positions. When a sink node stays at a residence position, the distance between nodes i and j is calculated using

$$d_{ij}^p = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{6}$$

where (x_i, y_i, z_i) denotes the 3D coordinates of node i. L_{ij}^p is

$$L_{ij}^p = \begin{cases} 1, & d_{ij}^p \le d_{\max} \\ 0, & others \end{cases} \tag{7}$$

By transmitting and receiving hello packets, the set of sensor nodes within the data collection range of the sink node V_s^p when the sink node stays at position p is obtained. Then, $V = \{V_s^p, V_c^p\}$.

If the mobile sink node at a certain position can receive the routing information of sensor nodes by multi-hop routing method, those sensor nodes transmit the data to the sink node through multi-hop routing. As the transmitting data of the sensor nodes are composed of sensing data, cached storage data from the previous time, and data received from neighbor nodes [13], the data flow constraint is calculated as

$$\sum_{j \in N_i^p} L_{ij}^p F_{ij}^p = S_i t_p + C_i^{p-1} + \sum_{j \in N_i^p} L_{ji}^p F_{ji}^p,\ \forall i \in V_s^p,\ \forall p \tag{8}$$

In practical wireless communication, the sensor nodes mainly consume energy due to data transmission and reception, but the data sensing, computing, message packet transmission, and other energy consumption is negligible and can be ignored. Therefore, the energy consumption of the data received by sensor node i only considers the electronic energy consumption of the node circuit when the sink node stays at position p, that is, $\sum_{j \in N_i^p} F_{ji}^p E_i^{elec}$, where E_i^{elec} is related to the characteristics of the node hardware. The energy consumption of the transmitting data includes the electronic energy

consumption of the node circuit and the electronic energy consumption of the signal amplifier, that is, $\sum_{j\in N_i^p} F_{ij}^p(E_i^{elec} + \varepsilon_i^{fs} d_{ij}^{p\gamma})$, where $\gamma \in [2, 4]$ denotes the loss coefficient of data transmission, and ε_i^{fs} is related to the characteristics of the node hardware. Then, the energy consumption constraint of the nodes [21] is given by

$$N_{Path}\left(\begin{array}{l}\sum_p \sum_{j\in N_i^p} L_{ji}^p F_{ji}^p E_i^{elec} \\ +\sum_p \sum_{j\in N_i^p} L_{ij}^p F_{ij}^p (E_i^{elec} + \varepsilon_i^{fs} d_{ij}^{p\gamma})\end{array}\right) \le E_i,\ \forall i \in V_s^p \tag{9}$$

As the bandwidth resource of the link and the total amount of the data transmitted by the link are limited, the link transmission constraint [13] can be expressed as

$$L_{ij}^p F_{ij}^p + L_{ji}^p F_{ji}^p \le t_p f_{\max},\ \forall j \in N_i^p,\ \forall i \in V_s^p,\ \forall p \tag{10}$$

C. ESTABLISHMENT OF OPTIMIZATION MODEL

When sensor nodes are within the data collection range of the sink node, the sensor nodes transmit the cached data to the sink node, and $C_i^p = 0$. When the sensor nodes are not within the data collection range of the sink node, the sensor nodes sense the data and store them in the cache. If the cache capacity exceeds its maximum capacity, then the earliest data are discarded to be able to store the latest data. The cached data of node when the sink node stays at the position p is updated.

$$C_i^{p+1} = \begin{cases} C_{\max}, & C_i^p + S_i t_p \ge C_{\max},\ \forall i \in V_c^p \\ C_i^p + S_i t_p, & C_i^p + S_i t_p < C_{\max}, \forall i \in V_c^p \\ 0, & \forall i \in V_s^p \end{cases} \tag{11}$$

where $C_{\max}$ denotes the maximum storage capacity of the sensor node. The lifetime of sensor node i is defined as the time when node i is in a working state.

$$\begin{aligned} T_i &= \left(\begin{array}{l} E_i / \sum_p \sum_{j\in N_i^p} L_{ji}^p F_{ij}^p E_{elec} \\ +\sum_p \sum_{j\in N_i^p} L_{ij}^p F_{ij}^p (E_{elec} + \varepsilon_{fs} d_{ij}^{p\gamma})\end{array}\right) \sum_p L_i^p t_p \\ &= N_{Path} \sum_p L_i^p t_p \end{aligned} \tag{12}$$

where L_i^p corresponds to whether sensor node i is within the data collection range of the sink node when the sink node stays at position p. The average lifetime of a node is defined as

$$T = \sum_i T_i / N \tag{13}$$

As mentioned above, after a network startup, a sink node stays at the center of the cubic grid and moves to collect the data of the sensor nodes. A sensor node senses data and judges itself within the data collection range of the sink node. In such a case, the sensor node transmits the data to the sink node in the multi-hop mode; otherwise, it updates the cache space. According to the above analysis, the data collection algorithm needs to maximize the network life and the coverage rate of sensor nodes. The problem is converted into the following optimization model that weighs node coverage rate and network lifetime :

$$\begin{aligned} &\max(N_{Path} \sum_i \sum_p L_i^p t_p * Cover) \\ &\text{s.t. moving path selection constraints: } Eqs.\ (2) - -(5), \\ &\quad \text{data collection constraints: } Eqs.\ (8) - -(11), \end{aligned} \tag{14}$$

where $Cover$ is the ratio of the number of covered sensor nodes to the total number of sensor nodes when the sink node moves along its moving path.

IV. ALGORITHM SOLUTION

Karaboga first systematically proposes artificial bee colony (ABC) algorithm and its improved algorithm [22]. Since then, many scholars have proposed new models and made improvements and performance analyses [23]–[25]. The ABC algorithm based on bee foraging behavior has been successfully applied in various fields and is proven to be more efficient than other optimization algorithms. Many algorithms improve ABC algorithm to solve the target optimization problems and achieve certain results. But those algorithms do not address the problem of sink node's moving path selection in 3D wireless sensor networks, and cannot directly be applied to solve the model ($Eq.14$). Therefore, we propose a modified ABC algorithm to solve the model ($Eq.14$). In particular, the moving path of a sink node is used as a food source, and the optimal solution of the optimization model is found by employed bee phase, onlooker bee phase, and scout bee phase.

A. FITNESS VALUE CALCULATION OF FOOD SOURCE

The moving path selection constraints mainly affect $\sum_p L_i^p t_p/(D_{ave})$, while the data collection constraints mainly affect N_{Path}. If the moving path of the sink node is known, which means that the food source of the bee has been found, then $\min(\sum_p L_i^p t_p)/(D_{ave})$ is known, and the model ($Eq.$ 14) is transformed into the following optimization model:

$$\begin{aligned} &\max(N_{Path} \sum_i \sum_p L_i^p t_p * Cover) \\ &\text{s.t. data collection constraints: } Eqs.\ (8) - -(10). \end{aligned} \tag{15}$$

By solving the optimization model ($Eq.$ 15), the optimal solution is obtained when the moving path of the sink node is known. Its fitness value is calculated using the following formula:

$$fitness = N_{Path} \sum_i \sum_p L_i^p t_p * Cover \tag{16}$$

The specific solving process of the optimization model ($Eq.15$) is as follows. Set $\bar{F}_{ij}^p = L_{ij}^p F_{ij}^p$ and $q = 1/N_{path}$, then

the optimization model (*Eq*. 15) is converted into

$$\min(q) \tag{17}$$

$$\sum_{j\in N_i^p} \bar{F}_{ij}^p = S_i t_p + C_i^{p-1} + \sum_{j\in N_i^p} \bar{F}_{ji}^p, \forall i \in V_s^p, \forall p \tag{17a}$$

$$\left(\sum_p \sum_{j\in N_i^p} \bar{F}_{ji}^p E_i^{elec} + \sum_p \sum_{j\in N_i^p} \bar{F}_{ij}^p (E_i^{elec} + \varepsilon_i^{fs} d_{ij}^{p\gamma})\right) \leq qE_i, \forall i \in V_s^p \tag{17b}$$

$$\bar{F}_{ij}^p + \bar{F}_{ji}^p \leq t_p f_{\max}, \forall j \in N_i, \forall i \in V_s^p, \forall p \tag{17c}$$

$$C_i^{p+1} = \begin{cases} C_{\max}, C_i^p + S_i t_p \geq C_{\max}, \ \forall i \in V_c^p \\ C_i^p + S_i t_p, C_i^p + S_i t_p < C_{\max}, \ \forall i \in V_c^p \\ 0, \ \forall i \in V_s^p \end{cases} \quad \forall p \tag{17d}$$

The optimization model (*Eq*. 17) is a typical linear programming problem, and it has many variables. If the moving path is known, then the data collection of the sink node along the moving path is converted into data collection optimization models at several certain positions. When the sink node stays at position p, the data collection optimization model is

$$\min(q_p) \tag{18}$$

$$\sum_{j\in N_i^p} \bar{F}_{ij}^p = S_i t_p + C_i^{p-1} + \sum_{j\in N_i^p} \bar{F}_{ji}^p, \ \forall i \in V_s^p \tag{18a}$$

$$e_i^p = \left(\sum_{j\in N_i^p} \bar{F}_{ji}^p E_i^{elec} + \sum_{j\in N_i^p} \bar{F}_{ij}^p (E_i^{elec} + \varepsilon_i^{fs} d_{ij}^{p\gamma})\right) \leq q_p E_i, \ \forall i \in V_s^p \tag{18b}$$

$$\bar{F}_{ij}^p + \bar{F}_{ji}^p \leq t_p f_{\max}, \ \forall j \in N_i^p, \ \forall i \in V_s^p \tag{18c}$$

where e_i^p denotes the energy consumption of sensor node i when the sink node stays at position p.

Theorem 1: The solution of the data collection optimization model (*Eq*. 17) of the sink node is composed of the solution of the data collection optimization models (*Eq*. 18) when the sink node stays at several positions, and $\hat{q} = \sum_p \tilde{q}_p$.

Proof: Let the optimal solution of the node communication in the model (*Eq*. 18) be $\tilde{F}_p$ and the minimum value of the objective function be $\tilde{q}_p$ when the sink node stays at any position p on its moving path. Let the optimal solution in the data collection optimization model (*Eq*. 17) of the sink node be $\hat{F}$ and the minimum value of the objective function be $\hat{q}$. As the sink node collects data at different positions, constraints (*Eq*. 18.a) and (*Eq*. 18.c) of the data collection optimization model when the static sink node stays at each position can meet the optimization model (*Eq*. 17), and $\sum_p \tilde{e}_i^p \leq \sum_p \tilde{q}_p E_i$.

Moreover, the optimal solution combination of the data collection optimization models (*Eq*. 18) when static sink node stays at each position can meet all constraints of the optimization model (*Eq*. 17). The combination solution constitutes a solution of the model (*Eq*. 17), and $\hat{q}$ is the minimum value. Therefore, $\hat{q} \leq \sum_p \tilde{q}_p$. Similarly, the optimal solution $\hat{F}$ of the optimization model (*Eq*. 17) is divided according to the positions, and it meets constraints (*Eq*. 18.a) and (*Eq*. 18.c) of the optimization model (*Eq*. 18). If at least one sensor node exists, and its energy is exhausted, then $\sum_p \hat{e}_i^p = \hat{q}E_i$. Let $\hat{q}_p = \hat{e}_i^p / E_i$, then $\sum_p \hat{q}_p = \hat{q}$. As the minimum value of the objective function is $\tilde{q}_p$ at position p, then $\hat{q}_p = \hat{e}_i^p / E_i \geq \tilde{q}_p$ and $\hat{q} \geq \sum_p \tilde{q}_p$. Therefore, $\hat{q} = \sum_p \tilde{q}_p$ is obtained.

In summary, when the artificial intelligence algorithm determines the food source of the bee according to the position of the sink node on this food source, it can solve the optimization model (*Eq*. 18) of the sink node, and obtain the current communication scheme [21]. After determining the communication schemes at all positions, we can obtain the optimal data collection scheme with the known sink node's path, and calculate the fitness value of the food source by *Eq*. (16).

B. MODIFIED ARTIFICIAL BEE COLONY ALGORITHM

In DCA_3D, the key aspect in solving the optimization model (*Eq*.14) by the modified artificial bee colony algorithm is knowing how to find the moving path of the sink node. The solving process includes the initialization of the food source and the employed bee phase, onlooker bee phase, and scout bee phase. The specific principle is discussed subsequently.

1) INITIALIZATION OF FOOD SOURCE

The main purpose of the initialization phase is to determine the initial scheme of the food sources, that is, the initial position of the sink node randomly assigned to a grid's center. According to the current position of the sink node, the sink nodes selects all centers of the neighbor grid never before visited, and randomly chooses one of them as the grid's center at which the sink node stays at the next moment. If there is no center of any nearby neighbor grids which is not visited, then the sink nodes chooses the nearest grid's center as the grid's center at the next moment. If the moving path length of the sink node reaches the specified threshold, then DCA_3D completes the moving path searching of the sink node and the food source initialization.

2) EMPLOYED BEES

In the employed bee phase, DCA_3D hires bees to optimize the path of the current food source S. For each sink node residence position, DCA_3D selects the neighbor grid's center of the current residence position. Then it randomly selects a quarter circle, and similarly selects the neighbor grid's center in the quarter circle. If the selected grid's center is on the path of the food source, then it exchanges the selected and next grid's centers directly; otherwise, it replaces the next grid's center directly by the selected grid's center. Then DCA_3D obtains a new food source S_i' after some cycles. If the fitness

value of the new food source S_i' is greater than that of S_i and larger than a specified threshold value, then the new food source S_i' is regarded a better moving path, and it replaces the food source S_i; otherwise, the path search fails, and the failure parameter β_i of the food source S_i adds 1.

3) ONLOOKER BEES

DCA_3D calculates the updating probability of the optimal fitness value after optimizing all the food sources.

$$P_i = f_i / \sum_j f_j \tag{19}$$

where f_i denotes the fitness value of food source i, and P_i denotes the updating probability of food source i. It generates a random number. If the random number is less than P_i, then it enters the onlooker bee phase. In this phase, DCA_3D randomly adds the non-staying grid's center to the moving path. Then it optimizes each path according to the shortest path selection and adjusts the order of the residence position appropriately to obtain a new moving path with a fixed length. If the fitness value of the new moving path is greater than that of the current moving path, then it replaces the current moving path directly.

4) SCOUT BEES

Following the employed bee phase and the onlooker bee phase, if the failure parameter β_i of the food source i is greater than a specified threshold, then the food source falls into a local optimal solution after several iterations, and DCA_3D reinitializes the food source. However, it cannot discard the optimal food source until it identifies a better food source.

C. ALGORITHM IMPLEMENTATION

As shown in Fig.3, the specific implementation steps of the DCA_3D are as follows:

Step 1: Initializes the parameters of the algorithm by setting the initial values of iteration number $m_1 = 1$, current food source ID $m_2 = 1$, maximum iteration number M_1, total number M_2 of food sources, etc.

Step 2: Initializes the moving paths of the food sources.

Step 3: Calculates the fitness value of food source m_2. If the fitness value is larger than the historical optimal value, then it updates the optimal food source, and $m_2 = m_2 + 1$. If m_2 is greater than M_1, it obtains the fitness value of each food source, global optimal food source, and its fitness value, and $m_2 = 1$, then proceed to step 4; otherwise, repeat step 3.

Step 4: Enter the employed bee phase and modify the moving paths of all food sources; that is, modify the current food source m_2 according to the information of the neighbor grid's center to obtain a new food source. If the fitness value of the new food source is greater than that of the current food source, then it replaces the current one ; otherwise, $\beta_i = \beta_i + 1$, and $m_2 = m_2 + 1$. If m_2 is greater than M_2, it completes the employed bee phase and calculates the updating probability P_i of the optimal fitness value. Then, $m_2 = 1$, and proceed to step 5; otherwise, repeat step 4.

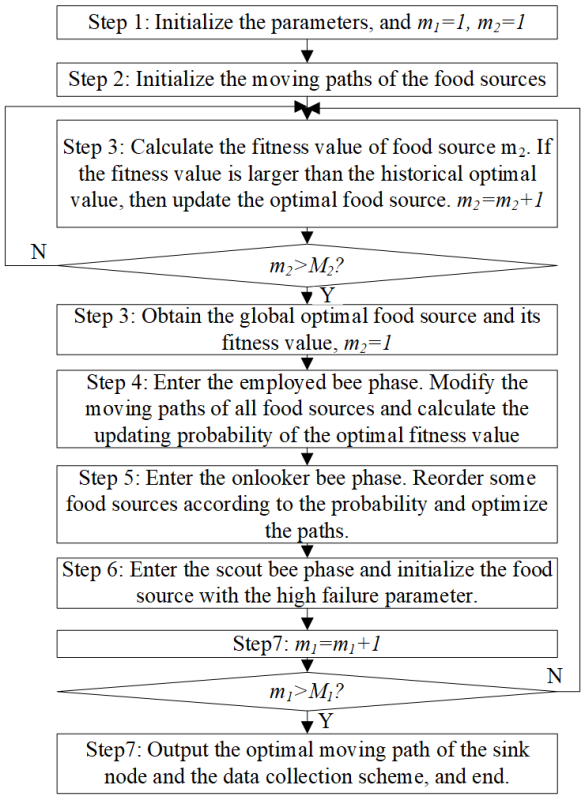


FIGURE 3. Flowchart of the DCA_3D.

Step 5: Enter the onlooker bee phase. Reorder some food sources according to the probability, and optimize the paths. Then it subsequently generates a random number R. If R is less than P_i, then the uncovered neighbor grid's centers in food source m_2 are added. The paths are reordered according to the shortest path principle. A new food source is obtained. If the fitness value of the new food source is greater than that of current food source, then the current food source is replaced; otherwise, $m_2 = m_2 + 1$. If m_2 is greater than M_2, the onlooker bee phase is completed, and $m_2 = 1$, then proceed to step 6; otherwise, repeat step 5.

Step 6: Enter the scout bee phase and initialize the food source with the high failure parameter. If the failure parameter β_i of food source m_2 is larger than the threshold value, then it reinitializes the food source, and $m_2 = m_2+1$. If m_2 is greater than M_2, it completes the scout bee phase, and $m_2 = 1$, then proceed to step 7; otherwise, repeat step 6.

Step 7: $m_1 = m_1 + 1$. If m_2 is larger than M_2, then it outputs the optimal moving path of the sink node and the data collection scheme, and ends the algorithm ; otherwise, repeat step 3.

DCA_3D executes key steps 3–7 cyclically, and obtains the optimal solution of the optimization model (*Eq.*14). Then it obtains the optimal moving path of the sink node and the data collection scheme which can improve network lifetime.

TABLE 2. Simulation Parameters.

Parameter name	Value
Side length of 3D square simulation area	1000 m
Grid length	200 m
Node communication radius	200 m
Electronic energy consumption ε_{fs} of signal amplifier when amplifying unit bit signal	100 pJ/bit/ m^2
Electronic energy consumption E_{elec} of the circuit when receiving and transmitting unit bit data wirelessly	50 nJ/bit
Maximum number of iterations M_1	50
Number of food sources M_2	50
Failure parameter β_i	5
Stay time at each grid	100s
Maximum data storage capacity of sensor node	1 Mbit
Sensing rate of sensor node	10 bit/s
Initial energy	1000J

V. ALGORITHM SIMULATION

A. SELECTION OF SIMULATION PARAMETERS

The monitoring area is divided into several grids, and the mobile sink node moves among the grids during algorithm simulation. If the mobile sink node moves to a grid's center, then the sensor nodes not covered by the sink node store the sensing data in the storage space; otherwise, the sensor nodes transmit the sensing data to the sink node. We select the parameters in Table 2 for the simulation experiments to obtain the performance parameters, such as coverage rate of sensor nodes, network lifetime, average data transmission amount of sensor nodes, average energy consumption variance of sensor nodes, and average packet loss rate of sensor nodes. Network lifetime is calculated according to *Eq.* (13). The coverage rate of the sensor nodes is defined as the ratio between the number of static sensor nodes, which can communicate with sink nodes when a sink node moves along the path, and the total number of sensor nodes. The average data transmission amount of the sensor nodes R is defined as the average sensing data amount that is successfully sent to sink node.

$$R = \sum_{g} \sum_{i} \tilde{D}_{is}^{g} / N \tag{20}$$

where $\tilde{D}_{is}^{g}$ denotes the amount of data that the sink node receives from sensor node i when the sink node stays at position g. The average energy consumption variance of the sensor nodes is defined as the variance of the energy consumption of the sensor nodes when the sink node moves along the path once. The sensor node loses its own data because of its limited data storage capacity. The update formula of the number of lost packets is defined as

$$D_i^{p+1} = \begin{cases} C_i^p + S_i t_p - C_{\max}, C_i^p + S_i t_p > C_{\max} \\ 0, others \end{cases} \quad \forall i \in V \tag{21}$$

where D_i^p denotes the number of lost packets of sensor node i when the sink node stays at position p. The average packet loss rate of the sensor nodes is defined as the average packet loss rate of the sensor nodes when the sink node moves along its path once, that is,

$$D_{ave} = \sum_{i} D_i^p / D_{total} \tag{22}$$

where D_{total} denotes the total number of data packets generated by the sensor nodes when the sink node moves along its path once.

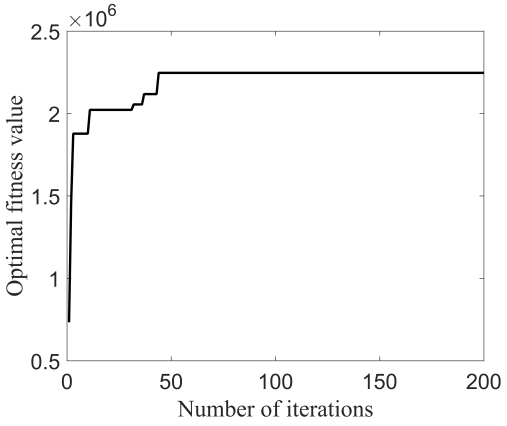


FIGURE 4. Convergence diagram of DCA_3D.

B. ANALYSIS OF SIMULATION RESULTS

First, we analyze the convergence of DCA_3D. Then we select the number of static sensor nodes 100, the maximum data collection hops of the sink node 3, and the parameters in Table 1. DCA_3D loops a total of 200 times to calculate the optimal fitness value of the current food source, then it obtains the convergence diagram. As shown in Fig. 4, the optimal fitness value of each iteration continues to increase after several iterations in the early stage. The optimal fitness value of each iteration converges to $2.247 * 10^6$ after 50 iterations. Therefore, DCA_3D is convergent, and can find the optimal solution of the optimization model consequently.

We select the parameters in Table 1, and cyclically calculate coverage rate, network lifetime, average data transmission amount, average energy consumption variance, and average packet loss rate in RAND [10], GREED [10], EDG_3D [20], ANT [26], and DCA_3D. Where RAND, GREED, EDG_3D and ANT are classical algorithms in path selection. RAND randomly selects the unpassed neighbor grids as sink node's next position in 3D environment. GREED randomly selects one of the closest grid in the sets that have not passed as sink node's next position in 3D environment. EDG_3D selects the linear movement path to traverse all grids according to the current number of grids in 3D environment. ANT uses ant colony algorithm to solve the model (*Eq.*14) and obtains the moving path of sink node. The next grid selection method of the mobile sink node is inconsistent in RAND, GREED, EDG_3D, and ANT, but the data collection method of the sink node staying at a grid's center is the same. We take the maximum data collection hops of 3 for the sink node and the number of 100 for the sensor nodes as examples, and compare the moving paths of

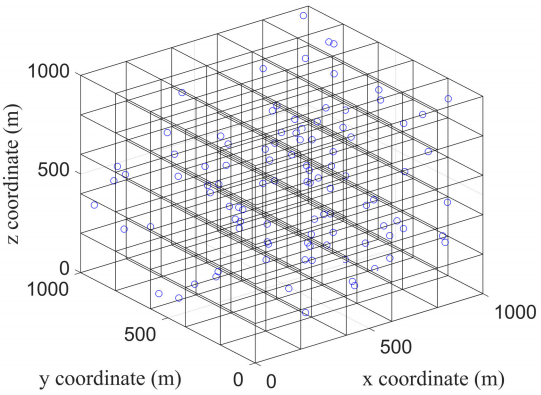


FIGURE 5. Grid division.

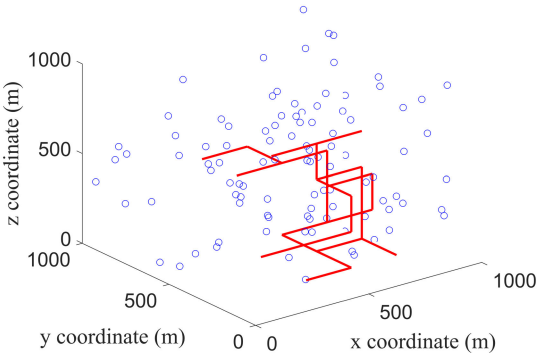


FIGURE 6. Date collection in RAND.

those algorithms. The monitoring area shown in Fig. 5 is divided into 5*5*5 monitoring centers, and all sensor nodes are randomly distributed in this area. As shown in Fig. 6, the sink node randomly selects the surrounding grid as the next residence grid in RAND, which means that the moving path of the sink node is random and blind, and the sink node only moves in a sub-area of the monitoring area. As shown in Fig. 7, the sink node randomly selects a non-repetitive grid in the nearest grid set as the next residence grid in GREED, so the moving path covers a wide range. As shown in Fig. 8, according to the number of current 3D grids, EDG_3D traverses all grids as the path and selects the first part as the linear moving path of the sink node. However, RAND, GREED, and EDG_3D do not consider the position distribution of the sensor nodes. As shown in Fig. 9, ANT selects the ant colony algorithm to solve the optimization model (*Eq*. 14). Given that we select the same number of iterations, ANT only finds the local optimal path, and its moving path concentrates in the area with more sensor nodes, but it does not consider the other areas. Therefore, the moving path lengths of the above four algorithms are not optimal. As shown in Fig. 10, DCA_3D establishes a data collection optimization model that considers both the moving path of the sink node and the communication among sensor nodes, and uses the modified artificial bee colony algorithm to solve effectively the optimization model. Then it obtains the optimal scheme, which is the most suitable approach for the current distribution of the

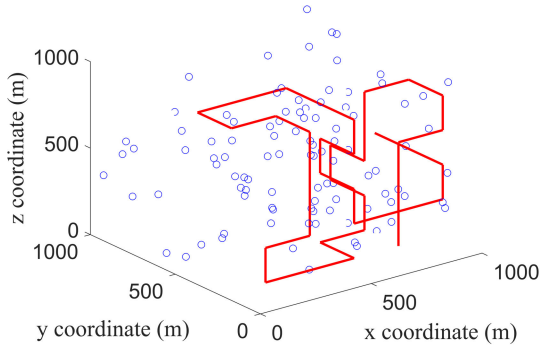


FIGURE 7. Date collection in GREEN.

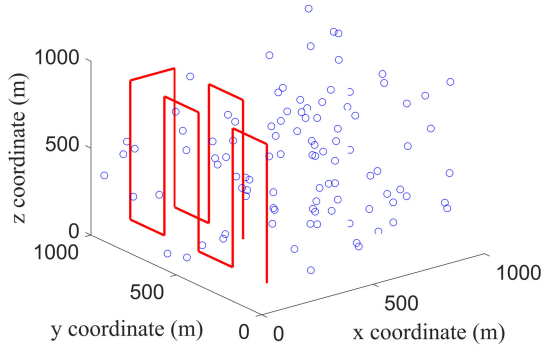


FIGURE 8. Date collection in EDG_3D.

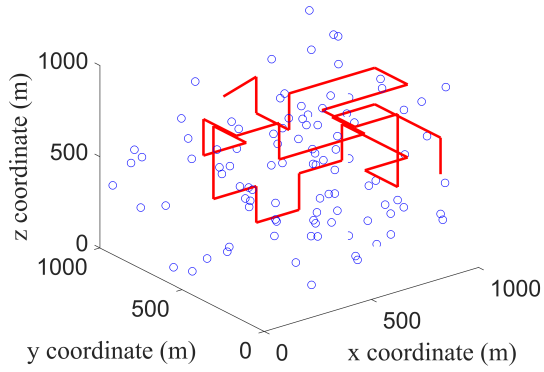


FIGURE 9. Date collection in ANT.

sensor nodes. Therefore, the moving path of the sink node in DCA_3D is distributed throughout the monitoring area, and it can cover the whole sensor nodes as much as possible under the constraint of limited moving path length.

1) SIMULATION RESULT ANALYSIS WHEN THE MOVING PATH LENGTH OF THE SINK NODE CHANGES

We select the moving path lengths of the sink node as 20, 30, 40, 50, 60, 70, 80, 90, and 100, the number of static sensor nodes 100, the maximum data collection hops of the sink node 3, and the parameters in Table 1. Then we randomly generate ten different topological structures of sensor nodes, and calculate the coverage rate, network lifetime, average data transmission amount, average energy consumption variance, and average packet loss rate in RAND, GREED, EDG_3D,

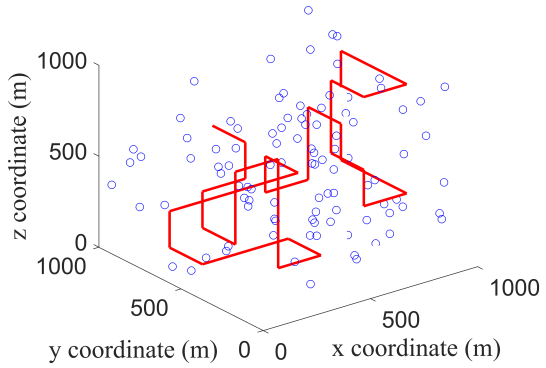


FIGURE 10. Date collection in DCA_3D.

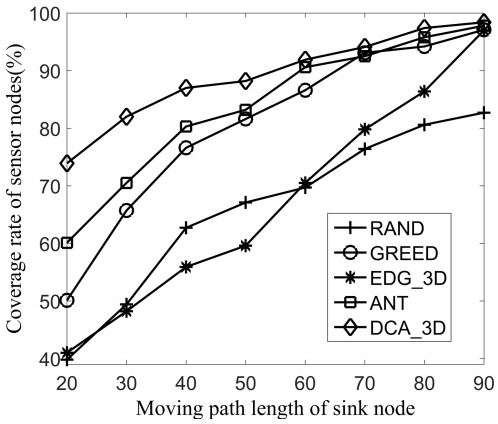


FIGURE 11. Coverage rate comparison of sensor nodes.

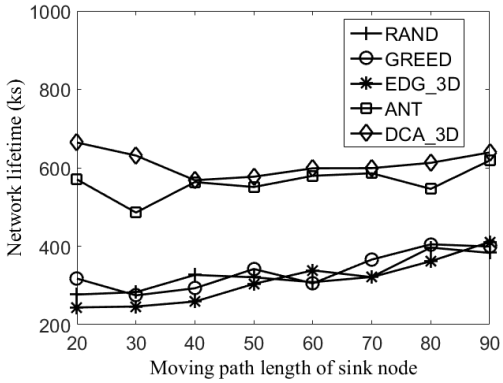


FIGURE 12. Network lifetime comparison.

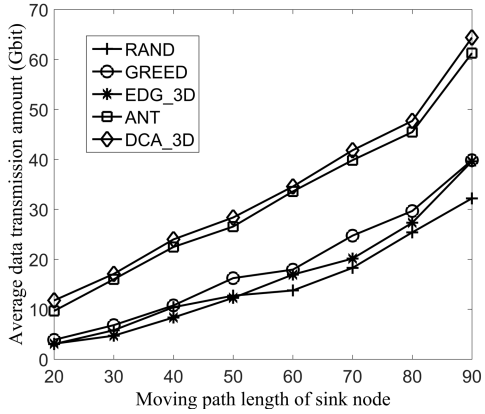


FIGURE 13. Average data transmission amount comparison of sensor nodes.

ANT, and DCA_3D under each topological structure. The average values are the simulation results. The obtained simulation results are presented hereafter.

As shown in Fig. 11, when the moving path length of the sink node increases, the coverage rates of the sensor nodes increase in RAND, GREED, EDG_3D, ANT, and DCA_3D, but the coverage rate of the sensor nodes in DCA_3D is higher than those in RAND, GREED, EDG_3D, and ANT. This result can be explained as follows. When the moving path length of the sink node increases, the range of coverage of the sink node increases, and more opportunities are provided for the sensor nodes to communicate with the sink node in the monitoring area, resulting in an increase in the coverage rate of the sensor nodes. In RAND, the mobile sink node selects a residence grid's center as the next position from neighborhood grid's center. The selection is blind and random, indicating that the coverage rate of the sensor nodes is low. EDG_3D searches for a linear moving path that traverses the centers of all grids and starts by selecting subpaths with fixed lengths from the beginning. The moving path of the sink node is located in the corner of the monitoring area, and the coverage rate of the sensor nodes is low when the moving path length of the sink node is short. When the moving path length of the sink node increases, the moving path of the sink node can pass through more than half of the grids, and the coverage rate of the sensor nodes increases significantly. GREED randomly selects an non-staying grid's center from the nearest neighbor grid's centers as the next residence grid. The above three algorithms do not attempt to find the optimal moving path, so their coverage rates of sensor nodes are low. DCA_3D regards the coverage rate of the sensor nodes as an important parameter of the path fitness value in bee foraging, which affects the search of food sources in the employed bee, onlooker bee, and scout bee phases, then it can find the optimal moving path of the sink node, which improves the coverage rate of the sensor nodes after many iterations.

As shown in Fig. 12, the network lifetime of DCA_3D is longer than those of RAND, GREED, EDG_3D, and ANT. This result can be attributed to the characteristic of DCA_3D that regards network lifetime as one of the important parameters in the fitness function, and it moves to the grid with more sensor nodes in the process of solving the food source. This scheme reduces the energy consumption of those nodes and subsequently improves network lifetime. RAND, GREED, and EDG_3D do not consider network lifetime when selecting the moving path, resulting in relatively short network lifetimes for the three algorithms. ANT considers network lifetime, but it only finds the local optimal solution when the same short iteration times is selected, resulting in a slightly lower network lifetime.

As shown in Fig. 13, when the moving path length of the sink node increases, the average data transmission amounts

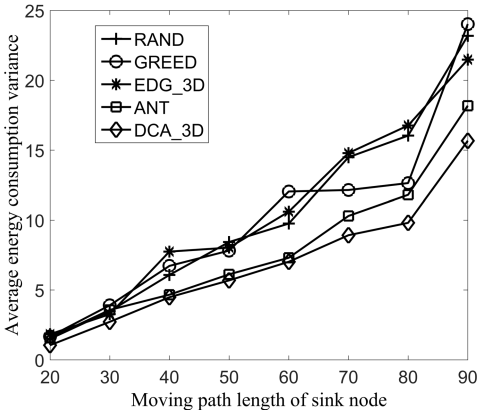


FIGURE 14. Average energy consumption variance comparison of sensor nodes.

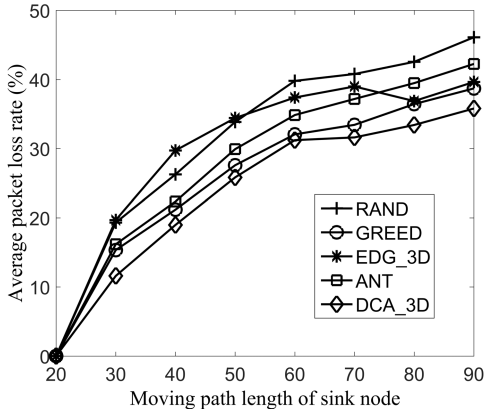


FIGURE 15. Average packet loss rate comparison of sensor nodes.

of the sensor nodes in RAND, GREED, EDG_3D, ANT, and DCA_3D increase. However, the average data transmission amount in DCA_3D is higher than those in RAND, GREED, EDG_3D, and ANT. This result can be explained as follows. When the moving path length of the sink node increases, the coverage area of the sink node's data collection increases, and more data of the sensor nodes can be collected, resulting in a relatively large average data transmission amount of the sensor node. However, DCA_3D finds a much better moving path. Then it can increase the coverage rate of the sensor nodes, the communication time between the sensor nodes and sink node, and the data amount, which is successfully sent by all sensor nodes, resulting in the largest average data transmission amount of the sensor nodes. The moving path of ANT is only locally optimal, so the average data transmission amount in ANT is second in the ranking. RAND, GREED, and EDG_3D do not consider the position distribution of sensor nodes when selecting the path, so the average data transfer amounts of those three algorithms are low.

As shown in Fig. 14, when the moving path length of the sink node increases, the average energy consumption variances in RAND, GREED, EDG_3D, ANT, and DCA_3D increase. However, the average energy consumption variance in DCA_3D is lower than those in RAND, GREED, EDG_3D, and ANT. This result can be explained as follows. When the moving path length of the sink node increases, the communication opportunity between some sensor nodes and sink node increases further. Some sensor nodes transmit data directly and are used as a relay node to transmit data from the other sensor nodes to the sink node. As a result, the energy consumption of the sensor nodes is unbalanced, and the average energy consumption variance increases. However, DCA_3D establishes the data collection optimization model of the 3D WSN. The optimal moving path of the sink node is obtained using the modified artificial bee colony algorithm to solve the model. The energy consumption of the sensor nodes is balanced, and the average node energy consumption variance is reduced by ensuring that the sensor nodes are as close as possible to the sink node.

As shown in Fig. 15, when the moving path length of the sink node increases, the average packet loss rates of one round of data collection in RAND, GREED, EDG_3D, ANT, and DCA_3D increase. However, the average packet loss rate of one round of data collection in DCA_3D is lower than those in RAND, GREED, EDG_3D, and ANT. This result can be explained as follows. When the moving path length of the sink node increases, the time of the sink nodes moving along its path becomes longer, and the communication distance between some sensor nodes and sink node becomes larger. This scenario increases the data amount of sensor nodes. Owing to the limited storage space of the sensor nodes, the average packet loss rate increases. At the same time, DCA_3D finds the optimal moving path of the sink node, indicating that more sensor nodes can communicate with the sink node to transmit their own data, and the corresponding average packet loss rate is the lowest. In order to highlight the comparison effect of the algorithms during the simulation, we set less sensor node's storage space and maximum data collection hops, which leads to higher average packet loss rates of the five algorithms when the moving path length of the sink node is large. In practical applications, we can adjust the sensor node's storage space and the maximum data collection hops to reduce the average packet loss rate.

Thus, regardless of how the moving path length of the sink node changes, DCA_3D can find the optimal moving path of a sink node, which improves the coverage rate of sensor nodes, network lifetime, and average data transmission amount, and reduces average node energy consumption variance and average packet loss rate.

2) SIMULATION RESULT ANALYSIS WHEN THE MAXIMUM DATA COLLECTION HOPS OF THE SINK NODE CHANGE

We select the maximum data collection hops of the sink node as 1, 2, 3, 4, 5, 6, and 7, the number of the sensor node as 100, the moving path length of the sink node as 30 and the parameters in Table 1. Then, we randomly generate ten different topological structures of the sensor nodes and calculate the coverage rate, network lifetime, average data transmission amount, average energy consumption variance,

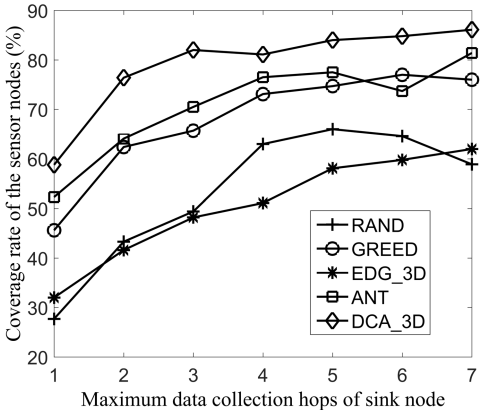


FIGURE 16. Coverage rate comparison of the sensor nodes when the maximum data collection hops of the sink node change.

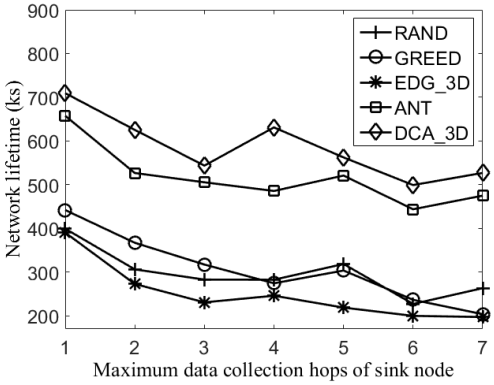


FIGURE 17. Network lifetime comparison when maximum data collection hops of the sink node change.

and average packet loss rate under each topology structure. The average values are the simulation results in RAND, GREED, EDG_3D, ANT, and DCA_3D. The specific simulation results are presented hereafter.

As shown in Fig.16, because the moving path length of the sink node is less, the sink node only passes through a part of the monitoring area. Then the coverage rate of the sensor nodes in the five algorithms is not larger than 90%. But the coverage rate of the sensor nodes in DCA_3D is higher than those in RAND, GREED, EDG_3D, and ANT. This result can be explained as follows. DCA_3D establishes an optimization model to weigh the network lifetime and coverage rate. And it uses the modified artificial bee colony algorithm to solve the optimization model, which takes the coverage rate of the sensor nodes as the main parameter, then both obtains the optimal value of the fitness function and the optimal moving path. This scheme improves the coverage rate of the sensor nodes. RAND, GREED, and EDG_3D do not consider the distribution of the sensor nodes, whereas ANT only finds the local optimal solution. As such, the coverage rates of the sensor nodes for the above four algorithms are low.

As shown in Fig. 17, the network lifetime of DCA_3D is longer than those of RAND, GREED, EDG_3D, and ANT. This result can be explained as follows. DCA_3D regards network lifetime as one of the important parameters of the fitness function and moves to the grid with more sensor nodes in the solving process of modified artificial bee colony. Moreover, DCA_3D reduces the energy consumption of the nodes and improves network lifetime. RAND, GREED, and EDG_3D do not consider network lifetime when selecting the moving path, whereas ANT only finds the local optimal solution. Thus, the network lifetimes in the above four algorithms are relatively short. The simulation results and specific reasons of the average node data transmission amount, average node energy consumption variance, and average packet loss rate of each algorithm when the maximum data collection hops of the sink node change are consistent with the simulation results and specific reasons when the moving path length of the sink node changes,as previously discussed in Section V.B.1.

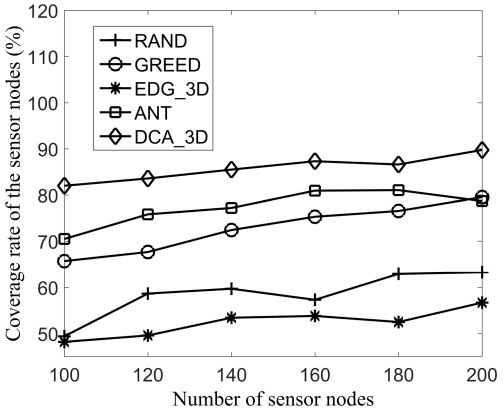


FIGURE 18. Coverage rate comparison of the sensor nodes when the number of the sensor nodes changes.

3) SIMULATION RESULT ANALYSIS WHEN THE NUMBER OF THE SENSOR NODES CHANGES

We select the maximum data collection hops of the sink nodes 3, the numbers of the sensor nodes as 100, 120, 140, 160, 180, and 200, the moving path length of the sink node as 30 and the parameters in Table 1. Then,we randomly generate ten different topological structures of the sensor nodes, and calculate the coverage rate, network lifetime, average data transmission amount, average energy consumption variance, and average packet loss rate under each topology structure. The average values are the simulation results in RAND, GREED, EDG_3D, ANT, and DCA_3D. The specific simulation results are presented hereafter.

As shown in Figs. 18 and 19, the coverage rate of the sensor nodes in DCA_3D is notably higher than those in RAND, GREED, EDG_3D, and ANT. The network lifetime of DCA_3D is longer than those in RAND, GREED, EDG_3D, and ANT. The coverage rate and network lifetime of the sensor nodes in DCA_3D are not affected by the number of static sensor nodes, that is, regardless of the number of the sensor nodes, DCA_3D can find the optimal moving path of the sink node by solving the optimization model (*Eq*. 14), which can balance and improve the coverage rate of the sensor nodes and network lifetime. By contrast, the other algorithms

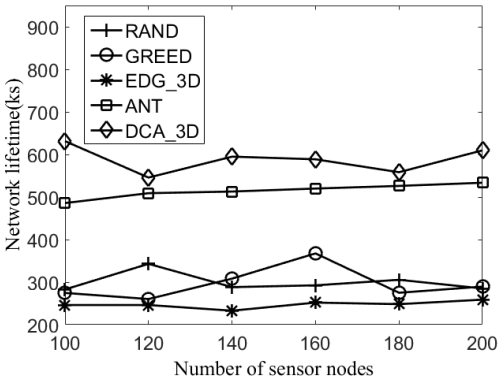


FIGURE 19. Network lifetime comparison when number of the sensor nodes changes.

present certain limitations. Consequently, DCA_3D has the highest coverage rate of the sensor nodes and network lifetime. The simulation results and specific reasons of average node data transmission amount, average node energy consumption variance, and average packet loss rate of each algorithm when the number of the sensor nodes changes are consistent with the simulation results and specific reasons when the moving path length of the sink node changes, as previously discussed in Section V.B.1.

VI. CONCLUSION

We propose a data collection algorithm of the 3D WSN (DCA_3D) that weighs node coverage rate and lifetime. First, we propose the algorithm assumptions, the cubic grid division method, and the basic principle in 3D environments. Second, we establish the data collection optimization model that weighs node coverage rate and lifetime with the constraints, such as the sink node's moving path selection constraints, data flow constraint, energy consumption constraint and link transmission constraint. Then, we propose the fitness value calculation method of the sink node's moving path to solve the data collection optimization problem with the known moving path of the sink node, and use the modified artificial bee colony algorithm to solve the moving path problem of the sink node and subsequently obtain the optimal moving path. Finally, we give the simulation parameters of the algorithm, and compare the performances of DCA_3D, RAND, GREED, EDG_3D, and ANT.

In summary, regardless of the moving path length of the sink node, the maximum data collection hops of the sink node and the number of the sensor nodes change. DCA_3D can find the optimal moving path of the sink node, which improves the coverage rate, network lifetime, and average data transmission amount of the sensor nodes, and reduces the average energy consumption variance and average packet loss rate of the sensor nodes. However, the time complexity of the modified artificial bee colony algorithm in DCA_3D is relatively high. Therefore, the succeeding stage of this research will study the heuristic algorithm to determine the optimal moving path of the sink node and the communication scheme of the sensor nodes, according to the information of surrounding sensor nodes through the communication between the sink node and sensor nodes, which can reduce the algorithm's time complexity.

REFERENCES

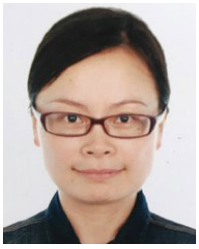
[1] M. Abdulkarem, K. Samsudin, and Z. Rokhani, "Wireless sensor network for structural health monitoring: A contemporary review of technologies, challenges, and future direction," *Nature*, vol. 19, pp. 693–735, Feb. 2020.
[2] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828–854, 2nd Quart., 2017.
[3] A. Mehto, S. Tapaswi, and K. K. Pattanaik, "Virtual grid-based rendezvous point and sojourn location selection for energy and delay efficient data acquisition in wireless sensor networks with mobile sink," *Wireless Netw.*, vol. 26, no. 5, pp. 3763–3779, Jul. 2020.
[4] X. Zhao, X. Xiong, Z. Sun, X. Zhang, and Z. Sun, "An immune clone selection based power control strategy for alleviating energy hole problems in wireless sensor networks," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 6, pp. 2505–2518, Jun. 2020.
[5] N. Sharmin, A. Karmaker, W. L. Lambert, M. S. Alam, and M. S. A. Shawkat, "Minimizing the energy hole problem in wireless sensor networks: A wedge merging approach," *Sensors*, vol. 20, no. 1, pp. 277–301, 2020.
[6] H. U. Yildiz, V. C. Gungor, and B. Tavli, "Packet size optimization for lifetime maximization in underwater acoustic sensor networks," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 719–729, Feb. 2019.
[7] Y. Wang and S. Zhang, "Research on network lifetime optimization based on transmission power and data packet size," *Mod. Electron. Techn.*, vol. 41, no. 5, pp. 41–46, 2018.
[8] J. Wang, Y. Gao, C. Zhou, R. Simon Sherratt, and L. Wang, "Optimal coverage multi-path scheduling scheme with multiple mobile sinks for WSNs," *Comput., Mater. Continua*, vol. 62, no. 2, pp. 695–711, 2020.
[9] A. Karimi and S. M. Amini, "Reduction of energy consumption in wireless sensor networks based on predictable routes for multi-mobile sink," *J. Supercomput.*, vol. 75, no. 11, pp. 7290–7313, Nov. 2019.
[10] C. Zhu, S. Zhang, G. Han, J. Jiang, and J. Rodrigues, "A greedy scanning data collection strategy for large-scale wireless sensor networks with a mobile sink," *Sensors*, vol. 16, no. 9, pp. 1432–1460, 2016.
[11] A. A. A. Ari, I. Damakoa, A. Gueroui, C. Titouna, N. Labraoui, G. Kaladzavi, and B. O. Yenké, "Bacterial foraging optimization scheme for mobile sensing in wireless sensor networks," *Int. J. Wireless Inf. Netw.*, vol. 24, no. 3, pp. 254–267, Sep. 2017.
[12] J. Zhong, Z. Huang, L. Feng, W. Du, and Y. Li, "A hyper-heuristic framework for lifetime maximization in wireless sensor networks with a mobile sink," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 1, pp. 223–236, Jan. 2020.
[13] M. E. Keskin, "A column generation heuristic for optimal wireless sensor network design with mobile sinks," *Eur. J. Oper. Res.*, vol. 260, no. 1, pp. 291–304, Jul. 2017.
[14] F. Tashtarian, K. Sohraby, and A. Varasteh, "Multihop data gathering in wireless sensor networks with a mobile sink," *Int. J. Commun. Syst.*, vol. 30, no. 12, p. e3264, 2017.
[15] L. Lu, J. Wang, C. Zong, and P. Zhao, "Simulation of 3D path planning approach for quad-rotor helicopter based on A~* algorithm," *J. Hefei Univ. Technol. (Natural Sci.)*, vol. 40, no. 3, pp. 304–309, 2017.
[16] G. Frontera, D. J. Martín, J. A. Besada, and D.-W. Gu, "Approximate 3D Euclidean shortest paths for unmanned aircraft in urban environments," *J. Intell. Robotic Syst.*, vol. 85, no. 2, pp. 353–368, Feb. 2017.
[17] N. T. Tam, D. T. Hai, L. H. Son, and L. T. Vinh, "Improving lifetime and network connections of 3D wireless sensor networks based on fuzzy clustering and particle swarm optimization," *Wireless Netw.*, vol. 2, no. 2, pp. 1477–1490, 2016.
[18] H. A. S. M. Sanwar, C. G. Hwan, and R. In-Ho, "An eccentricity based data routing protocol with uniform node distribution in 3D WSN," *Sensors*, vol. 17, no. 9, p. 2137, 2017.
[19] A. S. M. S. Hosen and G. H. Cho, "An eccentricity-based data routing protocol for 3D wireless sensor networks," *Int. J. Sensor Netw.*, vol. 24, no. 4, pp. 230–240, 2017.
[20] A. Mariam, J. Nadeem, K. Ayesha, I. Muhammad, S. Muhammad, and V. Athanasios, "Efficient data gathering in 3D linear underwater wireless sensor networks using sink mobility," *Sensors*, vol. 16, no. 3, p. 404, 2016.

[21] Y. Chen, Z. Wang, T. Ren, and H. Lv, "Lifetime optimization algorithm with mobile sink nodes for wireless sensor networks based on location information," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 8, 2015, Art. no. 857673.
[22] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Appl. Soft Comput.*, vol. 8, no. 1, pp. 687–697, 2008.
[23] G. Tian, Y. Ren, Y. Feng, M. C. Zhou, H. Zhang, and J. Tan, "Modeling and planning for dual-objective selective disassembly using and/or graph and discrete artificial bee colony," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2456–2468, Apr. 2018.
[24] H. Jia, H. Miao, G. Tian, M. Zhou, Y. Feng, Z. Li, and J. Li, "Multiobjective bike repositioning in bike-sharing systems via a modified artificial bee colony algorithm," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 2, pp. 909–920, Apr. 2020.
[25] H. Liu, M. C. Zhou, X. Guo, Z. Zhang, B. Ning, and T. Tang, "Timetable optimization for regenerative energy utilization in subway systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 9, pp. 3247–3257, Sep. 2019.
[26] N. Ghosh, I. Banerjee, and R. S. Sherratt, "On-demand fuzzy clustering and ant-colony optimisation based mobile data collection in wireless sensor network," *Wireless Netw.*, vol. 25, no. 4, pp. 1829–1845, May 2019.

BANTENG LIU received the B.S. and M.S. degrees in communication engineering from the Zhejiang University of Technology, Hangzhou, China, in 2009, and the Ph.D. degree in control theory and control engineering from Zhejiang University, Hangzhou, China, in 2017. Since 2009, he has been teaching in the School of Information Technology, Zhejiang Shuren University, where he is currently an Associate Professor. His research interests include wireless mobile communication and nondestructive flaw detection.

YOURONG CHEN received the B.S. degree in communication engineering, the M.S. degree in communication engineering communication and information system, and the Ph.D. degree in control theory and control engineering from the Zhejiang University of Technology, Hangzhou, China, in 2004, 2007, and 2012, respectively. Since 2007, he has been teaching in the School of Information Technology, Zhejiang Shuren University, where he is currently a Professor. His research interests include the Internet of Things engineering and wireless mobile communication.

JINHAO WAN received the B.S. degree in communication engineering from Zhejiang Shuren University, Hangzhou, China, in 2016, and the M.S. degree in computer system architecture from Changzhou University, Changzhou, China, in 2019. He is currently a Teaching Assistant with Zhejiang Shuren University. His research interest includes wireless and mobile communications.



PING SUN received the B.S. degree from Hangzhou Normal University, in 2001, and the M.S. degree in computer technology from Zhejiang University, in 2010. Since 2004, she has been teaching in the School of Information Technology, Zhejiang Shuren University, where she is currently an Associate Professor of computer science and technology. Her current research interests include the Internet of Things engineering and micro-curriculum design.

ZHANGQUAN WANG received the B.S. degree from the Department of Technical Teachers, Zhejiang Institute of Technology, Hangzhou, China, in 1991, and the M.S. degree in electrical engineering from Zhejiang University, Hangzhou, in 2004. Since 2001, he has been teaching in the School of Information Technology, Zhejiang Shuren University, where he is currently a Professor. His research interest includes wireless and mobile communications.

TIAOJUAN REN received the B.S. degree in dynamic testing and automation from Zhejiang University, Hangzhou, China, in 1987, and the M.S. degree in information management from Korea Giant Buddha University, South Korea, in 2005. Since 2001, she has been teaching in the School of Information Technology, Zhejiang Shuren University, where she is currently a Professor. Her research interest includes wireless and mobile communications.

ZEGAO YIN received the B.S. degree in technology and economy and the M.S. degree in hydraulics and river dynamics from the North China University of Water Conservancy and Electric Power, in 1999 and 2002, respectively, and the Ph.D. degree in civil engineering from Zhejiang University, in 2005. Since 2005, he has been teaching in the College of Engineering, Ocean University of China, where he is currently a Professor. His research interests include hydrodynamics for coastal and offshore engineering, and the utilization of marine renewable energy.

RENGONG ZHANG received the B.S. and M.S. degrees from the North China University of Water Conservancy and Electric Power, in 1998 and 2002, respectively, and the Ph.D. degree from the Zhejiang University of Technology, in 2014. Since 2014, he has been working with Zhejiang Yugong Information Technology Company Ltd. His research interests include water conservancy and hydropower engineering, and the utilization of marine renewable energy.