Wang Y, Liu YX, Zhu SJ *et al.* Approximation designs for energy harvesting relay deployment in wireless sensor networks. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 37(4): 779–796 July 2022. DOI 10.1007/s11390-022-1964-5

Approximation Designs for Energy Harvesting Relay Deployment in Wireless Sensor Networks

Yi Wang[1] (王 怿), Yi-Xue Liu[2] (刘忆雪), Shun-Jia Zhu[3] (朱舜佳)
Xiao-Feng Gao[4,*] (高晓沨), *Member, CCF, ACM, IEEE*, and
Chen Tian[2] (田 臣), *Senior Member, IEEE, Member, CCF, ACM*

[1] *School of Modern Posts, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
[2] *State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China*
[3] *ByteDance Inc., Shanghai 201100, China*
[4] *Key Laboratory of Artificial Intelligence of Ministry of Education, Department of Computer Science and Engineering Shanghai Jiao Tong University, Shanghai 200240, China*

E-mail: yiwang@njupt.edu.cn; mf20330052@smail.nju.edu.cn; zhushunjia@sjtu.edu.cn; gao-xf@cs.sjtu.edu.cn tianchen@nju.edu.cn

Received October 12, 2021; accepted June 6, 2022.

Abstract Energy harvesting technologies allow wireless devices to be recharged by the surrounding environment, providing wireless sensor networks (WSNs) with higher performance and longer lifetime. However, directly building a wireless sensor network with energy harvesting nodes is very costly. A compromise is upgrading existing networks with energy harvesting technologies. In this paper, we focus on prolonging the lifetime of WSNs with the help of energy harvesting relays (EHRs). EHRs are responsible for forwarding data for sensor nodes, allowing them to become terminals and thus extending their lifetime. We aim to deploy a minimum number of relays covering the whole network. As EHRs have several special properties such as the energy harvesting and depletion rate, it brings great research challenges to seek an optimal deployment strategy. To this end, we propose an approximation algorithm named Effective Relay Deployment Algorithm, which can be divided into two phases: disk covering and connector insertion using the partitioning technique and the Steinerization technique, respectively. Based on probabilistic analysis, we further optimize the performance ratio of our algorithm to $(5 + 6/K)$ where K is an integer denoting the side length of a cell after partitioning. Our extensive simulation results show that our algorithm can reduce the number of EHRs to be deployed by up to 45% compared with previous work and thus validate the efficiency and effectiveness of our solution.

Keywords approximation algorithm, constraint relay deployment, energy harvesting, wireless sensor network

1 Introduction

Wireless sensor networks (WSNs) which are composed of sensor nodes (SNs) have a broad range of applications such as disaster prevention [1], region mapping [2], and environment monitoring [3].

In large-scale WSNs, due to the limited communication range of devices, data routing and transmission are performed by the efforts of many nodes. One sensor node not only performs its sensing tasks but also has to forward data for other nodes. This additional burden leads to fast energy depletion of sensor nodes. To resolve this problem, some tiny and convenient devices called Relay have been widely used in WSNs, especially for data transmission. The relay node (RN) is a signal

Regular Paper
Special Section of Xia Peisu Young Scholars Forum 2021
This work was supported by the Key-Area Research and Development Program of Guangdong Province of China under Grant No. 2020B0101390001, the Shanghai Municipal Science and Technology Major Project of China under Grant No. 2021SHZDZX0102, the National Natural Science Foundation of China under Grant No. 62072228, the Fundamental Research Funds for the Central Universities of China, the Collaborative Innovation Center of Novel Software Technology and Industrialization of Jiangsu Province of China, and the Jiangsu Innovation and Entrepreneurship (Shuangchuang) Program of China.
*Corresponding Author

forwarding device to amplify and forward the wireless signals to users. Since it is tiny, convenient and mobile, with excellent signal amplification effect, deploying relays in multi-hop networks becomes an effective method to improve system capacity, wireless coverage, and service quality. Besides, relays also play an important role in the development of B5G and 6G [4–6]. Recent studies [7–9] separate the responsibilities of nodes in WSNs where SNs only transmit their own packets and RNs forward packets for SNs. Experimental results reveal that relays can significantly improve the energy efficiency and prolong the network lifetime [8].

Although these relays help with the consistency of networks, powering remains a challenging problem [10]. Recharging relays is unrealistic for nodes placed in the hard-to-reach, hazardous or toxic areas where there is no grid power connected. For nodes powered by batteries, long lifetime batteries may not be suitable due to the limited size of nodes while small capacity batteries require frequent replacement, which is time-consuming and harmful to the communications. In this case, energy harvesting relays (EHRs) have emerged and gained much attention in recent years. EHR is a new kind of relays with energy harvesting capabilities, which can be recharged by the surrounding environment [11–15]. Thanks to this energy sustainability, the incorporation of EHRs in WSNs can provide long-term dependable service and is going to be the trend of future. A lot of work has been done for the deployment of energy harvesting relays. In the paper [16], a location-aware multimedia relay node deployment (MRND) scheme was proposed to support the future quickly deployable wireless multimedia services around mmWave frequency band. Moreover, a Quality of Experience (QoE)-driven location-aware optimization algorithm of autonomous relay BS deployment scheme was proposed to achieve the maximized system-level QoE under energy constraints. However, the application of energy harvesting relays also introduces some new problems. For instance, EHRs may doze off if they do not harvest enough energy to consume, leaving the network disconnected. In fact, deploying EHRs in WSNs is a challenging work, with a great number of factors that need to be taken into account. Table 1 shows the research on ENRs in recent years.

The placement problem for normal relay nodes is also important and has been extensively studied. Numerous efforts have been made on Facility Allocation Problem (FAP) [26–28] and Unit Disk Cover Problem (UDC) [29–31] which solved the problem to some extent. However, the solutions do not consider the additional complexity of energy harvesting, i.e., the energy that SNs consume constrains the maximum number of nodes that each EHR can cover. In this case, these solutions may be inapplicable in practice. Although this area has many potential applications, there have been few studies about it yet. The existing work [23, 32] revealed the dramatic improvement in performance of WSNs after the deployment of EHRs but ignored the coexistent limitations of EHRs. For instance, the energy causality constraint, which means the energy a node consumes cannot exceed the energy it harvests, is indeed the primary limitation and challenge of designing an energy harvesting system [33].

Another problem is that the existing work [34] on the Relay Node Placement (RNP) problem requires a set of

Table 1. Research on Energy Harvesting Relays in Recent Years

Research Direction	Reference	Contribution	Year
Energy harvesting	[17]	Maximizing the energy efficiency while taking both radio transmission power and static circuit power into account	2015
	[18]	Characterizing R-E regions for multiple access broadcast (MABC) and time division broadcast protocols	2019
	[19]	Implementing the technique of energy harvesting in full duplex (FD) cognitive radio network	2019
	[20]	Diversity-multiplexing tradeoff (DMT) of non-orthogonal EH based amplify-and-forward (EH-NAF) protocol	2017
Relay selection	[21]	An optimal offline joint relay selection and power allocation scheme	2015
	[22]	Proposing a joint optimal power allocation and relay selection strategy using the Lagrangian multiplier method	2018
Relay deployment	[23]	Proposing a model to ensure sustainable coverage and prolong the network lifetime using integer linear programming (ILP)	2015
	[24]	Proposing a linear program-based heuristic (LPBH), a two-stage heuristic (TSH) and a sensing- and routing-integrated greedy heuristic (SRIGH)	2019
	[25]	A tractable framework of the location deployment and mobility management of EH-SCBSs	2019

candidate locations which constrains designs of the deployment. Misra *et al.*[34] studied the constrained relay node placement problem in an energy harvesting network. However, the EHRs are constrained to be placed at only the candidate locations, where the potential of the locations is known as a priori. In effect, to fully-tap the advantages of EHRs, we need algorithms that do not require any candidate locations and take the properties of energy harvesting into account. These reasons motivate our study on Energy Harvesting Relay Deployment (EHRD) in wireless sensor networks. We propose a solution that seeks a deployment scheme with the minimum number of EHRs to cover the whole WSN. With the help of these special relay nodes, all the SNs will become terminals, only sensing and transmitting their own data. Consequently, the workload of SNs is lessened and the lifetime of the network is prolonged. Our research may help offer comparatively low-cost and easy-to-implement solutions to upgrade the networks without energy harvesting capacities.

We first formally define and formulate the Energy Harvesting Relay Deployment problem. Then we propose an approximation algorithm named Effective Relay Deployment Algorithm (ERDA). ERDA consists of two phases: disk covering and connector insertion. In the first phase, we seek a degree constrained disk cover (DC-DC) of the WSN, using the partitioning technique and shifting policy. We divide the plane into small squares, called cells. The approximation ratio of this part is $(2+6/K)$, where K is the side length of the cells, and we will discuss this more precisely in Section 5. In the second phase, we apply an algorithm for Steiner Minimum Tree with Minimum Steiner Points (SMT-MSP) to connect the DC-DC, and the approximation ratio of this part is 3. Hence, the total approximation ratio of ERDA is $(5+6/K)$. Through extensive simulations, we find that ERDA can significantly reduce the deployment cost of EHRs.

To the best of our knowledge, this is the first one of its type to study the EHR deployment problem without candidate locations in conventional wireless sensor networks. This study will provide a reference on how to upgrade the existing WSNs, which has both theoretical and practical significance in the related area. Fig.1 illustrates the contributions of four important problems studying the EHR deployment.

The major intellectual contributions of our work can be summarized as follows.

1) We establish a comprehensive network model and formulate the Energy Harvesting Relays Deployment (EHRD) problem in wireless sensor networks. Meanwhile, we give the proof of its NP-hardness.

2) We propose a novel polynomial-time approximation algorithm ERDA and theoretically evaluate its performance. We also utilize a shifting policy to reduce the approximation ratio to $(5+6/K)$ where K is a constant value for plane partitioning.

3) We conduct extensive simulations to evaluate the performance of ERDA. The numerical results show that up to 45% reduction of the number of EHRs to be deployed can be achieved when compared with previous work and thus demonstrate that our algorithm is well designed and has a great practical value.

The rest of the paper is organized as follows. First, we review the related work in Section 2. Then, we give some definitions and introduce the system model of WSNs in Section 3. In Section 4, we formally define the EHRD problem. Next, we propose an adaptive approximation algorithm in Section 5. Subsequently, we analyze the algorithm's performance in Section 6. Finally, we show the simulation results in Section 7 and conclude the paper in Section 8.

2 Related Work

In this section, we briefly describe the existing research studies on the facility allocation problem (FAP), the unit disk covering (UDC) problem, the relay node placement (RNP) problem and the energy harvesting relay deployment (EHRD) problem.

The facility allocation problem is a strategic decision problem used to determine an optimal number and locations of facilities to be established, and also to design an optimal distribution of the network. Xu and Shen[35] studied the problem of Fault-Tolerant Facility Allocation (FTFA). FTFA aims to allocate each site a proper number of facilities. The objective is to find such an allocation that minimizes the total combined cost for facility operating and service accessing. Tsao *et al.*[36] proposed a continuous approximation (CA) model to determine where to locate the regional distribution centers (RDCs) and how to assign retail stores to RDCs. It is worth noting that FAP does not consider the distance constraint but increases the cost when the distance gets larger. However, in the EHRD problem, the exceeding of a certain distance is strictly undesirable which may cause disconnections.

The UDC problem is known to be NP-hard and several approximation algorithms have been designed over the past three decades. Given a set of points in the

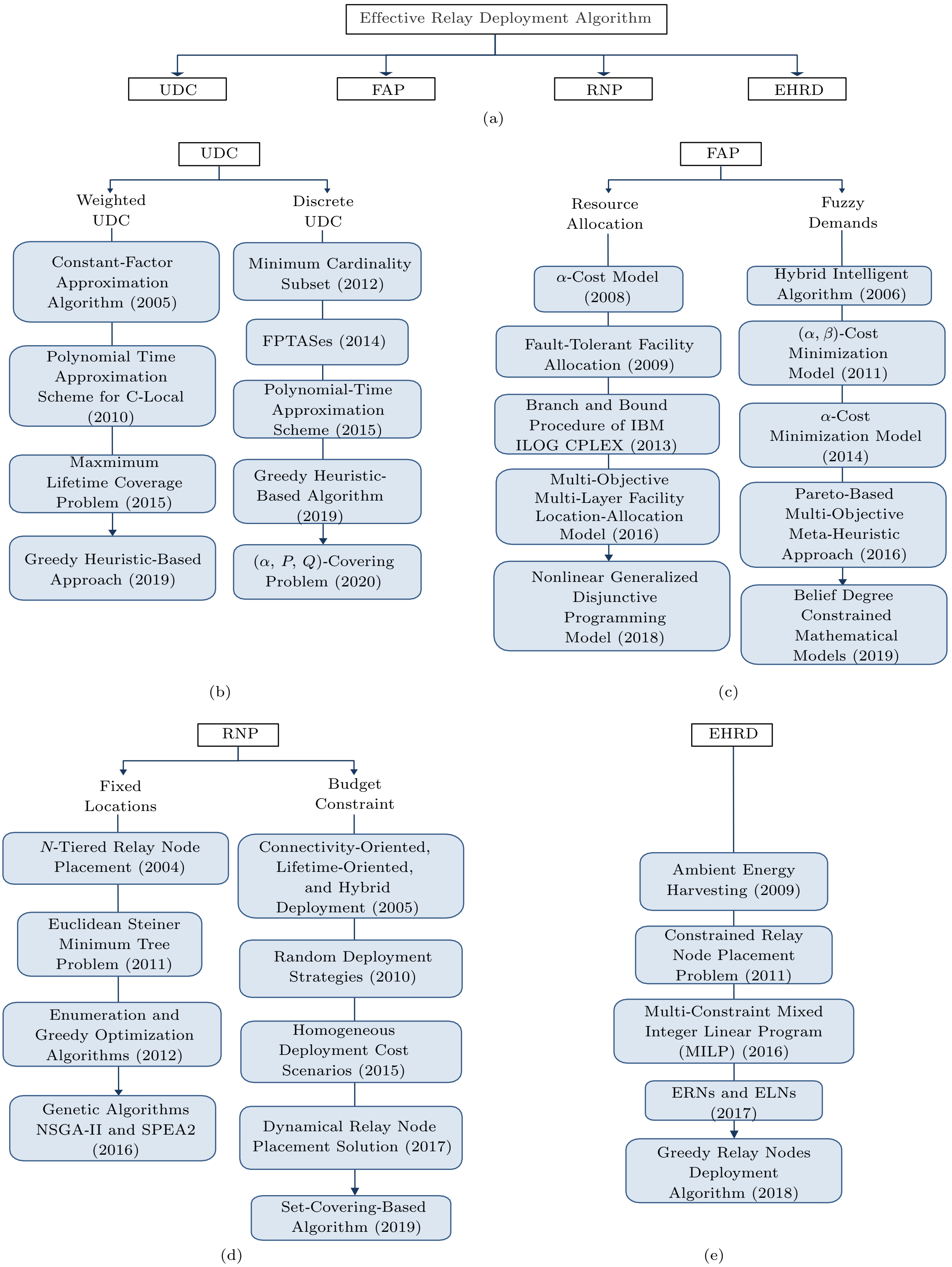


Fig.1. Introduction of related work on ERDA. (a) Key problems of EHR deployment. (b) Existing solutions to UDC. (c) Existing solutions to FAP. (d) Existing solutions to RNP. (e) Existing solutions to EHRD.

plane, the Unit Disk Covering problem computes the minimum number of unit disks that can cover all the points, along with a placement of the disks. The UDC problem has some similarities with the following RNP and EHRD problems and has applications in extending the sensor network lifetime by selecting the minimum number of nodes to cover each location. However, the UDC problem only focuses on the performance in time and connectivity, while the EHRD problem has more constraints for energy harvesting.

The RNP problem in WSNs, a well-known NP-hard problem, has been extensively studied, which aims at improving some properties, such as the lifetime. Table 2 briefly summarizes the problem formulations and constraints of some relay deployment algorithms. Most of RNPs in WSNs focus on two objectives: connectivity and survivability. Connectivity means that the graph induced from the WSN is connected while survivability has a higher requirement that the graph must be k-connected. For the RNP problem on the network connectivity, Lin and Xue [43] first gave a 5-approximation algorithm with a Steinerization technique called Steiner Minimum Tree with Minimum Number of Steiner Points (SMT-MSP). Then their studies were extended by [44–49] and the best known approximation ratio is 3 in [45]. Du *et al.* [45] also gave a randomized algorithm with a performance ratio of 2.5 and the probability of at least 1/2. For survivability, when $k = 2$, Kashyap *et al.* [50] presented a 10-approximation algorithm. For any $k \geqslant 2$, Bredin *et al.* [51] proposed polynomial time $O(1)$-approximation algorithms. Then Willson *et al.* [52] made a great progress. For $k = 1, 2$, they proposed a $(3 + \varepsilon)$-approximation algorithm. For $k \geqslant 3$, they presented an algorithm of which the performance ratio is $(4 + \varepsilon)$.

The above researches all studied the case that sensor nodes and relay nodes had the same communication range. Lloyed and Xue [53] first studied the RNP problem on connectivity where the communication ranges of sensor nodes and relay nodes were different, and presented a 7-approximation algorithm to resolve it. Then Zhang *et al.* [9] studied the RNP problem on survivability in this case. When $k = 2$, they gave a 14-approximation algorithm.

For the EHRD problem, Misra *et al.* [32] aimed to improve the network connectivity and survivability with a minimum number of added EHRs, using the Steinerization technique. Djenouri and Bagaa [23] considered placing EHRs near none-harvesting SNs to realize power-efficient coverage, reducing EHRD to a minimum weighted connected dominating set problem. These two studies both revealed the great potentials of placing EHRs. Nevertheless, they both ignored the limitations of EHRs. Their deployment strategies do not consider the energy causality constraint. This omission may lead to fast energy depletion of EHRs, thus leading to performance degradation or even the breakdown of the wireless sensor network.

3 System Model

In this section, we first give some definitions and notations in this paper and then introduce the system model. With thorough analysis, we conclude the two constraints in the EHRD problem: connectivity and sDegree constraint. The necessary notations used in our model are explained in Table 3.

3.1 Communication Model

There are two kinds of nodes in the EHRD problem: the original sensor nodes (SNs) and the energy harvesting relays (EHRs) to be deployed. They have some similar properties. For an SN, its service radius denoting the longest distance of its terminals is d_s and its

Table 2. Summary on Relay Deployment Algorithms with Problem Formulations and Constraints

Algorithm	Problem Formulation	Constraint
[37]	Optimal locations of mobile relay stations maximizing energy efficiency	Spectral efficiency requirement and SINR constraints
[38]	Trade-off among the throughput, the deployment budget, and coverage	Candidate positions, traffic demands
[39]	Deploying different types of RSs in Mobile Multihop Relay (MMR) networks	Network cost, transmission quality, deployment price, coverage
[40]	Establishing an energy dissipation model of multi-hop relay transmission scheme	A given BER requirement, modulation sizes
[41]	If a relay should be deployed and where to deploy in linear underwater acoustic networks	Open distance according to effective bandwidth and transmit power
[42]	Energy efficiency of three basic bidirectional relay transmission schemes from the angle of relay deployment	Should be deployed closer to the base station with the nonideal PA, the impact of small-scale fading

communication radius which indicates the longest distance between two SNs is d_c where $d_c = 2 \times d_s$. RNs, which are EHRs in this paper, have a similar transmission capacity to SNs. The service radius and the communication radius of RNs are d_s and d_c as well.

Table 3. Notation List for System Model

Symbol	Description
d_s	Service radius
d_c	Communication radius
P_e	Harvesting rate
Z_r	Workload of r
$z_{r,n}$	Number of source-destination pairs in the n-th block
C_e	Unit power consumption rate
T_b	Duration time of transmission blocks

For two nodes u and v, the Euclidean distance between u and v is denoted as $dist(u, v)$. For convenience, u and v also denote the locations of the nodes. The communication between nodes in the network is symmetric. Two nodes u and v can communicate with each other if their distance is less than or equal to their minimum communication range.

As shown in Fig.2, in the original WSN, an SN u is connected to another SN v if and only if $dist(u, v) \leqslant d_c$. However, in the new WSN after deployment, SNs all become terminals and only sense their own data, thereby there are no links between two SNs. An SN u can be connected to an RN r, which is an EHR in this paper, if and only if $dist(u, r) \leqslant d_s$. With the decrease of the communication range of SNs from d_c to d_s, the workload of SNs is relieved and energy is saved. For EHRs, an EHR r is connected to another EHR t if and only if $dist(r, t) \leqslant d_c$. The relationships are shown as Fig.3. We use disks with radius d_s to denote the EHRs. An EHR can communicate with SNs within its radius. In this paper, if an SN s and an EHR r are connected, we claim that r covers s and r is responsible for forwarding data for s. Next, we give the definition of sDegree.

Definition 1 (sDegree of EHRs). *For an EHR r, $sDegree(r)$ is the number of SNs that r covers.*

Note that $sDegree(r)$ is likely to be a fraction because there are some back-up EHRs and they will take turns to forward the data for one SN.

3.2 Network Topology

In this subsection, we give the models of the WSNs (before and after the deployment of EHRs), which are two undirected graphs. For the original WSN, the set of SNs S, the communication range d_c and the functional status of SNs, induce a communication graph $SG(V, E)$. After completing the deployment of EHRs, functional status of SNs will change: SNs will only carry out their own sensing tasks. Thus we can induce another communication graph $HG(V, E)$. For the sake of clarity, in this paper we use V and E to represent the vertex set and the edge set respectively. We have the following definitions.

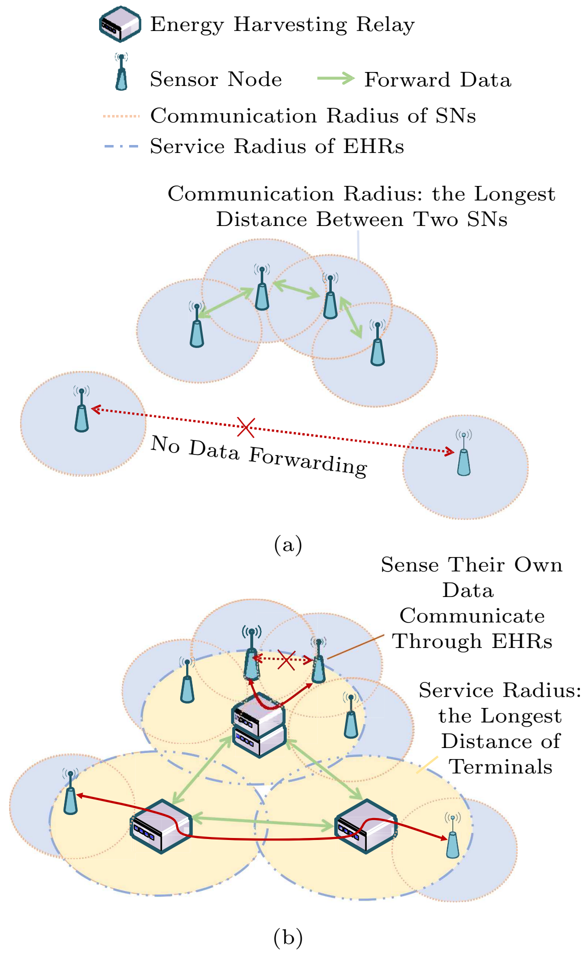


Fig.2. Comparison between (a) traditional WSN and (b) WSN with EHRs.

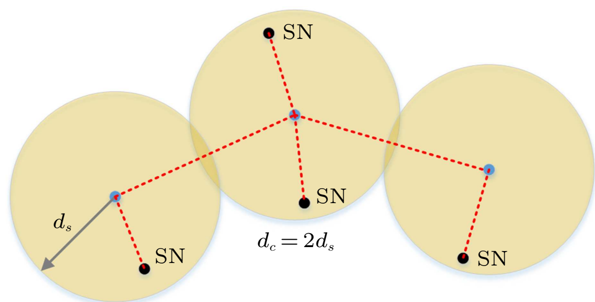


Fig.3. Communication between EHRs and SNs.

Definition 2 (Communication Graph Before Deployment). *Given the SN set S and the communication range d_c, the communication graph of the WSN before the relay deployment, denoted by $SG(V,E)$, is an undirected graph with the vertex set $V(SG) = S$ and the edge set $E(SG)$ consisting of edges that satisfy the following conditions: for two SNs u and v, the edge $(u,v) \in E(SG)$ if $dist(u,v) \leqslant d_c$.*

Definition 3 (Communication Graph After Deployment). *Given the SN set S, the EHR set R, the service range d_s and the communication range d_c, the communication graph of the WSN after the relay deployment, denoted by $HG(V,E)$, is an undirected graph with the vertex set $V(HG) = S \cup R$ and the edge set $E(HG)$ consisting of edges that satisfy the following conditions: for an SN u and an EHR r, the edge $(u,r) \in E(HG)$ if $dist(u,r) \leqslant d_s$ and r is responsible for u. For two EHRs r and t, the edge $(r,t) \in E(HG)$ if $dist(r,t) \leqslant d_c$.*

In the EHRD problem, the original WSN is connected and SG is also connected. After the deployment of EHRs, the new WSN induces another communication graph HG. It is evident that HG should be connected for the reason that the network connectivity in WSNs ensures that the data can be transmitted through the whole network. Therefore, the first requirement of the EHRD problem is connectivity.

3.3 Energy Harvesting Model

In this subsection, we introduce the model of EHRs and reduce the energy casualty constraint to an sDegree restriction. Here we refer to some characteristics of energy harvesting nodes in [33, 54, 55].

We divide time into transmission blocks with duration time T_b. In each transmission block, relay selection is performed [55]. Several relays will be selected to take part in the data transmission process.

We assume that the offline information of all EHRs is available. The harvesting rate donated as P_e is steady while the power consumption rate in a block is dependent on the workload of r. For an EHR r, it can harvest and spend energy simultaneously. Let $\boldsymbol{Z}_r = (z_{r,n})$ be the workload matrix of r, where $z_{r,n}$ is the number of source-destination pairs that r forwards the data for in the n-th block. Therefore, the energy consumption in the n-th transmission block can be expressed as $C(n,r) = z_{r,n}C_{\rm e}T_{\rm b}$, where $C_{\rm e}$ is the unit power consumption rate. Let $E_{\rm o}$ be the initial energy of r. Furthermore, we have the available energy of r after l transmission blocks:

$$E(r, lT_{\rm b}) = E_{\rm o} + \sum_{n=1}^{l} P_{\rm e}T_{\rm b} - \sum_{n=1}^{l} z_{r,n}C_{\rm e}T_{\rm b}.$$

For r, its utilization of harvested energy is constrained by the energy causality constraint [33], which can be expressed as

$$E(r, lT_{\rm b}) \geqslant 0, \quad \forall\, l \in N. \tag{1}$$

This constraint brings significant challenges for designing energy harvesting systems. An EHR will doze off if its energy level is low. In many real-world energy harvesting systems, in order to combat this challenge, back-up or extra devices are indispensable to allow energy harvesting devices to take turns to work as shown in Fig.4. They are able to avoid energy depletion [33, 56]. Moreover, to avoid the energy depletion of EHRs, we first reduce it to an sDegree constraint and then add some back-up EHRs.

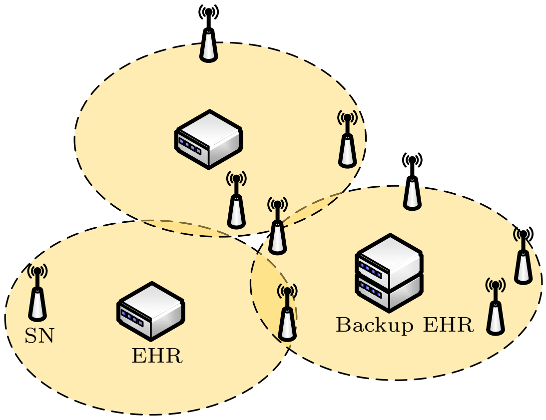


Fig.4. EHRs cover SNs.

EHRs at the same location can be seen as one relay node from the outside. From the interior, they are independent relays that take turns to work. The workload is also averaged out among EHRs and then every EHR will have enough time to absorb energy, avoiding violating the energy causality constraint. If m EHRs located at the same position cover n SNs, the total sDegree is averaged out among these EHRs, and for a specific EHR r, we have

$$sDegree(r) = \frac{n}{m}.$$

Although the energy causality constraint (constraint (1)) is very complicated and dynamic, the constraint can be easily satisfied when an EHR r covers at most $sDegree(r)$ SNs. In this case, we convert the energy causality constraint to a static sDegree constraint which is a sufficient, but not necessary condition to meet the constraint. Here we first give the following simple lemma.

Lemma 1. *We can seek a deployment where an EHR is connected to at most* 20 *EHRs.*

Proof. Let the radius of EHRs be d_s and all EHRs are connected to an EHR r. They can cover a large disk of radius $3d_s$ (see Fig.5). Moreover, since a disk of radius d_s can cover a $\sqrt{2}d_s \times \sqrt{2}d_s$ square, 21 EHRs including r can cover this large disk. If there is a deployment where an EHR is connected to more than 20 EHRs, it is not a reasonable solution since the EHRs are consuming more energy while not providing a larger covered area. As a result, we assume 20 is an upper bound of the number of EHRs connected to r although it may not be a tightest bound. □

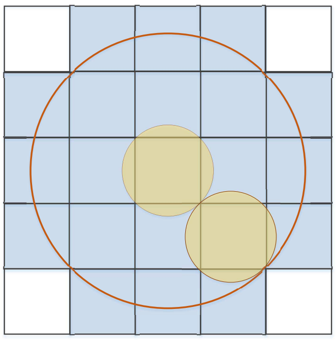


Fig.5. An EHR is connected to at most 20 EHRs.

Then we have the following theorem.

Theorem 1. *The energy causality constraint can be guaranteed by an sDegree constraint:* $\forall r \in R, sDegree(r) \leqslant D$, *where* D *is a constant.*

Proof. For an EHR r, we let k be the number of EHRs at the locations of r and s be the number of SNs within the radius of r. As Lemma 1 shows, an EHR can be connected to at most 20 EHRs. The k EHRs at the location of r take turns to forward the data among their 20 neighbors, where every EHR at the location of r has $\frac{20}{k}$ in average.

Then r is connected to at most $(sDegree(r) + \frac{20}{k})$ nodes including EHRs and SNs. To ensure the energy causality constraint, it suffices to let $\sum_{n=1}^{l} P_e T_b - \sum_{n=1}^{l} z_{r,n} C_e T_b \geqslant 0$. Then we have

$$\sum_{n=1}^{l} z_{r,n} \leqslant \sum_{n=1}^{l} \frac{P_e}{C_e}.$$

Since $z_{r,n}$ is the number of S-D pairs that r forwards the data for in the n-th block, we get

$$z_{r,n} \leqslant \frac{1}{2}\left(sDegree(r) + \frac{20}{k}\right).$$

Since $sDegree(r) = \frac{s}{k} \leqslant D$, then $k \geqslant \frac{s}{D}$, thereby we get

$$\frac{1}{2}\left(\frac{s}{k} + \frac{20}{k}\right) \leqslant \frac{1}{2}\left(D + \frac{20D}{s}\right) \leqslant 11D.$$

We can set D as $\frac{P_e}{11C_e}$, and then the energy causality constraint can thus be guaranteed. □

In the above proof, we just show how to find a feasible D. In fact, $D = \frac{P_e}{11C_e}$ is just a very strict condition to ensure the energy causality constraint. In most cases, with a well-designed relay selection schedule, we can find a D much greater than $\frac{P_e}{11C_e}$ [55]. However, in this paper, we do not focus on the designs of relay selection schedules. We will not analyze how to find the upper-bound of D. We just assume that we have found such a constant D as the sDegree constraint.

So far, we have found the second constraint of the EHRD problem, the sDegree constraint.

4 Problem Formulation

In this section, we give the formal definition of the EHRD problem and then the proof of its NP-hardness.

In the EHRD problem, we aim to seek a deployment of a minimum number of EHRs that covers the whole network. In the meantime, we should guarantee the connectivity and survivability by deploying a connected dominating set. The design objective is to minimize the size of the EHR set R. In Section 3, we have seen that there are two constraints in this problem, connectivity constraint and sDegree constraint. The connectivity constraint guarantees the data transmission through the network and ensures that the data of every SN can be forwarded among the nodes after we complete the deployment while the sDegree constraint guarantees the energy casualty constraint of EHRs. Each EHR can only be connected to at most $sDegree(r)$ SNs and 20 EHRs. The formal definitions are the following.

Definition 4 (Effective Relay Node Deployment). *Given a communication graph* $SG(S, E)$ *of a WSN, the service radius* d_s*, the communication radius* d_c *and an sDegree constraint* D*, a deployment of EHRs* R *is said to be an effective relay deployment (ERD) for SG if the induced communication graph after deployment* $HG(S \cup R, E)$ *is connected and* $\forall r \in R$*,* $sDegree(r) \leqslant D$*. An ERD is said to be a minimum effective relay deployment (M-ERD) if it has the minimum size among all ERDs.*

Definition 5 (Energy Harvesting Relay Deployment). *Given a communication graph* $SG(S, E)$ *of a WSN, the service radius* d_s*, the communication radius*

d_c and an sDegree constraint D, the Energy Harvesting Relay Deployment (EHRD) problem for (SG, d_s, d_c, D), denoted as EHRD(SG,d_s,d_c,D), is to find an M-ERND for SG.

The EHRD problem is an optimization problem with several complicated constraints, and it can be described in the form of a mixed integer nonlinear program. Here we introduce some variables. We use a vector of booleans $\boldsymbol{X} = (X_m)$, which represents the type of each node in V, i.e., $X_m = 1$ if node v_m is an EHR and we let $N(m)$ denote the set of neighboring nodes of v_m. The flow matrix of integers $F_{i,j}$, $(v_i, v_j) \in E$, and the vector $\boldsymbol{Y} = (Y_m)$, $v_i \in N(1)$ are additional variables used to model the connectivity [23]. Then the EHRD problem can then be formulated as

$$\begin{aligned}
\min \quad & |R| \\
\text{s.t.} \quad & sDegree(v_m) \leqslant D, \forall\, v_m \in R, && (2)\\
& X_m + \sum_{j\in N(m)} X_j \geqslant 1, \quad \forall\, v_m \in V, && (3)\\
& \sum_{i\in N(1)} F_{1,i} = \sum_{i\in V, i\neq 1} X_i, && (4)\\
& \sum_{j\in N(i)} F_{j,i} - \sum_{j\in N(i)} F_{i,j} = X_i, \quad \forall\, i \in V,\ i \neq 1, && \\
& 0 \leqslant F_{i,j} \leqslant nX_j, \quad \forall\, (v_i, v_j) \in E,\ j \neq 1, && (5)\\
& \sum_{i\in N(1)} Y_i \leqslant 1 + X_1(|N(1)| - 1), && (6)\\
& F_{i,j} \leqslant nY_i, \quad \forall\, v_i \in N(1), && (7)\\
& F_{i,j} = 0, \quad \forall (v_i, v_j) \in E,\ \text{or}\ j = 1, && (8)
\end{aligned}$$

where constraint (2) is the sDegree constraint based on the energy causality constraint denoted in constraint (1). The remaining constraints guarantee that R is a connected cover of S which is also a connected dominating set of HG. The constraint represented by constraint (3) is to guarantee that v_i is an EHR (X_i=1), or it has a neighbor EHR (there exists at least one neighbor j, X_j=1). For constraint (4)–constraint (8), the principle is to generate a flow, only from vertex v_1. The sum of the flow is equal to the number of EHRs covering the rest WSN, i.e., $\sum X_i$ if v_i is not an EHR (X_1=0), or $\sum X_i - 1$ if v_i is an EHR (X_1=1). In the former case, v_1 should have at least one edge towards an EHR which is a dominating vertice. Every flow goes towards the EHRs while none of the flows go to dominated vertices. The flow fades at every unit as it traverses the EHRs. Since the flows indicate the number of EHRs dominating the rest WSN, they are bounded by 0 and n represented by constraint (5). The binary vector, $\boldsymbol{Y}$, indicates whether the neighbors of node v_1 can be traversed, i.e., $Y_i = 1$ if the flow is permitted from v_1 to v_i ensured by constraint (7). Constraint (6) lets $Y_i = 1$ when v_1 is not an EHR; otherwise it is bounded by the number of its neighbors. Finally, it is guaranteed by constraint (8) that the flow travels only through existing edges and no flow enters v_1. Then the data can be forwarded exclusively through EHRs. Since the EHRD problem can be formed as a mixed integer nonlinear programming problem, it is an NP-hard problem [57]. Next, we will prove this in a more intuitive way.

Theorem 2. *The EHRD problem is NP-hard.*

Proof. We can reduce the EHRD problem from the unit disk covering (UDC) problem, which has been proved to be NP-hard [58]. Given n points in the Euclidean plane, we find a minimum number of unit disks to cover all given points. A unit disk is a disk with diameter 1.

Suppose we have an algorithm A_d to compute the EHRD problem. For any instance of the UDC problem, we can construct the same graph. The service radius can be set as $1/2$. The sDegree constraint D can be set as a large number which is bigger than the number of SNs $|S|$, and the communication radius d_c of EHRs is large so that any two relays in the plane are always connected to each other. Then we get an instance for the EHRD problem, which can be computed by A_d. The output can be seen as the solution to the UDC problem as well. Therefore, according to the Cook's reduction, the EHRD problem is NP-hard. □

For the talking convenience, in Sections 5–7, we let the communication radius $d_c = 1$ and the service radius $d_s = \frac{1}{2}$. Hence, an EHR is a unit disk with diameter 1.

5 Effective Relay Deployment Algorithm

In this section, we present a polynomial-time algorithm for the EHRD problem. To find an ERND of the WSN, the Effective Relay Deployment Algorithm (ERDA) consists of two phases. The first phase computes an sDegree constraint disk cover (DC-DC) of SG, while the second phase connects the DC-DC computed in the first phase. The final ERND is the union of the results of the two phases. Fig.6 shows how our algorithm solves the constraints of the EHRD problem and the steps of the solution.

5.1 Computing a DC-DC

In this phase, we compute a DC-DC of SG with the partitioning technique. First of all, we divide the plane

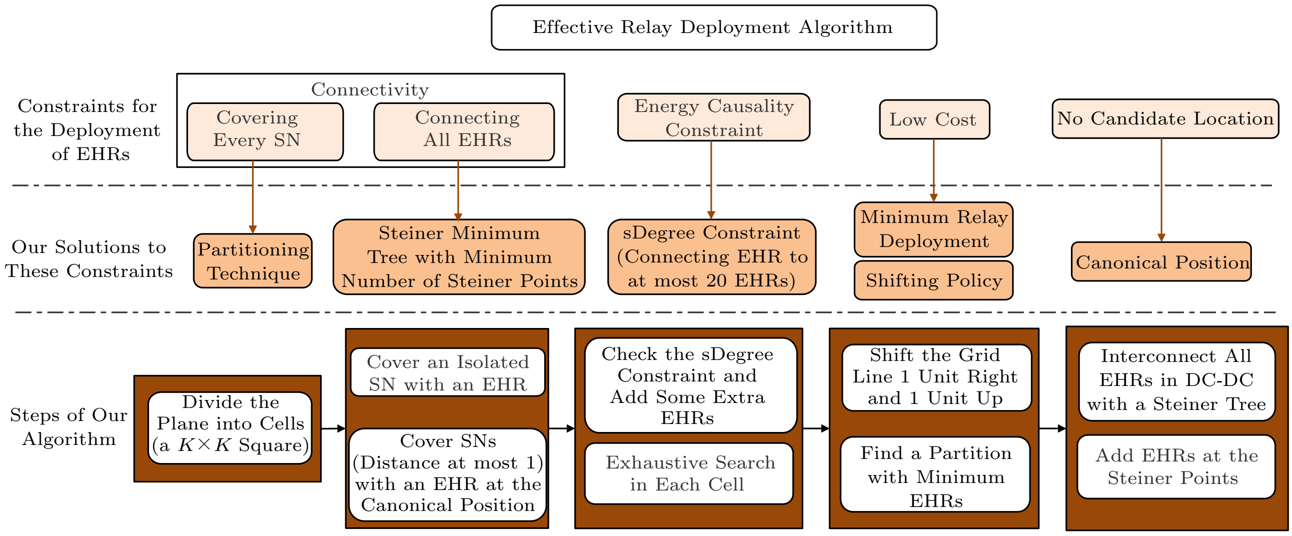


Fig.6. Explanation for our proposed Effective Relay Deployment Algorithm.

Q where SG lies into small squares, called cells. For simplicity, a cell is a $K \times K$ square for some constant K, excluding the top and right boundaries as shown in Fig.7. Then we pick those nonempty cells that contain SNs and ignore the remaining ones. Next, we solve the problem DC-DC for each cell. The solution for a single cell also contains two parts. First, we compute a disk cover in the cell with an exhaustive search algorithm. Second, we check the sDegree constraint and add some extra disks if necessary. Finally, we take the union of the solutions of all cells to form a DC-DC $\boldsymbol{A}$ of the whole graph SG.

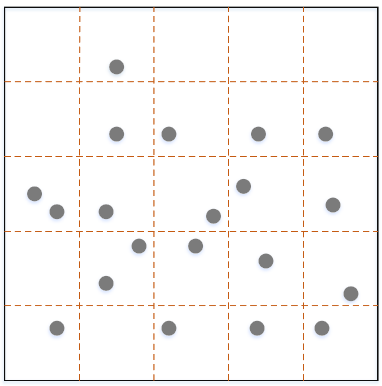


Fig.7. Partitioning the plane into cells.

The pseudo-code of the algorithm for computing a DC-DC is presented in the following.

Here we briefly explain why a minimum disk cover in a cell can be computed in polynomial time $n^{O(K^2)}$ [59], where n is the number of SNs. Since a unit disk can cover a $\frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2}$ square, a cell of size $K \times K$ can be covered with at most $\lceil \sqrt{2}K \rceil^2$ disks. For a cell e, we can find a minimum disk cover with an exhaustive search algorithm. If an SN has a distance greater than 1 from any other SN, we have to cover it with an isolated disk. If an SN u has a distance at most 1 from some other SN v, then we can use one disk to cover u and v. In this case, we can move this disk to a canonical position of u and v so that this disk is more likely to cover other SNs. For u and v, their canonical positions are the locations of disks ensuring both u and v lie on their boundary. There are two possible canonical positions for u and v as showed in Fig.8. Hence, there are at most $\binom{n_e}{2}$ canonical positions in e, where n_e is the number of SNs in e. Together with the observation that e can be covered by at most $\lceil \sqrt{2}K \rceil^2$ disks, we see that, in the exhaustive search algorithm, we need to inspect at most $\binom{n_e}{2}^{\lceil \sqrt{2}K \rceil^2} = n_e^{O(m^2)}$ solutions to find a minimum disk cover in a cell.

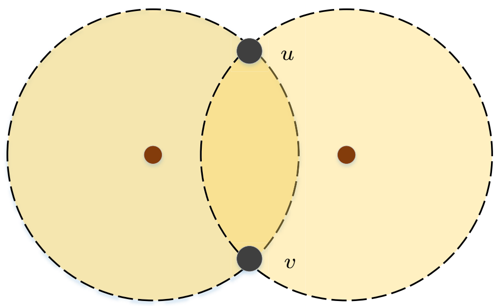


Fig.8. Canonical positions of u and v.

For the SNs in a single cell, we first find all canonical positions of these SNs. Then we run an exhaustive search algorithm according to these positions and seek a minimum disk cover. Since we partition a large scale of SNs to small cells which contain only a part of the nodes, the amount of calculation and time will be reduced.

5.2 Improvement Based on Shifting Policy

The performance of Algorithm 1 is closely related to the partition. Thus we can use a shifting strategy to move grid lines to find a partition with a small number of disks intersecting with grid lines. We move every grid line 1 unit right and 1 unit up to its original positions Q, which can be seen in Fig.9. Then we run Algorithm 1 to get a DC-DC of SG. Since the sidelength of each cell is K, we do this process $(K-1)$ times and choose the minimum DC-DC $\boldsymbol{A}$ as the result of phase 1.

Algorithm 1. Computing the DC-DC

Input: a communication graph $SG(S, E)$, lying in square Q; an sDegree constraint D; an integer $K \geqslant 1$
Output: a DC-DC of SG

```
Divide Q into cells, each of size K × K;
cell(Q) ← the set of all nonempty cells in Q;
for e ∈ cell(Q) do
    Find a minimum unit disk cover A(e) for nodes in e
    for a ∈ A(e) do
        if sDegree(a) > D then
            Add (⌈C(a)/D⌉ − 1) disks into A(e) where a is
            located;
return A = ∪_{e∈cell(Q)} A(e)
```

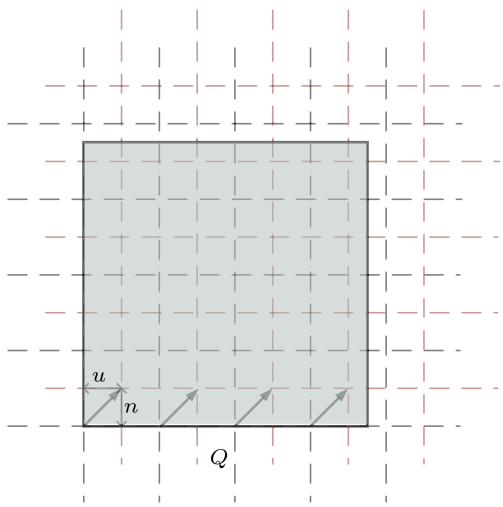


Fig.9. Shifting policy.

5.3 Connecting a DC-DC

After Algorithm 1, we get a DC-DC $\boldsymbol{A}$ of SG. However, this DC-DC is not necessarily connected. In the following step, we will add some extra EHRs to this DC-DC to form a feasible ERND.

Here we connect the DC-DC with a Steiner minimum tree with minimum number of Steiner points (SMT-MSP). In this paper, we adopt a 3-approximation algorithm proposed in [44]. In details, a Steiner tree interconnecting all terminals with the minimum number of Steiner points is proposed such that the Euclidean length of each edge is no more than a constant according to the given n terminals in the plane and a positive constant. In addition, the minimum spanning tree yields a polynomial-time approximation with the performance ratio of exactly 4 and there exists a polynomial-time approximation scheme under certain conditions. The pseudo-code of the algorithm for connecting the DC-DC is presented in Algorithm 2.

Algorithm 2. Connecting the DC-DC

Input: a communication graph SG; a DC-DC $\boldsymbol{A}$ of SG; a 3-approximation algorithm $\hbar$ computing a Steiner tree with the minimum number of Steiner points
Output: an ERND of SG

```
Use ħ to compute a Steiner tree T with input A;
P ← the Steiner points of T;
Add unit disks at the positions of each point in P;
C ← the set of added disks;
return A ∪ C;
```

Although the pseudo-code seems simple, what Algorithm 2 solves is indeed another challenging NP-hard problem [43]. Here we briefly introduce the principle of the 3-approximation algorithm in [44]. This algorithm studies the geometric property of nodes in the planes and classifies the nodes into three different categories. Then it puts Steiner points near every node in a way depending on its category.

5.4 Ensuring sDegree Constraint

After connecting the DC-DC of SG in Algorithm 2, the output ERND is a connected dominating set of SG which can cover all SNs and forward the data for them to save the limited energy. However, we have to guarantee that each EHR we deploy in WSN is connected to at most $sDegree(r)$ SNs and 20 EHRs. For those EHRs whose neighbors exceed the bound, we need to add back-up EHRs that take turns to forward data.

6 Performance Analysis

6.1 Analysis of Computational Complexity

In this subsection, we analyze the computational complexity of ERDA. Table 4 shows the notations needed for performance analysis.

Table 4. Notation List of Performance Analysis

Symbol	Description
$A^*(e)$	Optimal solution to the DC-DC problem
$A_o(e)$	Optimal solution to the disk cover problem
N	Set of disks satisfying the sDegree constraint
O	Disks which cover too many nodes
P_o	Original partition
P_n	Partition after n shifts
$B^*(e)$	Set of disks that intersects cell e
H	Set of disks that intersects two horizontal strips
V	Set of disks that intersects two vertical strips
Opt^*	A minimum effective relay node deployment
C^*	A Steiner tree with the minimum number of Steiner points

Theorem 3. *The ERDA algorithm runs in polynomial time.*

Proof. To see that our algorithm runs in polynomial time, we claim that both the two phases run in polynomial time. This claim is obvious in the second phase because Algorithm 2 is an approximation algorithm whose computational complexity is $O(n^4)$ [44].

The case in the first phase is a little complex. To see that Algorithm 1 runs in polynomial time, we need to illustrate that the operation in each cell can be completed in polynomial time. Let e be a single cell containing n_e SNs and its side length be a constant K. To find a DC-DC of e, we first compute a minimum disk cover and then we check the sDegree constraint. According to Section 5, we have known that the minimum disk cover restricted to e can be computed in time $n_e^{O(K^2)}$. Second, to meet the sDegree constraint, we check whether every disk covers too many nodes. There are $O(K^2)$ disks in the cell to check. Checking one disk is composed of two steps: computing the number of covering nodes and adding backup disks, which can be finished in linear time $O(n_e)$. Thus the computational time of checking all disks in e is $O(K^2) \times O(n_e) = O(n_e)$. Furthermore, the operation in a single cell can be completed in polynomial time which is

$$n_e^{O(K^2)} + O(n_e) = n_e^{O(K^2)}.$$

Thus, over all nonempty cells, the total time of the loop from line 3 to line 7 in Algorithm 1 is

$$\sum_{e \in cell(e)} n_e^{O(K^2)} \leqslant \Big(\sum_{e \in cell(e)} n_e\Big)^{O(K^2)} = n^{O(K^2)}.$$

It is easy to see that Algorithm 1 runs in polynomial time. Combining the two phases, we see that ERDA is a polynomial-time algorithm. □

6.2 Performance Ratio

In this subsection, we give the performance ratio of ERDA. First, we look at the performance ratio of DC-DC $A(e)$ in a single cell e. Let $A^*(e)$ be an optimal solution to the DC-DC problem in e and $A_o(e)$ be an optimal solution to the disk cover problem in e. The number of nodes in e is n_e.

Theorem 4. *The size of DC-DC $A(e)$ in the cell e obtained by Algorithm 1 is within 2 times of the optimum one.*

Proof. In the cell e, we first obtain an optimal disk cover $A_o(e)$. Then we divide $A_o(e)$ into two sets according to sDergee, N and O. N is the set of disks satisfying the sDegree constraint while O is the remaining disks which cover too many nodes.

According to the operation in Algorithm 1, the DC-DC $A(e)$ follows that

$$\begin{aligned}|A(e)| &\leqslant |A_o(e)| + \sum_{a \in O}\Big(\Big\lceil\frac{sDegree(a) - D}{D}\Big\rceil\Big)\\ &\leqslant |A_o(e)| + \sum_{a \in O}\frac{sDegree(a)}{D}\\ &\leqslant |A_o(e)| + \frac{n_e}{D}.\end{aligned}$$

Meanwhile, $A^*(e)$ is a disk cover in e and it follows the sDegree constraint. Thus we have

$$|A^*(e)| \geqslant |A_o(e)| \text{ and } |A^*(e)| \geqslant \frac{n_e}{D}.$$

Then we finally get

$$|A(e)| \leqslant 2|A^*(e)|.$$ □

We have proved that the approximation ratio of DC-DC in a single cell is 2. Next, we analyze the approximation ratio of DC-DC $\boldsymbol{A}$ computed by Algorithm 1 of the whole WSN. Following the general approach of the analysis of approximation algorithms designed by the restriction method, we consider an optimal solution $\boldsymbol{A}^*$ in the first phase and modify it to a feasible solution. Subsequently, we compare this feasible solution with $\boldsymbol{A}$ obtained by Algorithm 1.

Theorem 5. *Algorithm 1 always outputs a DC-DC with its number of disks within eight times of the optimum one.*

Proof. For each disk in $\boldsymbol{A}^*$ that intersects more than one cells, we make additional copies of the disk and use them to cover points in different cells. If a disk intersects k cells, where $2 \leqslant k \leqslant 4$, then we make $k-1$ additional copies. Each copy is used to cover points in a different cell. Note that one unit disk can intersect at most four cells, and the modification will get a feasible solution F, where $|F| \leqslant 4|\boldsymbol{A}^*|$.

Now we can prove that the solution $\boldsymbol{A}$ obtained by Algorithm 1 satisfies that $|\boldsymbol{A}^*| < 2|\boldsymbol{F}|$. Since disks in F in a cell e are a DC-DC of e, F has a lower bound $\cup_{e\in cell(Q)} A^*(e)$. According to Theorem 5, we get

$$\begin{aligned}|\boldsymbol{A}| &= \cup_{e\in cell(Q)}\, A(e)\\ &\leqslant \cup_{e\in cell(Q)}\, 2A^*(e)\\ &\leqslant 2|\boldsymbol{F}|.\end{aligned}$$

Hence, we have

$$|\boldsymbol{A}| \leqslant 2|\boldsymbol{F}| \leqslant 8|\boldsymbol{A}^*|. \qquad \square$$

Here we will see that the shifting policy greatly improves the results of phase 1.

Theorem 6. *The computed DC-DC after $K-1$ shifts is $(2+6/K)$-approximation of the optimal solution.*

Proof. Let P_0 be the original partition and P_n be the partition after n shifts. For a partition P_a where $0 \leqslant a \leqslant K-1$, the DC-DC of SG is denoted by $\boldsymbol{A}_a$ and the DC-DC in cell e is denoted by $A_a(e)$. Let $B_a^*(e)$ denote the set of disks in the optimal $\boldsymbol{A}^*$ that intersects cell e in P_a. Let the collection of all cells in P_a that lie along a horizontal or vertical line be a strip. Let H_a and V_a be two sets of disks in $\boldsymbol{A}^*$. A disk $d \in H_a(V_a)$ if d intersects two horizontal (or vertical) strips. We know that a disk intersects more than two cells, and it must belong to both H_a and V_a. Then we have

$$|\boldsymbol{A}_a| = \cup_{e\in cell(Q_a)} A_a(e) \leqslant 2\ \cup_{e\in Cell(Q_a)} |B_a^*(e)|, \quad (9)$$

and

$$\cup_{e\in Cell(Q_a)} |B_a^*(e)| \leqslant |\boldsymbol{A}^*| + |H_a| + |V_a| + |V_a|. \quad (10)$$

Note that a cell is excluding its right and top boundaries and the diameter of the disk is 1, and for partition P_a and P_b $(a \neq b)$, a disk cannot be in both H_a (or V_a) and H_b (V_b). Thus,

$$\sum_{i=0}^{K-1} |H_i| \leqslant |\boldsymbol{A}^*| \quad \text{and} \quad \sum_{i=0}^{K-1} |V_i| \leqslant |\boldsymbol{A}^*|. \quad (11)$$

Therefore, combining inequalities (9), (10) and (11), we have

$$\sum_{i=0}^{K-1} |\boldsymbol{A}_a| \leqslant 2\sum_{i=0}^{K-1} (|\boldsymbol{A}^*| + |H_a| + 2|V_a|) \leqslant 2(K+3)|\boldsymbol{A}^*|. \quad (12)$$

Hence,

$$\frac{1}{K}\sum_{i=0}^{K-1} |\boldsymbol{A}_a| \leqslant (2+\frac{6}{K})|\boldsymbol{A}^*|.$$

According to the probabilistic method, we know that there exists a value of $a \in \{0, 1, ..., K-1\}$ that $|\boldsymbol{A}_a| \leqslant (2+\frac{6}{K})|\boldsymbol{A}^*|$, the DC-DC after $(K-1)$ shifts is a $(2+6/K)$-approximation of the optimal DC-DC. $\square$

Combining the two phases, we can get the solution to the EHRD problem. Next, we analyze its performance ratio.

Theorem 7. *The performance ratio of ERDA is* $(5+\frac{6}{K})$.

Proof. Let $\boldsymbol{Opt}^*$, which is also a connected DC-DC, be a minimum effective relay node deployment (M-ERND), and $\boldsymbol{C}^*$ be a Steiner tree connecting $\boldsymbol{A}$ with a minimum number of Steiner points. The solution obtained by our algorithm is $\boldsymbol{A} \cup \boldsymbol{C}$. Because $\boldsymbol{Opt}^*$ is also DC-DC, we have

$$|\boldsymbol{A}| \leqslant (2+\frac{6}{K})|\boldsymbol{A}^*| \leqslant (2+\frac{6}{K})|\boldsymbol{Opt}^*|. \quad (13)$$

Let d be a disk in $\boldsymbol{A}$. Suppose v is a node covered by d, and d^* is a disk in $\boldsymbol{Opt}^*$ covering v. Then we know that $dist(d, d^*) \leqslant 1$ (d_c). Hence, $\forall\ d \in \boldsymbol{A}$, $\exists d^* \in \boldsymbol{Opt}^*$, and d is connected to d^*. In other words, $\boldsymbol{Opt}^*$ will be a feasible solution of the second phase. Then we have

$$|\boldsymbol{C}| \leqslant 3|\boldsymbol{C}^*| \leqslant 3|\boldsymbol{Opt}^*|. \quad (14)$$

Combining inequalities (13) and (14), we have

$$|\boldsymbol{A} \cup \boldsymbol{C}| \leqslant |\boldsymbol{A}| + |\boldsymbol{C}| \leqslant (5+\frac{6}{K})|\boldsymbol{Opt}^*|.$$

Therefore, the performance ratio of the whole algorithm is $5+\frac{6}{K}$. $\square$

7 Simulation

In this section, we use Python to conduct some simulations and evaluate the performance of ERDA. In our simulations, we randomly deploy sensor nodes in a 2D virtual space. Each two nodes can be connected to each other when the distance between them is at most 1. Moreover, we also ensure that the generated graph is connected. Since the initial deployment of SNs is random, we run the algorithm for 10 times at each following plot with different initial SNs and the results are presented with 90% confidence interval.

Fig.10 is a comparison among the solutions obtained by ERDA with different values of K. We run ERDA on the same graph with $K = 1$, 2 and 3 respectively. The number of SNs varies from 50 to 1 000 at intervals of 50. The sDegree constraint is 5.

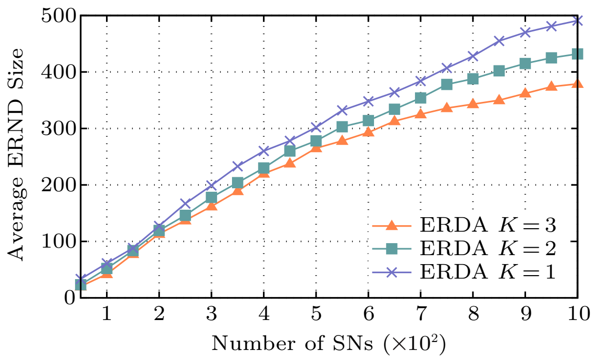


Fig.10. ERDA with different values of K.

From Fig. 10, it is clear that the case $K = 3$ has the best performance while the performance of the case $K = 1$ is comparatively poor. This case is reasonable in that the approximation ratio of ERDA is $5+6/K$ and a larger K will increase the probability of finding a good solution. Meantime, we also observe that the running time of ERDA will increase as K increases during the process of the simulation.

To confirm the effectiveness and efficiency of our solution, we compare ERDA with some other algorithms. Given that no previous work uses a similar model, we compare our algorithm with: 1) a solution based on [23] (MCDS), and 2) a simple greedy deployment solution (GDS) which directly deploys an ERN at the location of the nearest uncovered SN.

A set of comparisons between running results of ERDA, MCDS and GDS are shown in Fig. 11. The inputs are WSNs with 10, 50, 100, 200 and 500 SNs. According to Fig.11, ERDA always adds the least number of EHRs. It is easy to confirm that the ERND of ERDA is an optimal deployment.

Fig. 12 shows the average ERND size against the number of SNs. The number of nodes varies from 50 to 1 000 at intervals of 50. The sDegree constraint D is set to 5 and the sidelength K of cells is set to 2. It is clear that the difference among ERDA, MCDS and GDS increases with the number of nodes. To confirm the better performance of ERDA in a larger scale network, we run the three algorithms on samples with a greater number of SNs. The increase in SNs inevitably causes an increase in the number of added EHRs, but ERDA remains reasonable and does not exceed 500 nodes on average for the scenario with 1 000 nodes, which shows

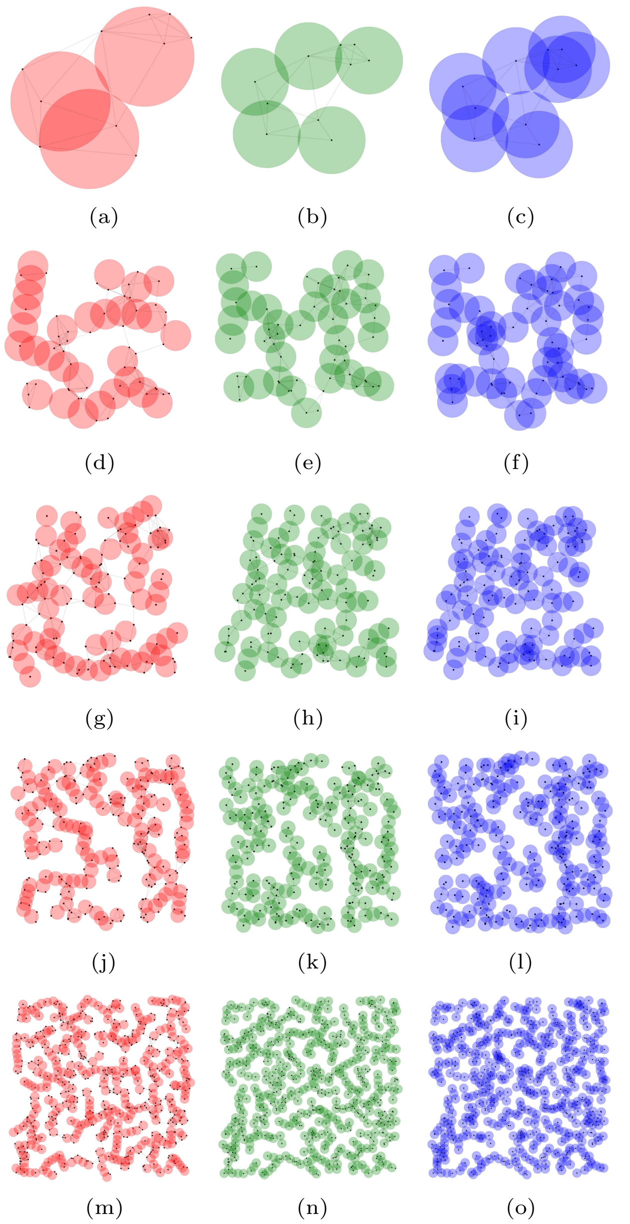


Fig.11. Samples of WSNs with different numbers of SNs. (a) Deployments of ERDA with 10 SNs. (b) Deployments of MCDS with 10 SNs. (c) Deployments of GDS with 10 SNs. (d) Deployments of ERDA with 50 SNs. (e) Deployments of MCDS with 50 SNs. (f) Deployments of GDS with 50 SNs. (g) Deployments of ERDA with 100 SNs. (h) Deployments of MCDS with 100 SNs. (i) Deployments of GDS with 100 SNs. (j) Deployments of ERDA with 200 SNs. (k) Deployments of MCDS with 200 SNs. (l) Deployments of GDS with 200 SNs. (m) Deployments of ERDA with 500 SNs. (n) Deployments of MCDS with 500 SNs. (o) Deployments of GDS with 500 SNs.

the great superiority of ERDA. With ERDA, the number of EHRs deployed can be reduced by up to 45% compared with MCDS and up to 29% compared with GDS. Fig. 13 shows EHRs of ERDA when the input numbers of SNs are 800, 900 and 1 000 respectively. The numbers of added EHRs do not show a large increase with the addition of SNs which can also be seen in Fig.12.

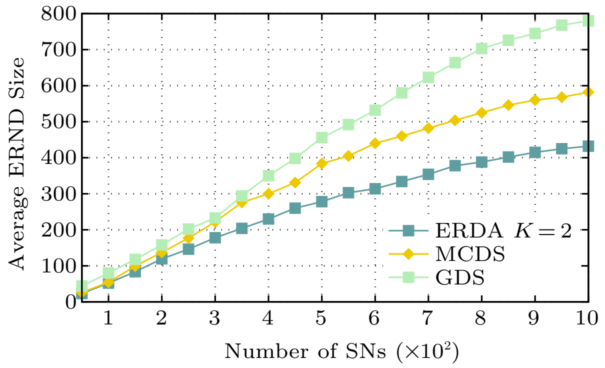


Fig. 12. Average numbers of EHRs deployed compared with MCDS and GDS.

8 Conclusions

In this work, we investigated the energy harvesting relay deployment (EHRD) problem in the conventional wireless sensor network, aiming to upgrade the conventional WSN with the advantage of energy harvesting capacities. This is the first of its type, to our knowledge, on energy harvesting relay deployment in a conventional WSN without candidate locations, which is both theoretically and practically significant in this area. We reduced the EHRD problem to a constrained disk cover problem and proposed an approximation algorithm named Effective Relay Deployment Algorithm (ERDA) to resolve this issue. ERDA consists of two phases: disk covering and connector insertion. In these two phases, ERDA utilizes the partitioning technique and the Steinerization technique respectively. It is general and easy to incorporate future development in algorithm designs for the SMT-MSP problem. To evaluate the performance of ERDA, we proved its performance ratio is $(5 + 6/K)$, where K is an integer denoting the side length of a cell after partitioning. Our simulation results also exhibited the outstanding performance of ERDA. To guarantee the survivability and reliability of the WSN, we converted the energy causality constraint to a static sDegree constraint ensuring every EHR will have enough time to absorb energy.

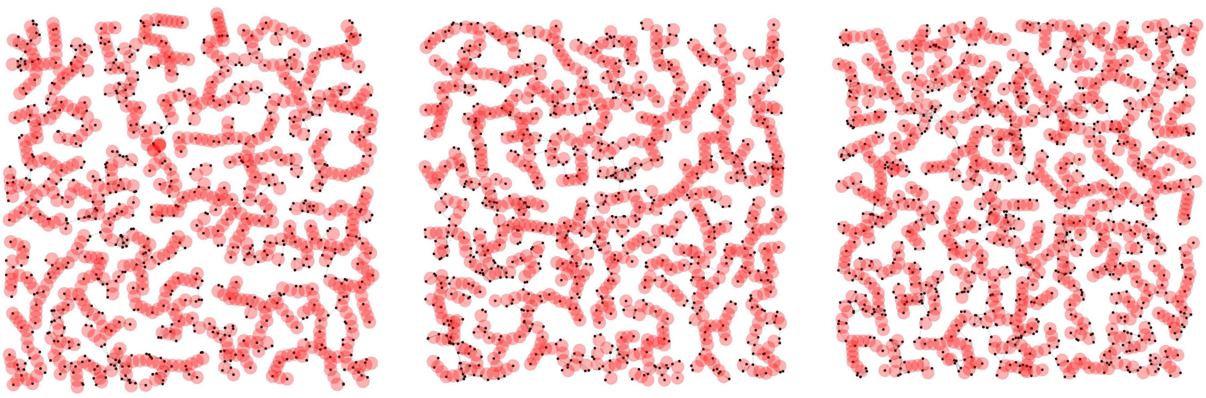
Acknowledgement(s) The authors would like to thank anonymous reviewers for their valuable comments.

References

[1] Wahyono I, Asfani K, Mohamad M, Rosyid H, Afandi A, Aripriharta. The new intelligent wireless sensor network using artificial intelligence for building fire disasters. In *Proc. the 3rd International Conference on Vocational Education and Electrical Engineering*, October 2020. DOI: 10.1109/ICVEE50212.2020.9243210.

[2] Feng S, Shi H, Huang L, Shen S, Yu S, Peng H, Wu C. Unknown hostile environment-oriented autonomous WSN deployment using a mobile robot. *Journal of Network and Computer Applications*, 2021, 182: Article No. 103053. DOI: 10.1016/j.jnca.2021.103053.

[3] Seo J, Park H. Forest environment monitoring application of intelligence embedded based on wireless sensor networks. *KSII Transactions on Internet and Information Systems*, 2016, 10(4): 1555-1570. DOI: 10.3837/tiis.2016.04.005.



(a) (b) (c)

Fig.13. ERNDs by ERDA with increasing number of SNs. (a) Deployment of ERDA with 800 SNs. (b) Deployment of ERDA with 900 SNs. (c) Deployment of ERDA with 1 000 SNs.

[4] Zhou Y, Liu L, Wang L, Hui N, Cui X, Wu J, Peng Y, Qi Y, Xing C. Service-aware 6G: An intelligent and open network based on the convergence of communication, computing and caching. *Digital Communications and Networks*, 2020, 6(3): 253-260. DOI: 10.1016/j.dcan.2020.05.003.

[5] Zhou Y, Tian L, Liu L, Qi Y. Fog computing enabled future mobile communication networks: A convergence of communication and computing. *IEEE Communications Magazine*, 2019, 57(5): 20-27. DOI: 10.1109/MCOM.2019.1800235.

[6] Zhou Y, Liu H, Pan Z, Tian L, Shi J. Cooperative multicast with location aware distributed mobile relay selection: Performance analysis and optimized design. *IEEE Transactions on Vehicular Technology,* 2017, 66(9): 8291-8302. DOI: 10.1109/TVT.2017.2682903.

[7] Hou Y, Shi Y, Sherali H, Midkiff S. Prolonging sensor network lifetime with energy provisioning and relay node placement. In *Proc. the 2nd Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, September 2005, pp.295-304. DOI: 10.1109/SAHCN.2005.1557084.

[8] Rahman M N, Matin M A. Efficient algorithm for prolonging network lifetime of wireless sensor networks. *Tsinghua Science and Technology*, 2011, 16(6): 561-568. DOI: 10.1016/S1007-0214(11)70075-X.

[9] Zhang W, Xue G, Misra S. Fault-tolerant relay node placement in wireless sensor networks: Problems and algorithms. In *Proc. the 26th IEEE International Conference on Computer Communications*, May 2007, pp.1649-1657. DOI: 10.1109/INFCOM.2007.193.

[10] Ku M L, Li W, Chen Y, Ray Liu K J. Advances in energy harvesting communications: Past, present, and future challenges. *IEEE Communications Surveys Tutorials*, 2016, 18(2): 1384-1412. DOI: 10.1109/COMST.2015.2497324.

[11] Atapattu S, Jiang H, Evans J, Tellambura C. Time-switching energy harvesting in relay networks. In *Proc. the IEEE International Conference on Communications*, June 2015, pp.5416-5421. DOI: 10.1109/ICC.2015.7249185.

[12] Ji B, Chen Z, Li Y, Chen S, Zheng G, Wen H. Energy harvesting and information transmission scheme with UAV relay cooperation. *EURASIP Journal on Wireless Communications and Networking*, 2019: Article No. 278. DOI: 10.1186/s13638-019-1598-7.

[13] Sharma S, Roy S D, Kundu S. Secure communication with energy harvesting multiple half-duplex DF relays assisted with jamming. *Wireless Networks*, 2020, 26(2): 1151-1164. DOI: 10.1007/s11276-018-1859-0.

[14] Raghunathan V, Kansal A, Hsu J, Friedman J, Srivastava M B. Design considerations for solar energy harvesting wireless embedded systems. In *Proc. the International Symposium on Information Processing in Sensor Networks*, April 2005, pp.457-462. DOI: 10.1109/IPSN.2005.1440973.

[15] Meninger S, Mur-Miranda J O, Amirtharajah R, Chandrakasan A, Lang J H. Vibration-to-electric energy conversion. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2001, 9(1): 64-76. DOI: 10.1109/92.920820.

[16] He S, Wang W. Location-aware QoE-driven wireless relay deployment for energy efficient multimedia communications. In *Proc. the 17th International Symposium on Web and Wireless Geographical Information Systems*, May 2019, pp.160-173. DOI: 10.1007/978-3-030-17246-613.

[17] Zhao M, Zhao J, Zhou W, Zhu J, Zhang S. Energy efficiency optimization in relay-assisted networks with energy harvesting relay constraints. *China Communications*, 2015, 12(2): 84-94. DOI: 10.1109/CC.2015.7084404.

[18] In C, Kim H, Choi W. Achievable rate-energy region in two-way decode-and-forward energy harvesting relay systems. *IEEE Transactions on Communications*, 2019, 67(6): 3923-3935. DOI: 10.1109/TCOMM.2019.2901783.

[19] Bhowmick A, Chatterjee A, Verma T. Performance of DF relaying in an energy harvesting full duplex cognitive radio network. In *Proc. the International Conference on Vision Towards Emerging Trends in Communication and Networking*, March 2019. DOI: 10.1109/ViTECoN.2019.8899470.

[20] Rao Y S, Madhukumar A S, Sirigina R P. RF energy harvesting based relays: Non-orthogonal AF vs dynamic DF. In *Proc. the 20th International Symposium on Wireless Personal Multimedia Communications*, December 2017, pp.506-510. DOI: 10.1109/WPMC.2017.8301865.

[21] Jiang D, Zheng H, Tang D, Tang Y. Relay selection and power allocation for cognitive energy harvesting two-way relaying networks. In *Proc. the 5th IEEE International Conference on Electronics Information and Emergency Communication*, May 2015, pp.163-166. DOI: 10.1109/ICEIEC.2015.7284511.

[22] Song X, Xu S. Joint optimal power allocation and relay selection in full-duplex energy harvesting relay networks. In *Proc. the 10th International Conference on Communication Software and Networks*, July 2018, pp.80-84. DOI: 10.1109/ICCSN.2018.8488216.

[23] Djenouri D, Bagaa M. Energy harvesting aware relay node addition for power-efficient coverage in wireless sensor networks. In *Proc. the 2015 IEEE International Conference on Communications*, June 2015, pp.86-91. DOI: 10.1109/ICC.2015.7248303.

[24] Zhu X, Li J, Zhou M. Optimal deployment of energy-harvesting directional sensor networks for target coverage. *IEEE Systems Journal*, 2019, 13(1): 377-388. DOI: 10.1109/JSYST.2018.2820085.

[25] Qiao G, Leng S, Zhang K, Yang K. Joint deployment and mobility management of energy harvesting small cells in heterogeneous networks. *IEEE Access*, 2017, 5: 183-196. DOI: 10.1109/ACCESS.2016.2635878.

[26] Verdonck L, Beullens P, Caris A, Ramaekers K, Janssens G K. Analysis of collaborative savings and cost allocation techniques for the cooperative carrier facility location problem. *The Journal of the Operational Research Society*, 2016, 67(6): 853-871. DOI: 10.1057/jors.2015.106.

[27] Wu S, Yang J, Peng R, Zhai Q. Optimal design of facility allocation and maintenance strategy for a cellular network. *Reliability Engineering System Safety*, 2021, 205: Article No. 107253. DOI: 10.1016/j.ress.2020.107253.

[28] Garcia V, Zhou Y, Shi J. Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering. *IEEE Transactions on Wireless Communications*, 2014, 13(8): 4297-4308. DOI: 10.1109/TWC.2014.2316500.

[29] Basappa M, Acharyya R, Das G K. Unit disk cover problem in 2D. *Journal of Discrete Algorithms*, 2015, 33: 193-201. DOI: 10.1016/j.jda.2015.05.002.

[30] Biniaz A, Liu P, Maheshwari A, Smid M. Approximation algorithms for the unit disk cover problem in 2D and 3D. *Computational Geometry: Theory and Applications*, 2017, 60: 8-18. DOI: 10.1016/j.comgeo.2016.04.002.

[31] Zhou Y, Liu H, Pan Z, Tian L, Shi J, Yang G. Two-stage cooperative multicast transmission with optimized power consumption and guaranteed coverage. *IEEE Journal on Selected Areas in Communications*, 2014, 32(2): 274-284. DOI: 10.1109/JSAC.2014.141208.

[32] Misra S, Majd N E, Huang H. Approximation algorithms for constrained relay node placement in energy harvesting wireless sensor networks. *IEEE Transactions on Computers*, 2014, 63(12): 2933-2947. DOI: 10.1109/TC.2013.171.

[33] Luo Y, Zhang J, Letaief K B. Optimal scheduling and power allocation for two-hop energy harvesting communication systems. *IEEE Transactions on Wireless Communications*, 2013, 12(9): 4729-4741. DOI: 10.1109/TW.2013.081413.122021.

[34] Misra S, Majd N E, Huang H. Constrained relay node placement in energy harvesting wireless sensor networks. In *Proc. the 8th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems*, October 2011, pp.25-34. DOI: 10.1109/MASS.2011.137.

[35] Xu S, Shen H. The fault-tolerant facility allocation problem. In *Proc. the 20th International Symposium on Algorithms and Computation*, December 2009, pp.689-698. DOI: 10.1007/978-3-642-10631-6_70.

[36] Tsao Y C, Mangotra D, Lu J C, Dong M. A continuous approximation approach for the integrated facility-inventory allocation problem. *European Journal of Operational Research*, 2012, 222(2): 216-228. DOI: 10.1016/j.ejor.2012.04.033.

[37] Chen W, Chen H, Zhao F. Energy and spectral efficient mobile relay deployment in relay-aided cellular networks. In *Proc. the 2015 International Conference on Wireless Communications & Signal Processing*, October 2015. DOI: 10.1109/WCSP.2015.7340992.

[38] Chang J, Chen Y. A relay station deployment scheme with a rotational clustering algorithm for multi-hop relay networks. In *Proc. the 2016 International Conference on System Science and Engineering*, July 2016. DOI: 10.1109/ICSSE.2016.7551543.

[39] Chang B, Liang Y, Su S. Analyses of QoS-based relay deployment in 4G LTE-A wireless mobile relay networks. In *Proc. the 1st Asia-Pacific Conference on Communications*, October 2015, pp.62-67. DOI: 10.1109/APCC.2015.7412581.

[40] Gan X, Lu H, Yang G Y. Improving energy efficiency by optimizing relay nodes deployment in wireless sensor networks. In *Proc. the 9th IEEE International Conference on Communication Software and Networks*, May 2017, pp.306-310. DOI: 10.1109/ICCSN.2017.8230125.

[41] Li Y, Zhang Y, Zhou H, Jiang T. To relay or not to relay: Open distance and optimal deployment for linear underwater acoustic networks. *IEEE Transactions on Communications*, 2018, 66(9): 3797-3808. DOI: 10.1109/TCOMM.2018.2822287.

[42] Cui Q, Yang X, Hämäläinen J, Tao X, Zhang P. Optimal energy-efficient relay deployment for the bidirectional relay transmission schemes. *IEEE Transactions on Vehicular Technology*, 2014, 63(6): 2625-2641. DOI: 10.1109/TVT.2013.2295413.

[43] Lin G H, Xue G. Steiner tree problem with minimum number of Steiner points and bounded edge-length. *Information Processing Letters*, 1999, 69(2): 53-57. DOI: 10.1016/S0020-0190(98)00201-4.

[44] Chen D, Du D Z, Hu X D, Lin G H, Wang L, Xue G. Approximations for Steiner trees with minimum number of Steiner points. *Theoretical Computer Science*, 2001, 262(1/2): 83-99. DOI: 10.1016/S0304-3975(00)00182-1.

[45] Du D Z, Wang L, Xu B. The Euclidean bottleneck Steiner tree and Steiner tree with minimum number of Steiner points. In *Proc. the 7th Annual International Conference on Computing and Combinatorics*, August 2001, pp.509–518. DOI: 10.1007/3-540-44679-6_57.

[46] Măndoiu I I, Zelikovsky A. A note on the MST heuristic for bounded edge-length Steiner trees with minimum number of Steiner points. *Information Processing Letters*, 2000, 75(4): 165-167. DOI: 10.1016/S0020-0190(00)00095-8.

[47] Liu L, Zhou Y, Yuan J, Zhuang W, Wang Y. Economically optimal MS association for multimedia content delivery in cache-enabled heterogeneous cloud radio access networks. *IEEE Journal on Selected Areas in Communications*, 2019, 37(7): 1584-1593. DOI: 10.1109/JSAC.2019.2916280.

[48] Liu L, Zhou Y, Yuan W Z, Tian L. Tractable coverage analysis for hexagonal macrocell-based heterogeneous UDNs with adaptive interference-aware CoMP. *IEEE Transactions on Wireless Communications*, 2019, 18(1): 503-517. DOI: 10.1109/TWC.2018.2882434.

[49] Liu L, Zhou Y, Garcia V, Tian L, Shi J. Load aware joint CoMP clustering and inter-cell resource scheduling in heterogeneous ultra dense cellular networks. *IEEE Transactions on Vehicular Technology*, 2018, 67(3): 2741-2755. DOI: 10.1109/TVT.2017.2773640.

[50] Kashyap A, Khuller S, Shayman M A. Relay placement for higher order connectivity in wireless sensor networks. In *Proc. the 2006 IEEE International Conference on Computer Communications*, April 2006. DOI: 10.1109/INFOCOM.2006.273.

[51] Bredin J, Demaine E D, Hajiaghayi M T, Rus D. Deploying sensor networks with guaranteed capacity and fault tolerance. In *Proc. the 6th ACM Interational Symposium on Mobile Ad Hoc Networking and Computing*, May 2005, pp.309-319. DOI: 10.1145/1062689.1062729.

[52] Willson J, Zhang Z, Wu W, Du D. Fault-tolerant coverage with maximum lifetime in wireless sensor networks. In *Proc. the 2015 IEEE International Conference on Computer Communications*, April 26-May 1, 2015, pp.1364-1372. DOI: 10.1109/INFOCOM.2015.7218513.

[53] Lloyd E L, Xue G. Relay node placement in wireless sensor networks. *IEEE Transactions on Computers*, 2007, 56(1): 134-138. DOI: 10.1109/TC.2007.250629.

[54] Huang C, Zhang R, Cui S. Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints. *IEEE Transactions on Wireless Communications*, 2014, 13(2): 1074-1087. DOI: 10.1109/TW.2013.121813.130953.

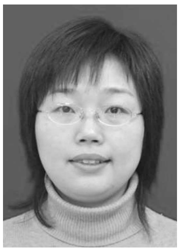
[55] Luo Y, Zhang J, Letaief K B. Transmit power minimization for wireless networks with energy harvesting relays. *IEEE Transactions on Communications*, 2016, 64(3): 987-1000. DOI: 10.1109/TCOMM.2016.2519418.

[56] Sudevalayam S, Kulkarni P. Energy harvesting sensor nodes: Survey and implications. *IEEE Communications Surveys and Tutorials*, 2011, 13(3): 443-461. DOI: 10.1109/SURV.2011.060710.00094.

[57] Lee J, Leyffer S. Mixed Integer Nonlinear Programming. Springer, 2012. DOI: 10.1007/978-1-4614-1927-3.

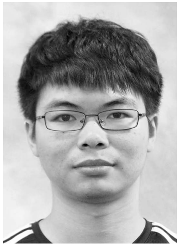
[58] Fowler R J, Paterson M, Tanimoto S L. Optimal packing and covering in the plane are NP-complete. *Information Processing Letters*, 1981, 12(3): 133-137. DOI: 10.1016/0020-0190(81)90111-3.

[59] Du D Z, Ko K I, Hu X. Design and Analysis of Approximation Algorithms. Springe, 2012.



Yi Wang received her B.S., M.S., and Ph.D. degrees in communication engineering from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, in 2000, 2003, and 2009, respectively. She is currently a lecturer with the School of Modern Posts, Nanjing University of Posts and Telecommunications, Nanjing. Her research interests include cloud computing.

Yi-Xue Liu received her B.S. degree in computer science and technology from the College of Computer, Sichuan University, Chengdu. She is currently a Master student with State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing.



Shun-Jia Zhu received his B.S. degree in computer science from Shanghai Jiao Tong University, Shanghai. He received his Master's degree in computer science from Brown University, Providence. He is now working at ByteDance Inc., Shanghai. His research interests include distributed systems, network optimization, time-series databases, and system observability. He participated in this work when he was an undergraduate student.

Xiao-Feng Gao received her B.S. degree in information and computational science from Nankai University, Tianjin, in 2004, her M.S. degree in operations research and control theory from Tsinghua University, Beijing, in 2006, and her Ph.D. degree in computer science from The University of Texas at Dallas, USA, in 2010. She is currently a professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai. Her research interests include data engineering, data mining, and network optimizations. She has published more than 200 peer-reviewed papers in the related area, including well-archived international journals such as IEEE TC, TKDE, TMC, TPDS, JSAC, and also in well known conference proceedings such as SIGKDD, WWW, IJCAI, NeurIPS, and INFOCOM. She has served on the editorial board of Journal of Combinatorial Optimization, and Discrete Mathematics, Algorithms and Applications, and as a PC member and a peer reviewer respectively for a number of international conferences and journals.

Chen Tian received his B.S., M.S., and Ph.D. degrees in communication engineering from Huazhong University of Science and Technology, Wuhan, in 2000, 2003, and 2008, respectively. He is currently a professor with the State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing. From 2012 to 2013, he was a postdoctoral researcher with the Department of Computer Science, Yale University, New Haven. His current research interests include data center networks, network function virtualization, distributed systems, and Internet streaming.