Healing Coverage Holes for Big Data Collection in Large-Scale Wireless Sensor Networks

Jie Feng[1] · Hongbin Chen[1]

Abstract

The quality of service is severely degraded by coverage holes in wireless sensor networks. This paper focuses on the coverage hole healing (CHH) problem for big data collection in a large-scale wireless sensor network (LS-WSN) where the LS-WSN containing both static sensors and mobile sensors with the topology control of LEACH algorithm. Meanwhile, the data volume transmitted by each sensor node may be inconsistent. Specifically, the target of the CHH problem is to find an optimal subset of mobile nodes from all mobile nodes while maximizing the transmission times (TT) that all dispatched mobile nodes can transmit in their lifetime. Hence, from the data-centric perspective, we propose a greedy healing algorithm (GHA) via the greedy-based heuristic strategy with low computational complexity to solve this CHH problem. Simulation results show that the proposed GHA can efficiently heal the coverage holes which significantly prolongs the network lifetime and observably enhances the quality of service (QoS) of WSNs while increasing the TT, transmitted data volume (TDV) and average residual energy of all dispatched mobile nodes.

Keywords Coverage holes · Large-scale wireless sensor networks · Big data · Coverage hole healing · Greedy healing algorithm

1 Introduction

Wireless sensor networks, which consist of many tiny sensor nodes, have lots of applications such as battlefield surveillance, environmental monitoring [1], industrial diagnostics [2], integrated space and terrestrial Things. A concept has been proposed in [3] which refers to the second economy has been increased by big data on processors, sensors, connectors, and executors. It is estimated that by 2030, the size of the second economy will approach that of the current conventional physical economy. More and more WSNs become LS-WSNs and bring great economic benefits to the second economy. Nowadays, some early definitions about LS-WSN can be found in [4, 5]. Authors in [4] applied a LS-WSN for field surveillance which has been designed to scale up to 330 nodes. Nasipuri et al. [5] presented a LS-WSN using 122 nodes to monitor equipment. In addition, with the proliferation of wireless devices (e.g., mobile phones, wireless sensors, wireless smart meters, unmanned vehicles and drones) and wireless communication technologies (e.g., Internet of things [6, 7], WiFi, Bluetooth, Zigbee, RFID, cognitive radio [8–11], 4G, 5G), wireless networks become a global information infrastructure incurring a fast escalation of data volume known as big data [12]. Although the data generated by an individual sensor node in normal WSNs may not appear to be significant, the overall data generated across numerous sensor nodes in LS-WSNs can produce a significant portion of the big data, which can be called big data for WSNs actually. Therefore, the collection of big data in LS-WSNs is necessary and crucial. Next, some schemes for big data collection will be introduced.

With the emergence of LS-WSNs, it is highly challenging that the next generation of big data systems will need to deal with big data from these forms of sensor networks. Currently, many studies have concentrated on big data collection in LS-WSNs [13–19]. Takaishi et al. [13] proposed a new mobile sink routing and data gathering method through network clustering based on the modified expectation-maximization technique. Authors in [14] presented a structure fidelity data collection framework leveraging the spatial correlations between nodes to reduce the number of active sensor nodes while maintaining the low structural distortion of the collected data. In [15], Zhu et al. devised a four-phase mobility-assisted data collecting protocol consisting of network clustering, routes planning, routes combination, and data collecting. Moreover,

✉ Hongbin Chen
chbscut@guet.edu.cn

1 Key Laboratory of Cognitive Radio and Information Processing, Guilin University of Electronic Technology, Guilin 541004, China

two heuristic routes planning algorithms were presented to build a set of trajectories which satisfy the deadline constraint and have the minimum overall movement cost. For the risk analysis of industrial operation, Ding et al. [16] proposed a real-time big data gathering algorithm for an indoor WSN. In this algorithm, sensor nodes can screen the data collected from the environment and equipment according to the requirements of risk analysis. Authors in [17] presented an energy-efficient big data algorithm for a WSN for real-time data collection and established clustering communication on the basis of a received signal strength indicator and residual energy of sensor nodes. Din et al. [18] designed a novel technique by using a hybrid algorithm for clustering and cluster member selection in the wireless multi-sensor system. After the selection of cluster heads and member nodes, the proposed data fusion technique was used for partitioning and processing the data. Ang et al. [19] presented analytical approaches to determine the node energy consumption for mobile data collector schemes in a LS-WSN, and MULE (data collection using the data mule) model and SENMA (sensor network with mobile access point) model for determining the optimal number of clusters for minimizing the energy consumption were given. However, the above works did not refer to the problem of coverage holes in LS-WSNs for big data collection, which will significantly degrade the QoS of LS-WSNs.

Compared with normal WSNs, LS-WSNs have a significant issue to be handled for big data collection: coverage holes in wide geographic areas that are required to be monitored. The data in the areas is lost or cannot be transmitted to other nodes due to the exhaustion of its own very limited energy or external damage in a sensor network, where these areas are called coverage holes. In addition, coverage, as an important measure of the QoS of WSNs, reflects the environment perception service that the WSN can provide, including information acquisition and effective transmission. Therefore, healing coverage holes for big data collection in LS-WSNs becomes the key to improve the QoS of LS-WSNs.

Currently, solving the coverage hole healing (CHH) problem is a hot issue in WSNs. A series of efforts have been devoted to this problem over the past years [20–29]. In [20], the scenario where all nodes were mobile nodes and the concept of "virtual force" in physics was introduced. The mobile nodes between overlapping areas have "repulsion", and the adjacent mobile nodes in the covering hole have "attraction". The schemes proposed by Deng et al. [21] aimed at efficiently healing the confident information coverage holes while minimizing the total moving energy consumption of the dispatched mobile nodes, or maximizing the mobile nodes' average remaining energy after movement, or minimizing the maximum mobile energy consumption of each dispatched mobile node. For healing coverage holes, authors in [22] presented a geometric method based on the Voronoi diagram. Han et al. [23]

designed the centralized heuristic method to keep the balance of mobile sensor nodes consumption while minimizing the movement energy consumption of mobile sensor nodes. Latif et al. [24] proposed the avoidance of the coverage hole in underwater wireless sensor network and coverage hole repair of underwater wireless network operation. In [27], the proposed algorithms considered limited mobility of the nodes and can select the mobile nodes based on their degree of coverage overlapping. In order to heal coverage holes of the sensor network, nodes with higher degree of density are moved to maintain uniform network density without increasing the coverage degree of the neighbors of a mobile node. A new game-theoretic method of distributed recover coverage holes was proposed in [28]. Authors [29] presented the tree-based coverage hole healing scheme, which partitions each large hole into some smaller holes and then deploys additional sensors to patch the position of each small sub-coverage hole.

Nevertheless, almost all earlier works [13–40] lack some considerations. Firstly, the CHH in WSNs was only built on small-scale scenarios and the total amount of data transmitted by sensor nodes in WSNs was small without considering large-scale situations. Furthermore, most of previous works assumed that the data volume transmitted by sensor nodes were consistent, which was too ideal for practical applications. Last but not least, it is not feasible and inappropriate that the existing algorithms did not consider healing coverage holes in LS-WSNs from the data-centric and big data perspective because LS-WSNs are actually data-centric networks and work in big data environments in the future.

Considering the problems in the above literatures, in this work, we investigate how to energy-efficiently solve the CHH problem in the LS-WSN by a greedy healing algorithm (GHA), where the LS-WSN under the topology control of the LEACH protocol contains a great mass of static sensor nodes and mobile sensor nodes. The static nodes are responsible for monitoring environment, while the mobile nodes are used to heal coverage holes. The main contributions of this paper are summarized as follows:

1) We further consider the CHH problem in the LS-WSN for big data collection, in which the data volume transmitted by sensor nodes is inconsistent. The main goal of the CHH is to select an optimal subset of mobile nodes from the set of all mobile nodes to be dispatched to coverage holes while maximizing the transmission times that all dispatched mobile nodes can transmit in their lifetime.

2) For handing the CHH problem, we devise a GHA based on the greedy-based heuristic solution. The GHA aims at efficiently increasing the transmission times that all dispatched mobile nodes can transmit in their lifetime to heal coverage holes while observably prolonging the network lifetime and dramatically enhancing the QoS of WSNs. In addition, the GHA is more appropriate and

feasible for healing coverage holes for big data collection in LS-WSNs due to its quite low computational complexity.

3) Simulation results validate that the proposed GHA can effectively heal the coverage holes which observably extends the lifetime of the LS-WSN and dramatically enhances the QoS of WSNs while increasing the TT, transmitted data volume (TDV) and average residual energy of all dispatched mobile nodes.

The rest of this paper is organized as follows. Section 2 briefly introduces the LS-WSN model and the energy consumption model. Section 3 defines the CHH problem. Section 4 proposes the GHA to effectively solve the CHH problem. Section 5 shows the performance in the simulation results. Finally, the conclusion is made in Section 6.

2 System model

In this section, we first briefly illustrate the LS-WSN model. Then, the energy consumption model is introduced.

2.1 Large-scale wireless sensor network model

Figure 1 illustrates all the constituent elements of the LS-WSN. We define a few terms about the sensor network model that will be used in the following sections.

Coverage hole In this paper, the data in the areas is lost or cannot be transmitted to other nodes due to the exhaustion of its own very limited energy in the WSN, where these areas are called coverage holes (such as h_1 in Fig. 1).

Sink The sink (such as b_1 in Fig. 1) takes charge of collecting data from cluster heads.

Static node The static node (such as s_1, s_2 or s_3 in Fig. 1) is responsible for monitoring environments, transmitting data to other nodes and receiving data from other nodes.

Dead node The dead node (such as d_1 in Fig. 1) is disabled to work due to energy depletion.

Mobile node The mobile node (such as m_1 in Fig. 1) takes charge of healing coverage holes that replaces dead node to work.

Intra-cluster node The intra-cluster node (such as s_1, s_2 or s_3 in Fig. 1) which is static sensor node or mobile sensor node that is responsible for monitoring environments, transmitting data to cluster head and receiving data from other nodes.

Cluster head The cluster head (such as c_1 in Fig. 1) is responsible for collecting overall data of nodes among its cluster and fuses the data from intra-cluster nodes.

The topology control of the LS-WSN adopts the LEACH algorithm which is a classic adaptive clustering algorithm [41]. The process of execution is in the form of working rounds, and each working round is divided into two stages including the stage of cluster establishment and the stage of data communication. In the stage of cluster establishment, neighboring nodes form clusters dynamically and randomly
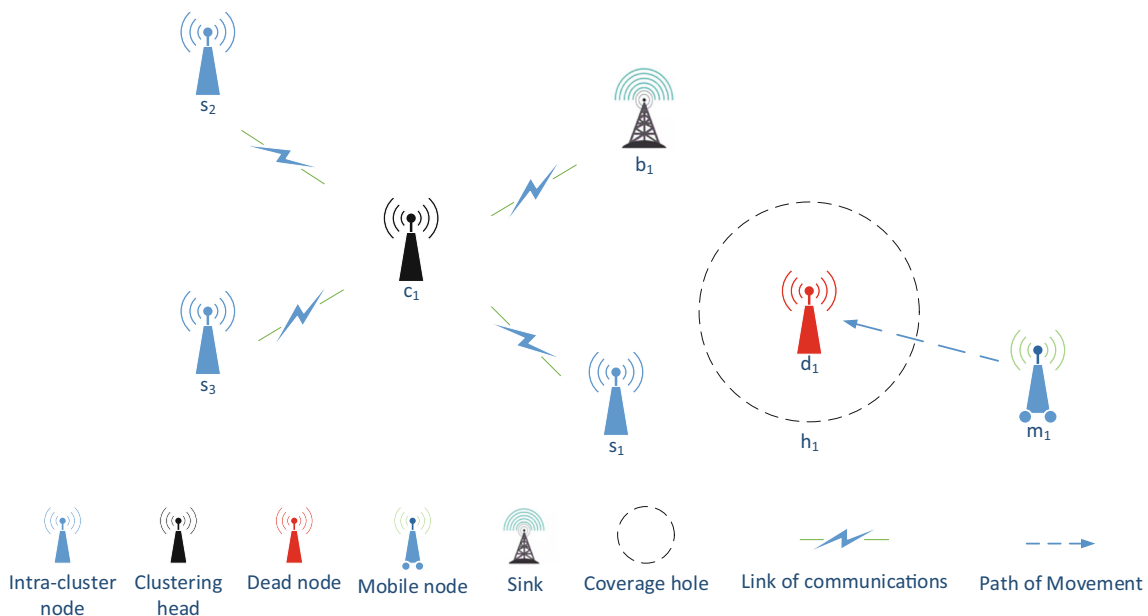

Fig. 1 Elements of the LS-WSN

elect cluster heads. In the data communication stage, the nodes in a cluster transmit data to the cluster head, which fuses the data and sends the results to the sink node. Meanwhile, the LEACH algorithm can make each node serves as a cluster head with equal probability and consumes energy in a relative balance. Next, we summarize some steps for the data collection process in the LS-WSN as follows:

Firstly, a static node transmits sensor data to a cluster head that is close to it. Secondly, the cluster head integrates the self-perception data and data of all transmission nodes in the cluster. Finally, the cluster head transfers the fused data to the sink node.

In addition, the mobile nodes move to the position of the coverage hole by the shortest path that is straight-line distance between a dead node and a mobile node (such as the dead node d_1 to the mobile node m_1 in Fig. 1) for healing coverage holes developed by the energy exhaustion of sensor nodes.

For the LS-WSN, we make the following assumptions:

1) The LS-WSN, which is deployed over wide geographic areas (two-dimensional space) that need to be sensed and contain a lot of sensor nodes consisting of static sensor nodes and mobile sensor nodes, generates the amount of data volume transmitted by sensors that can be called big data.
2) The data volume transmitted by each node is inconsistent and remains unchanged.
3) The static nodes and mobile nodes are randomly deployed.
4) Each node is aware of its own location via GPS or some other localization methods [42, 43], and it is considered that the immediate one-hop neighbors of a dead node know locations of the holes in the sensor network based on locations of the dead nodes [24].
5) The information of locations of coverage holes is known by one-hop neighbors of a dead node communicating with mobile nodes.

2.2 Energy consumption model

The energy consumption model plays an important role in the system model. In this paper, we refer to the first-order wireless communication model. Energy consumption can be divided into four parts: transmission circuit consumption, amplifying circuit consumption, receiving circuit consumption, and data fusion consumption. We set $E_{elec} = 50$ nJ/bit energy dissipation unit parameters for the sending or receiving circuit, $\epsilon_{amp} = 100$ pJ/bit/m^2 amplifying circuit of energy dissipation unit parameters, $E_{fus} = 5$ nJ/bit energy dissipation unit parameters for the data fusion, as listed in Table 1.

Table 1 Radio characteristics

Type of Energy Dissipation	Parameters
Transmitter dissipation	50 nJ/bit
Receiver dissipation	50 nJ/bit
Amplifier dissipation	100 pJ/bit/m^2
Data fusion dissipation	5 nJ/bit

In addition, the transmission and receiving energy consumption for a k-bit packet over distance d can be calculated by following two equations that are transmission circuit consumption $E_{Tx}(k, d)$ and receiving circuit consumption $E_{Rx}(k)$ respectively.

$$E_{Tx}(k, d) = E_{elec} * k + \epsilon_{amp} * k * d^2 \qquad (1)$$

where E_{elec} is the energy dissipation unit parameters for the sending or receiving circuit, k is the number of bits for the packet, ϵ_{amp} is the energy dissipation of transmission amplifier and d is the shortest distance between two nodes.

$$E_{Rx}(k) = E_{elec} * k \qquad (2)$$

where E_{elec} is the energy dissipation unit parameters for the sending or receiving circuit and k is the number of bits in the packet.

3 Problem definition and formulation

In this section, we describe the occurrence of coverage holes and the coverage hole healing problem as follows. Firstly, we define a few terms that are used in this paper:

> M: A collection of all mobile nodes denoted by $M = \{m_1, \ldots, m_L\}$.
> H: A collection of all coverage holes denoted by $H = \{h_1, \ldots, h_K\}$.

Transmission times (TT) The times that a node (nodes) can transmit in its (their) lifetime. It can be described as

$$T(E_r, B_e) = \frac{E_r}{B_e \times E_{elec}} \qquad (3)$$

where E_r is the residual energy of a single node, B_e is the transmitted data volume per working round, and E_{elec} is defined in Subsection 2.2 above. For example, the TT refers to that a node which has 10 J of energy and 1000 bits to transmit per working round can transmit 20 thousand times in its lifetime.

Transmitted data volume (TDV) The data volume that a node can transmit in its lifetime. It can be described as

$$D(E_r) = \frac{E_r}{E_{elec}} \tag{4}$$

where E_r is the residual energy of a single node, and E_{elec} is defined in Subsection 2.2 above. For example, the TDV means a node which has 10 J of energy can transmit 200 million bits in its lifetime.

In this paper, we do not consider the detection issue for coverage holes and only focus on how to effectively heal the generated coverage holes $H = \{h_1, …, h_K\}$ in an energy-efficient way. The process of healing coverage holes can be described as follows:

During the normal operation of the LEACH protocol, a set $H = \{h_1, …, h_K\}$ of coverage holes are generated in the LS-WSN due to energy exhaustion of sensor nodes and required to be covered by mobile sensor nodes, where the LS-WSN is described in Subsection 2.1 above. Next, for healing coverage holes, a subset $M'(M' \in M)$ of mobile nodes has been respectively dispatched to locations for the set $H = \{h_1, …, h_K\}$ of coverage holes, which are locations of dead nodes. Finally, the set $H = \{h_1, …, h_K\}$ of coverage holes has been successfully healed when the subset $M'(M' \in M)$ of mobile nodes arriving locations of coverage holes respectively. Next, the CHH problem will be defined.

Definition 1: CHH problem *The CHH problem primary focuses on how to select a subset $M'(M' \in M)$ of mobile nodes to heal coverage holes H such that after the movement, coverage holes H can be healed while maximizing the TT of mobile sensor nodes.*

According to the definition of the CHH problem, we consider that the following function (5) is sufficient to characterize the healing process of the CHH problem. Therefore, how to select a subset M' of mobile nodes from the set M of all mobile nodes to satisfy the objective function become the key.

The objective function is to maximize the TT of all mobile nodes to heal coverage holes, which is described as

$$\underset{m_l \in M'}{Max \sum} x_i(E_r, B, d_m^i) \tag{5}$$

where x_i is the TT of a single mobile node, which is calculated by

$$x_i(E_r, B, d_m^i) = \frac{E_r - e_m \times d_m^i}{B \times E_{elec}}, e_m \times d_m^i < E_r \tag{6}$$

where E_r is the residual energy of a single node, e_m is the energy required by unit distance (step), d_m^i is the is shortest distance between a coverage hole i and a node m, B is the data volume transmitted by a single sensor node, and E_{elec} is defined in Subsection 2.2 above.

In the fact, from the data-centric perspective, maximizing the TT for mobile nodes can be acquired through the optimized moving trace in the LS-WSN. In addition, the moving trace at the optimized strategy will significantly raise the remaining energy of mobile nodes since the movement of mobile nodes consumes a large quantity of energy, and it is effective that the way prolongs the lifetime of WSN because the energy consumption is used to measure the lifetime of WSN.

4 Solution for the problem

In this section, from data-centric perspective, we propose an energy-efficient solution named as GHA via the greedy-based heuristic method to effectively solve the CHH problem with significantly lower complexity due to some specific properties of LS-WSNs in the big data environment, limited computational capabilities of sensor nodes, lager-scale areas and transmitted data volume among sensor nodes. More specifically, the GHA finds a suboptimal subset M' of mobile sensors from the set M of all mobile sensors, where M' is obtained by maximizing the TT of a single dispatched mobile node one by one via a greedy-based heuristic strategy. The details of the GHA are listed in Algorithm 1.

Execute Steps 1 to 11 when the set H of coverage holes is detected, and the sink broadcasts the location of coverage hole $h_i \in H$ in Step 2, where $i = 1, 2, ..., K$.

In Steps 3 to 9, for each mobile sensor node $m_l \in M$, calculate the TT $x_i(E_r, B, d_m^i)$ that can transmit times in its lifetime after healing the current coverage hole h_i by (6) and select the maximum TT $x_i'(E_r, B, d_m^i)$ for the mobile node m_l' from the set M of all mobile nodes, where $l = 1, 2, ..., L$. Finally, the dispatched mobile node m_l' is removed from the set M of all mobile nodes.

In Step 10, the mobile node m_l' having the maximum TT $x_i'(E_r, B, d_m^i)$ constructs the subset M' of mobile nodes, and the subset M' of mobile nodes is dispatched to heal the current set H of coverage holes. Next, some explains about GHA will be introduced as follows.

There is no efficient algorithm to find the optimal solution to satisfy the objective function (6) since the CHH problem is a NP-hard problem for which a globally optimal solution is difficult to be obtained. Although the optimal solution can be found by exhaustive search, its exponential complexity is inappropriate and is too ideal for big data collection in LS-WSNs. Though the calculation result of GHA does not meet the optimal solution of the objective function (6), the computational complexity is quite low and the result is close to the optimal solution.

From Steps 1 to 11, the computational complexity of the GHA is $O(KL)$, which is lower than $O(L^K)$ of exhaustive search, obviously.

Algorithm 1   Greedy Healing Algorithm

1.	**while** $H = \{h_1, \ldots, h_i\}$ is detected
2.	The sink broadcasts the location of $h_i \in H$
3.	**for** each coverage hole $h_i \in H$
4.	**for** each mobile node $m_l \in M$
5.	calculate the TT $x_i(E_r, B, d_m^i)$ for m_l by (6)
6.	select the maximum TT $x_i'(E_r, B, d_m^i)$ for m_l' from M
7.	remove the m_l' from the set M
8.	**end for**
9.	**end for**
10.	dispatch M' to heal H
11. **end while**	

5 Simulation results

In this section, we firstly define some parameters for three simulations. Then, we give some descriptions of experiment results. Finally, we summarize the results according to our environments.

5.1 Simulation parameters

The environment used for the simulation is MATLAB. The LEACH [41] has been used as the topology control algorithm in these simulations, which allows each node to equally serve as a cluster head and consume energy relatively balanced. A wide geographic area (1000×1000 m^2) of LS-WSN exceeding 700 nodes which transmit data volume of 5000 to 8000 bits has been used. The experiments are done for three kinds of node arrangements (500 static nodes and 300 mobile nodes, 400 static nodes and 300 mobile nodes, and 500 static nodes

Table 2   Network parameters

Parameters	Values
Topological control algorithm	LEACH
Area size	1000×1000 m^2
Types of nodes	Static, mobile
Number of nodes	700 or 900 nodes
Nodes distribution	Random distribution
Transmission radius	100 m
Data volume of a node	5000–8000 bits
Working round	1500 times

and 400 mobile nodes), respectively. The distribution of all nodes is randomly generated. The transmission radius of each sensor node is set to 100 m. The specific parameters are listed in Table 2.

The energy consumption of transmit-receive circuits has been given in Subsection 2.2. Next, we set other energy consumption parameters. The initial energy of a static node and a mobile node are set to 10 J and 100 J, respectively. The mobile energy consumption for mobile sensor nodes is set to 0.5 J/m. The specific parameters are in listed Table 3.

5.2 Performance results and comparison

We conduct a series of experimental simulations to verify the TT, TDV and average residual energy of mobile nodes, and evaluate the gain of the proposed scheme of GHA compared with the method of no healing and random healing. All experiments below with above simulation parameters are performed in the MATLAB platform.

In the first experiment, as shown in Figs. 2 and 3, it evaluates the effect of the working rounds on the TT with different schemes and numbers of static nodes and mobile nodes. With the different schemes and numbers of nodes in Fig. 2 and

Table 3   Energy consumption parameters

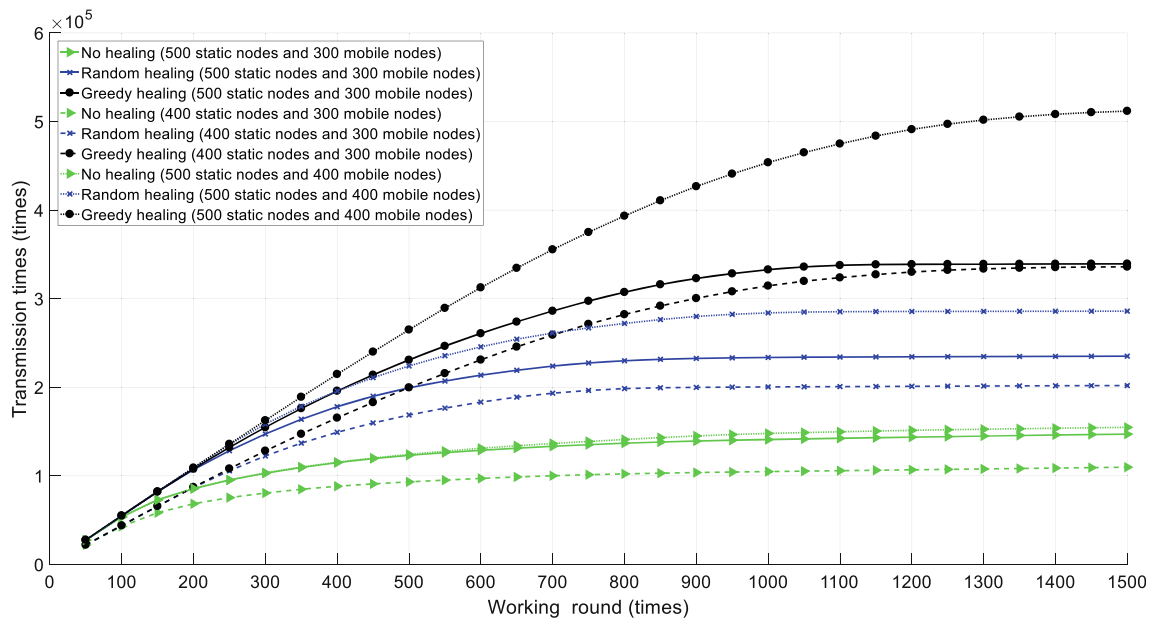| Parameters	Values
Initial energy of a static node	10 J
Initial energy of a mobile node	100 J
Energy consumption of a mobile node	0.5 J/m

Fig. 2 Comparison of TT

Fig. 3, we can observe it is no difference in TT at early working rounds due to the absence of coverage holes; as the working round grows, the TT using GHA as the measure is significantly greater than no healing and random healing, which is because the more the required hole healing, the more mobile nodes need to be dispatched. On average, the GHA can raise to 118.39% (500 static nodes and 300 mobile nodes), 145.47% (400 static nodes and 300 mobile nodes), and 172.22% (500 static nodes and 400 mobile nodes) compared with random healing at 1500 working rounds, respectively. In addition, decreasing the amount of static nodes can significantly have more TT when the number of mobile nodes is constant (e.g., 500 static nodes, 400 static nodes, and 300 mobile nodes), reducing the number of mobile nodes can dramatically add more TT under the same the number of static nodes (e.g., 400 mobile nodes, 300 mobile nodes, and 500 static nodes), and vice versa.

In the second experiment, as shown in Figs. 4 and 5, we compare the data volume by varying the working rounds with different schemes and numbers of static nodes and mobile nodes. With the different methods and numbers of nodes in Figs. 4 and 5, there is no markedly difference in data volume at early working rounds because of the lack of coverage holes; as the working round grows, the data volume is significantly greater than no healing and random healing when GHA is used, which is because more mobile nodes has been dispatched with the increasement of coverage holes. On average, the GHA can have 120.2739% (500 static nodes and 300 mobile nodes), 148.21% (400 static nodes and 300 mobile nodes), and 172.46% (500 static nodes and 400 mobile nodes) more data volume than random healing at 1500 working rounds, respectively. In addition, the data volume has been raised with cutting down the amount of static nodes when the number of mobile nodes is constant (e.g., 500 static nodes,

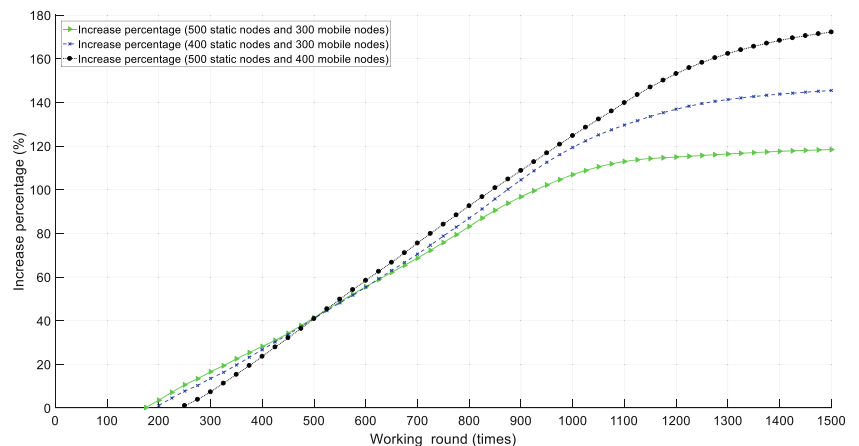**Fig. 3** Increased percentage of TT (GHA vs. random healing)
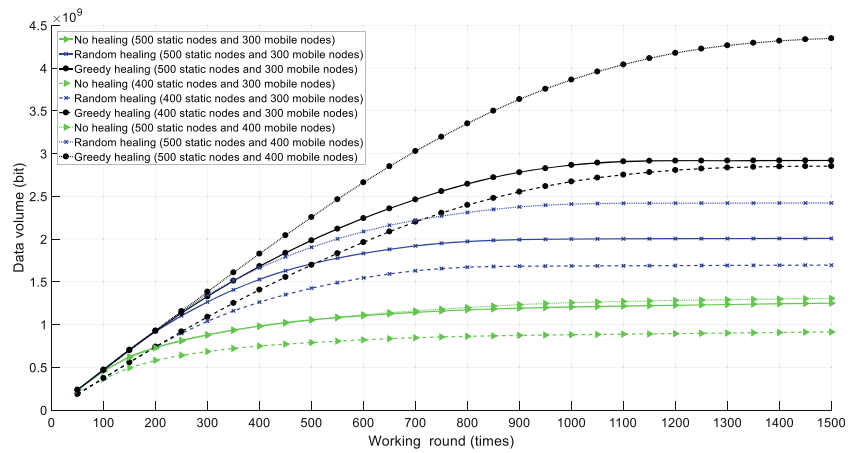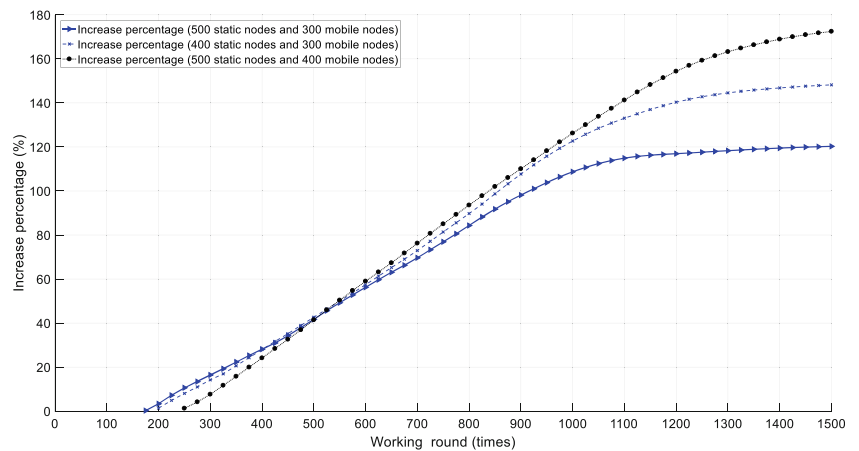
Fig. 4 Comparison of TDV

**Fig. 5** Increased percentage of TDV (GHA vs. random healing)

400 static nodes, and 300 mobile nodes), the data volume has been increased with the number of mobile nodes under the same the number of static nodes (e.g., 400 mobile nodes, 300 mobile nodes, and 500 static nodes), and vice versa.

The third experiment explores the effect of working rounds on the total residual energy with different schemes and numbers of static nodes and mobile nodes. In Fig. 6, it is same as

the average residual energy at early working rounds. That is because the number of dispatched mobile nodes are greater than that of the coverage holes with the augment of working rounds; thus, coverage holes are healed increasingly. From Fig. 6, we can observe that the proposed GHA can have more residual energy than random healing, and can significantly reduce 72.83% (500 static nodes and 300 mobile nodes),

Fig. 6 Comparison of average residual energy
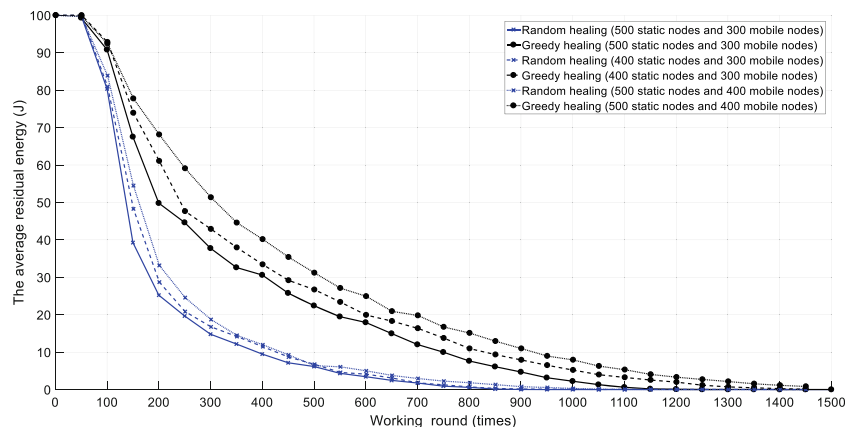
82.04% (400 static nodes and 300 mobile nodes), and 95.32% (500 static nodes and 400 mobile nodes) of the total moving energy, respectively. In addition, decreasing the amount of static nodes can obviously have more average residual energy when the number of mobile nodes is constant (e.g., 500 static nodes, 400 static nodes, and 300 mobile nodes), reducing the number of mobile nodes can dramatically improve more average residual energy under the same number of static nodes (e.g., 400 mobile nodes, 300 mobile nodes, and 500 static nodes), and vice versa.

According to the above simulation results, we can summarize the experiments as follows:

1) The proposed GHA always outperforms the random healing scheme in terms of higher TT, higher TDV and higher average residual energy of mobile nodes. This is due to that the GHA solution can fully maximize the TT of a single dispatched mobile node by a greedy-based heuristic strategy, and make full use of the data-centric principle in LS-WSNs.

2) The increasing ratio of the TT and TDV in the GHA scheme increases significantly with the number of mobile nodes under the same number of static nodes, and increase with decreasing the amount of static nodes when the number of mobile nodes is invariant, and vice versa.

3) The GHA can effectively heal the coverage holes which observably prolongs the lifetime of the LS-WSN and enhances the QoS of the LS-WSN (GHA vs. random healing and no healing) because TT, TVD and average residual energy are important measures for the LS-WSN.

6 Conclusions

For handing the CHH problem, in this paper, the GHA for big data collection in a LS-WSN was proposed, in which the data volume transmitted by sensor nodes may be inconsistent among them. The key for the CHH problem is to find an optimal subset from all residual mobile nodes to repair coverage holes. The GHA can acquire a suboptimal subset with low computational complexity, which was based on the energy-efficient heuristic method and the data-centric perspective. Simulation results have validated that the GHA can effectively heal coverage holes which observably extends the lifetime of WSNs and significantly enhances the QoS of WSNs while increasing the TT, TDV and average residual energy for all dispatched mobile nodes.

References

1. Oliveira LML, Rodrigues JJPC (2011) Wireless sensor networks: a survey on environmental monitoring. J Commun 6(2):143–151
2. Wang B (2011) Coverage problems in sensor networks: a survey. ACM Comput Surv 43(4):32
3. Arthur WB (2011) The second economy, McKinsey Quart
4. Liu Y, He Y, Li M, Wang J, Liu K, Li X (2013) Does wireless sensor network scale? A measurement study on green orbs. IEEE Trans Parallel Distrib Syst 24(10):1983–1993
5. Nasipuri A, Cox R, Conrad J, Van Der Zel L, Rodriguez B, and McKosky R (2010) Design considerations for a large-scale wireless sensor network for substation monitoring, in Proc. 5th IEEE Int. Workshop Pract. Issues Build. Sens. Netw. Appl., pp. 866–873
6. Jia M, Yin Z, Guo Q, Liu G, Gu X (2018) Toward improved offloading efficiency of data transmission in the IoT-cloud by leveraging secure truncating OFDM. IEEE Internet of Things J 99:1–8
7. Jia M, Yin Z, Guo Q, Liu G, Gu X (2018) Downlink design for spectrum efficient IoT network. IEEE Int of Things Journal 5(5): 3397–3404
8. Jia M, Liu X, Gu X, Guo Q (2017) Joint cooperative spectrum sensing and channel selection optimization for satellite communication systems based on cognitive radio. Int J Satellite Commun Netw 35(2):139–150
9. Jia M, Gu X, Guo Q, Xiang W, Zhang N (2016) Broadband hybrid satellite-terre- strial communication systems based on cognitive radio towards 5G. IEEE Wirel Commun 23(6):96–106
10. Jia M, Liu X, Yin Z, Guo Q, Gu X (2016) Joint cooperative spectrum sensing and spectrum opportunity for satellite cluster communication networks. Ad Hoc Net 58:231–238
11. Jia M, Li D, Yin Z, Guo Q, Gu X (2018) High spectral efficiency secure communications with non-orthogonal physical and multiple access layers. IEEE Internet of Things J. 99:1–8
12. Cao X, Liu L, Cheng Y, Shen X (2018) Towards energy-efficient wireless networking in the big data era: a survey. IEEE Commun Surveys & Tutorials 20(1):303–332, First Quarter
13. Takaishi D, Nishiyama H, Kato N, Miura R (2014) Toward energy efficient big data gathering in densely distributed sensor networks. IEEE Trans Emerg Top Comput 2(3):388–397
14. Wu M, Tan L, Xiong N (2015) A structure fidelity approach for big data collection in wireless sensor networks. Sensors 15:248–273
15. Zhu J, Yin X, Bai J, Wang Y (2016) Mobility-assisted big data collecting in wireless sensor networks. Int J Distrib Sens Netw 12(8):18
16. Ding X, Tian Y, Yu Y (2016) A real-time big data gathering algorithm based on indoor wireless sensor networks for risk analysis of industrial operations. IEEE Trans. Ind. Inform. 12(3):1232–1242
17. Rani S, Ahmed SH, Talwar R, Malhotra J (2017) Can sensors collect big data? An energy-efficient big data gathering algorithm for a WSN. IEEE Trans Ind Inform 13(4):1961–1968
18. Din S, Ahmed A, Paul A, Rathore MMU, Jeon G (2017) A cluster-based data fusion technique to analyze big data in wireless multi-sensor system. IEEE Access 5:5069–5083
19. Ang KL, Seng JKP, Zungeru AM (2018) Optimizing energy consumption for big data collection in large-scale wireless sensor networks with mobile collectors. IEEE Syst J 12(1):616–626

20. Senouci MR, Mellouk A, Assnoune K (2014) Localized movement-assisted sensor deployment algorithm for hole detection and healing. IEEE Trans. Parallel Distrib. Syst. 25(5):1267–1277

21. Deng X, Tang Z, Yang LT, Lin M, Wang B (2018) Confident information coverage hole healing in hybrid industrial wireless sensor networks. IEEE Trans. Ind. Informat. 14(5):2220–2229

22. Qiu C, Shen H, and Chen K (2015) An energy-efficient and distributed cooperation mechanism for k-coverage hole detection and healing in WSNs, in Proc. IEEE 12th Int. Conf. Mobile Ad Hoc Sens. Syst., Oct. pp. 73–81

23. Han G, Liu L, Jiang J, Shu L, Hancke G (2017) Analysis of energy-efficient connected target coverage algorithms for industrial wireless sensor networks. IEEE Trans Ind Informat 13(1):135–143

24. Latif K, Javaid N, Ahmad A, Khan ZA, Alrajeh N, Khan MI (2016) On energy hole and coverage hole avoidance in underwater wireless sensor networks. IEEE Sensors J 16(11):4431–4442

25. Wang YC, Hu CC, Tseng YC (2008) Efficient placement and dispatch of sensors in a wireless sensor network. IEEE Trans Mobile Comput 7(2):262–274

26. Yan F, Vergne A, Martins P, Decreusefond L (2015) Homology-based distributed coverage hole detection in wireless sensor networks. IEEE/ACM Trans Netw 23(6):1705–1718

27. Sahoo PK, Liao W (2015) HORA: a distributed coverage hole repair algorithm for wireless sensor networks. IEEE Trans Mob Comput 14(7):1397–1410

28. Abolhasan M, Maali Y, Rafiei A, Ni W (2016) Distributed hybrid coverage hole recovery in wireless sensor networks. IEEE Sensors J 16(23):8640–8648

29. Li W, Wu Y (2016) Tree-based coverage hole detection and healing method in wireless sensor networks. Comput Netw 103(24):33–43

30. Liu B, Ren F, Shen J, Chen H (2010) Advanced self-correcting time synchronization in wireless sensor networks. IEEE Commun Lett 14(4):309–311

31. Zhang W, Yin Q, Chen H, Gao F, Ansari N (2013) Distributed angle estimation for localization in wireless sensor networks. IEEE Trans Wirel Commun 12(2):527–537

32. Liu B, Chen H, Zhong Z, Poor HV (2010) Asymmetrical round trip based synchronization-free localization in large-scale underwater sensor networks. IEEE Trans Wirel Commun 9(11):3532–3542

33. Wang G, Chen H, Li Y, Jin M (2012) On received-signal-strength based localization with unknown transmit power and path loss exponent. IEEE Wirel Commun Lett 1(5):536–539

34. Chen H, Liu B, Huang P, Liang J, Gu Y (2012) Mobility-assisted node localization based on TOA measurements without time synchronization in wireless sensor networks. ACM Mob Net App 17(1):90–99

35. Chen L, Chen W, Wang B, Zhang X, Chen H, Yang D (2011) System-level simulation methodology and platform for mobile cellular systems. IEEE Commun Mag 49(7):148–155

36. Wang G, Chen H (2011) An importance sampling method for TDOA-based source localization. IEEE Trans Wirel Commun 10(5):1560–1568

37. Chen H, Gao F, Marins MHT, Huang P, Liang J (2013) Accurate and efficient node localization for mobile sensor networks. ACM Mob. Net. App. 18(1):141–147

38. Chen H, Wang GWZ, So HC, Poor HV (2011) Non-line-of-sight node localization based on semi-definite programming in wireless sensor networks. IEEE Trans Wirel Commun 11(1):108–116

39. Huang P, Chen H, Xing G, Tan Y (2009) SGF: a state-free gradient-based forwarding protocol for wireless sensor networks. ACM Trans Sensor Net 5(2):1–14

40. Chen H, Shi Q, Tan R, Poor HV, Sezaki K (2010) Mobile element assisted cooperative localization for wireless sensor networks with obstacles. IEEE Trans Wirel Commun 9(3):956–963

41. Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun 1(4):660–670

42. Amundson I, Koutsoukos XD (2009) A survey on localization for mobile wireless sensor networks, in Proc. 2nd Int. Conf. Mobile Entity Localization Tracking GPS-Less Environ., pp. 235–254

43. Liu Y, Yang Z (2010) Location, localization, and localizability: location-awareness Technology for Wireless Networks. Springer, New York