Computer Networks 104 (2016) 108–121

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Energy-sustainable relay node deployment in wireless sensor networks

Nusrat Mehajabin[a], Md. Abdur Razzaque[a,*], Mohammad Mehedi Hassan[b], Ahmad Almogren[b], Atif Alamri[b]

[a] Department of Computer Science and Engineering, University of Dhaka, Bangladesh
[b] College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:
Received 9 May 2015
Revised 11 May 2016
Accepted 16 May 2016
Available online 17 May 2016

Keywords:
Wireless sensor networks
Energy harvesting sensors
Optimal deployment
k-connected sustainable network
Network lifetime

ABSTRACT

Emergence of diverse renewable energy harvesting technologies and their incorporation into tiny sensor devices have given birth to Energy Harvesting Wireless Sensor Networks (EH-WSNs), where the problem domain has shifted from energy conservation to energy sustainability of the network. Renewable energy harvesting and depletion of sensor devices are stochastic and thus, energy availability in the devices is sporadic rather than continuous. Therefore, the optimal deployment of data routing devices (i.e., *relay* nodes) and their activity scheduling to ensure that, the data from all source sensors could be routed to the sink while keeping the network functional perpetually, is a challenging research problem. In this paper, we develop a multi-constraint mixed integer linear program (MILP) to minimize the number of *relay* nodes to be deployed in the network, while considering connectivity, sustainability and unpredictable energy harvesting and depletion rates. We refer to this problem as SMRMC (sustainable minimum-relay maximum-connectivity deployment) which is proved to be NP-hard. A light weight *k*-**connected** greedy solution to the SMRMC problem has been developed first for $k = 1$, and thereafter, a generalized solution has been presented for any k ($k \geq 2$) by constructing convex-polytopes among the existing *relay* nodes. Extensive simulation experiments have been conducted to validate the performance of the proposed deployment strategies. Performance studies carried out in MATLAB, show that the proposed SMRMC algorithms can achieve up to twice the network lifetime compared to state-of-the-art approaches whilst deploying minimum number of *relay* nodes.

1. Introduction

Wireless sensor networks (WSNs) provide long term and low cost solution to many emerging applications including surveillance [1], precision agriculture [2], environment quality sensing [3], machine and structural health monitoring [4], etc. Traditionally, sensor devices are equipped with chemical batteries having a limited lifespan [5]. Even with efficient energy conserving mechanisms, the battery would eventually drain out and the network would die [6]. A new direction of research focuses on arming the sensor devices with small renewable energy harvesters and super-capacitors for maximizing the network lifetime [7], giving birth to energy harvesting wireless sensor networks (EH-WSNs). Substituting the networking device's power supply with its renewable counterpart would prolong its lifetime [8] to some extent; however demise of the network is unavoidable, as the energy harvesting module would doze off to harvest energy eventually. Furthermore, energy depletion rate is higher than energy harvesting rate as of the current state of energy harvesters compatible for sensor devices [9]. However, optimal deployment of backup devices for the ones dozing off, would prevent the demise therefore, the network would be a sustainable one (*i.e.*, the network lives perpetually), given no hardware failure occurs.

For a large scale EH-WSN, most of the sensors are not within the sensing range of the sink node. Therefore, small sensor devices not only perform sensing tasks but also *relay* [10] the data packets of other sensors toward the sink. The routing solutions would require all sensor devices to be awake most of the time. As a result, these sensors are strained upon by additional processing and communication burdens [11], which lead to fast energy depletion and subsequently dropping their energy levels below the required threshold. This forces them to go to energy harvesting mode and contributes to minify the lifespan of the EH-WSN. To mitigate this problem, responsibility distribution of the networking devices has been used, where the *source sensor* nodes only perform the

* Corresponding author. Fax: +88-02-9667222.
E-mail address: razzaque@du.ac.bd (Md.A. Razzaque).

http://dx.doi.org/10.1016/j.comnet.2016.05.014

sensing tasks and *relay* nodes carry the sensed data of the *source sensor* nodes toward the sink in multi-hop fashion [12]. Basically, the *relay* nodes form an overlay network comprising of only *relay* nodes and the sink. Using various routing schemes [13], they can forward the data packets toward the sink node. This segregation of responsibilities immensely contributes to the longevity of the *source sensor* nodes. Surely, incorporation of *relay* nodes would aid in lifetime maximization but introduces yet another problem. That is, the network will be left disconnected when these *relay* nodes doze-off to harvest energy. In battery powered WSNs, this problem is solved by multi-tier *relay* node deployment strategies [14]. But, to the best of our knowledge, such an approach, for energy harvesting *relay* nodes, is scarce. A deployment strategy of *relay* nodes that not only maintains connectivity but also checks sustainability of the network using minimum number of energy harvesting *relay* nodes would be a great contribution in EH-WSNs. In this work, we aim to address the SMRMC (sustainable minimum-relay maximum-connectivity) problem for energy harvesting wireless sensor networks (EH-WSNs).

Planning for the placement or deployment of the resources is a fundamental task in WSNs, and an efficient placement strategy surely imparts efficient resource utilization, load balancing, lifetime maximization, etc. Literature for this discipline is quite rich as well. Various approximation algorithms have been developed to deploy *relay* nodes for lifetime maximization. Linear programs are also very popular solutions to ORP (optimal relay placement) and lifetime optimization problems. However, they are mostly NP-hard problems as the WSNs comprise of large area and huge number of nodes. New age solutions [15,16] are also studied adequately by incorporating energy harvesting components in sensor nodes. In [15], authors have deployed energy harvesting relays for α-lifetime maximization, *i.e.,* α% of the total nodes are responsible for sensing functions and (1−)% nodes perform *relay* duties. However, authors do not consider service disruption due to the energy harvesting process. In [16], *source sensor* nodes are allowed to perform *relay* duties when *relay* nodes are unavailable. Imposing *relay* responsibilities on low-powered sensor nodes is inefficient. Furthermore, energy harvesting is a stochastic process and running dynamic estimation algorithms to estimate the energy harvesting and depletion rates in real time is erroneous for EH-WSNs [17]. We intend to overcome these problems by planning optimal yet sustainable $k-connected$ *relay* node assignment. The SMRMC maintains connectivity and sustainability even if $k-1$ random *relay* nodes go into energy harvesting mode in the neighborhood of a *source sensor* node.

We first formulate the SMRMC problem as an MILP, optimizing the number of *relay* nodes for k connectivity. Since the optimal solution to the SMRMC is an NP-hard problem, we then develop suboptimal greedy alternative solutions. In greedy approach, we develop SMRMC ($k = 1$) algorithm for 1-connected *relay* node network and later extend the solution to a generalized one. The generalized approach ensures $k \geq 2$-connected network by exploiting the convex polytopes that can be formed among the existing *relays*. We summarize the major contributions of this work below:

- We design a novel optimization framework using mixed integer linear program, which minimizes the number of *relay* nodes in EH-WSNs and ensures ***k*-connected** sustainable sensor network.
- We develop a novel node density aware, sub-optimal, greedy energy harvesting *relay* node placement algorithm for ***k*-connected SMRMC** problem by exploiting convex polytopes among the *relays*.
- We formulate necessary constraints upon satisfying which the EH-WSN would achieve perpetual lifetime and provide an upper bound estimation for the number of *relay* nodes required to maintain connectivity throughout the network.
- Irrespective of the location of sink node and source node deployments, our SMRMC algorithms can create the ***k*-connected** *relay* node network. It also requires minimum installation and maintenance costs.
- Performance evaluation shows that, our SMRMC algorithm conveys up to twice the network lifetime than the other state-of-the-art node deployment algorithms.

The rest of the paper is organized as follows. Existing *relay* node deployment approaches in the literature have been presented in Section 2. Section 3 defines network and energy models and the problem of optimal *relay* node placement is defined in terms of network connectivity and sustainability. Section 4 presents our optimization framework and our proposed SMRMC greedy algorithms for EH-WSNs. In Section 5, results of the comparative performance evaluation experiments, carried out in MATLAB [18], for α-lifetime [15], $OPT-PRIM$ [16] and the proposed SMRMC systems are presented and discussed in detail. Finally, we conclude the paper in Section 6.

2. Related work

In this section, we discuss existing research works related to *relay* node placement and energy harvesting in WSNs. The literature studies for *relay* node deployment approaches can broadly be categorized into two types. *Firstly*, we discuss the well explored area of *relay* node placement in general sensor networks. Then, we review the energy harvesting *relay* node placement strategies. The first group of approaches are concerned with approximation algorithms for the planning problem. The second category is concerned with optimizing the energy demands of the EH-WSNs. Some of these works involve allocating communication responsibilities depending on the energy profile of the individual nodes.

Relay node placement for k connectivity, where $k = 1$ has been well studied in the literature. Steinerization technique [19] is one of the pioneer *relay* node placement strategies in WSNs. The authors have proved the optimal *relay* node placement problem to be NP-hard, and developed a minimum spanning tree (MST) based 5-approximation algorithm that solves *relay* node placement problem for $k = 1$. Their work has been studied and extended several times and used as the foundation for many other works [20–22]. Most of these works are closely related and provided good approximation algorithms to solve the problem. Cheng *et al.* [22] proposed an upgrade on [19] and developed a 3-approximation algorithm and a randomized algorithm with performance ratio 2.5. In [21], authors have extended Steinerization technique for $k = 2$ or 2-vertex connectivity, using a 10-approximation algorithm. Bredin *et al.* [23] proposed a *full fault tolerant relay node placement* which uses minimum number of *relay* nodes to achieve k-vertex ($k \geq 2$) connected network. However, none of these works accounts for energy harvesting wireless sensor networks. Lately, there have been a good number of research contributions in the field of *relay* node placement in ambient energy harvesting sensor networks.

Practical problems regarding the EH-WSNs have been studied by Misra *et al.* [24], where the authors proposed an approximation algorithm to solve *RNP* (relay node positioning) problem in EH-WSNs. They presume that a list of candidate locations (with higher energy harvesting potential) is at their disposal. They exploit this list to place *relay* nodes to a subset of the candidate locations. However, they have neither mentioned on the procurement of the list nor they have developed any mechanism on producing it. In [25], the authors considered a solar-based EH-WSNs and the sensor devices can dynamically alter their responsibilities according to the environment. They used LP (linear program) to fully utilize the residual energy of the sensor nodes so as to optimize the lifetime. However, some of the dynamic tasks may consume significant amount of energy so that the harvested amount of energy might prove to be insufficient. In addition, dynamic decision making inflicts computation and communication overheads. Another noteworthy contribution in this field is [26], where the authors developed an opportunistic routing protocol for WSNs running solely on energy harvesting nodes. The authors also presented the basic architectural overview and characteristics of EH-sensor nodes.

Abu-Baker *et al.* [15] developed a strategy to achieve α-lifetime maximization using $(1-\alpha)\%$ nodes of the network as *relays*. That is, the network is assumed to be functional or alive if data from α portion of the sensors can be collected by the base station. The authors formulated an LP (linear program) to maximize the lifetime of the network and discussed about the break point for infinite lifetime. However, it lacks in proper analysis of the break point and necessary conditions to achieve it. Again, the arrangement would collapse when energy drains out from any of the $(1-\alpha)\%$ nodes. In [16], the authors proposed an optimization framework with two LPs. The first one is for lifetime maximization and the second one is for minimizing hop counts to the sink. As optimal *relay* node placement problem is an NP-hard, the authors proposed a greedy algorithm $OPT - PRIM$ for *relay* node placement. They also considered the *source sensor* nodes are combination of energy harvesters and battery powered nodes. The source sensor nodes will pick up *relay* responsibilities when the *relays* are asleep. This choice definitely reduces the network lifetime by putting extra load on to the battery powered devices. Furthermore, the presence of redundancy and redundancy removal steps of $OPT - PRIM$ increase complexity. The arrangement is not a sustainable one either.

Although our work is motivated from and has a good level of similarity with the solution architecture of $OPT - PRIM$ [16], there are some noteworthy differences between the two. *Firstly*, our optimization framework achieves sustainable lifetime by deploying minimum number of *relay* nodes for ***k*-connected** network of *relays*; whereas, the $OPT - PRIM$ increases network lifetime (not sustainable) through flow maximization and hop count minimization. *Secondly*, our ***k*-connectivity** greedy algorithm is connectivity and density aware, producing a near optimal solution to the SMRMC problem; whereas, the $OPT - PRIM$ models connectivity as a consequence rather than a constraint. Hence, more than optimal number of *relays* could be placed to ensure connectivity. *Thirdly*, $OPT - PRIM$ requires a redundancy removal step which adds to the complexity of the algorithm and our algorithm expands from the sink onward, using coverage factor of each location and thus it requires no redundancy check for connectivity. *Finally*, $OPT - PRIM$ is only concerned with $k = 1$, it accounts for any *relay* node dozing-off by inflicting *relay* responsibilities to *source sensor* nodes. However, our solution implements **$k \geq 1$ connectivity**, never causing *source sensor* nodes to take over the *relay* duties. Unlike most of the works discussed above, our algorithm is independent of sink's position and the network topology. This work would enable EH-WSN application developers and researchers to plan node placements of their projects in a cost effective way.

3. Network model and problem definition

We assume the application terrain is accessible, for example, structural health monitoring, precision agriculture or any indoor applications that allow us to pre-compute the placement of the *relay* nodes and plant them accordingly. We also assume the placement of the *source sensor* nodes and the sink node are known a *priori* and it is our knowledge base. The *relay* nodes are equipped with energy harvesting capabilities (e.g., solar power, thermal energy, wind energy, salinity gradients, etc.) [27] and the network is known as EH-WSN (Energy Harvesting Wireless Sensor Network). Energy harvesting is a stochastic process, as the harvesting rates are completely dependent on temporal (likely to gain different values at different time instances) and spatial (likely to vary depending on the location of the nodes) variabilities. The energy harvesting rates may be less than energy depletion rates depending on the situations.

3.1. Network topology and energy model

We consider a single sink at a random location Y and many *source sensor* nodes scattered randomly all over the terrain and their locations constitute a set S. A *source sensor* node can transmit its sensed data through a *relay* node or directly to the sink. All *relay* nodes have the same transmission range T as the *source sensor* nodes. Two *relay* nodes with locations i and j can communicate with each other when the Euclidian distance between them satisfies $dist(i, j) \leq T - \delta$. Here, δ is a small distance used to mitigate the effects of anisotropic path loss. We discuss more about δ in Section 4.3. Similarly, a *relay* node at location i can carry the data packets produced by a *source sensor* node at location $s \in S$ when $dist(s, i) \leq T - \delta$. The sink node is assumed to be connected to power mains or have infinite power.

For energy model, we consider the characteristics of energy harvester sensors [28], where sensors will be operational only if sufficient amount of energy has been harvested. We call it activation energy E. A *relay* node can harvest and spend energy simultaneously. Hence the energy harvesting rate γ and depletion rate λ could have one of the following three relationships: $\lambda \equiv \gamma$, this situation is presented by the blue line in Fig. 1; $\lambda < \gamma$, denoted by olive lines, means that the node is able to save some amount of harvested energy after expenditure; and, $\lambda > \gamma$ states that the energy is draining out faster than harvesting, depicted using red lines. When the residual energy E' of a node drops below a minimum threshold $\mathbb{M}$, it stops all communications (*i.e.* turns off its transceiver) and switches to energy harvesting mode. Fig. 1 depicts this scenario by the green lines for a particular *relay* node.

Traditionally, network lifetime for WSNs is defined as the time duration after which the first sensor node runs out of its energy [29]. However, in this work, we customize the definition of network lifetime for energy harvesting wireless sensor networks (EH-WSNs). The time duration after which a *source sensor* node (with data packets to transmit) fails to find a path to the sink node is redefined as the network lifetime for EH-WSNs.

3.2. Problem definition

In this section, we define the SMRMC problem in network connectivity and sustainability domains.

3.2.1. Network connectivity

Network connectivity in WSNs refers to the communication foundation where all the *source sensor* nodes can successfully transmit their sensed data to the sink either directly or via *relay* node(s). Fig. 2 shows a typical network, where most of the *source sensor* nodes are out of sink's coverage area and thus they have to depend on *relay* nodes for data delivery. The *relays* need to be placed in such a way that all the *source sensors* can establish connection with the sink. An unplanned deployment may require huge number of *relays* to ensure connectivity, causing extra cost as well as maintenance overhead. Therefore, an optimal deployment strategy is required that would ensure connectivity of the *source sensors* to the sink with minimum number of *relay* nodes. That is, even if a single *relay* is taken off the network, one or more *source sensor* nodes loose their connectivity; similarly, addition of a new *relay* node after optimal deployment would not discover any *source sensor* node for the first time. And, no other deployment of *relay* nodes can achieve the above with fewer number of *relay*

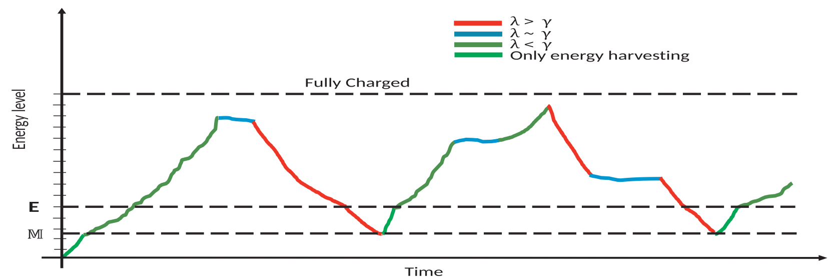


Fig. 1. Charging and discharging cycles of a *relay* node. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

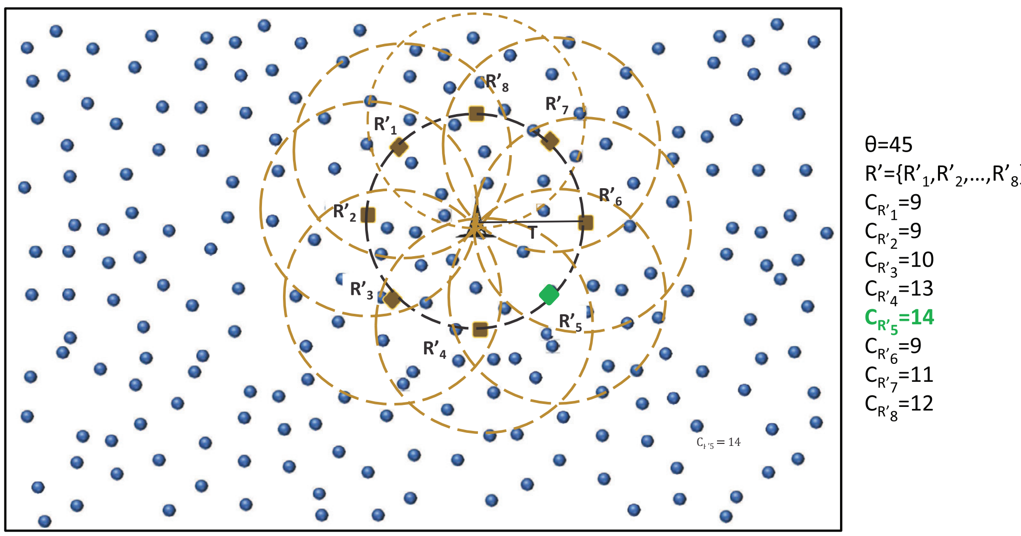


Fig. 2. Deployment process of **1-connected** SMRMC in EH-WSN.

nodes. This problem is referred to as **1-connected SMRMC** problem. More explicitly, if R' is the candidate *relay* node locations (all possible locations for *relay* nodes) and R is the set of *relay* node locations which can achieve **1-connected SMRMC**, then $\mathbb{P}_{<|R|}(R')$ cannot form a **1-connected** network of *relays*, where $\mathbb{P}(R')$ is the set of all possible subsets of R'.

3.2.2. *Network sustainability*

In EH-WSNs, the energy level of some of the *relay* nodes would eventually diminish below minimum threshold $\mathbb{M}$ and they switch into energy harvesting mode, leaving the network disconnected. To ensure sustainability, some redundant *relay* nodes must be placed to perform the responsibilities of the departing ones. A naive solution might be to place redundant *relay* nodes at very close proximity to previously placed *relay* nodes. That is, to choose the best locations for connectivity to meet sustainability. This is not a feasible solution as the energy harvesting rates of the locations are unknown and placing the next set of *relay* nodes at exactly (or approximately) the same locations would cause all sets of *relay* nodes to encounter same adverse environmental effects such as, shades, barriers, physical destruction or damage to the location, stamping by animals, etc. Moreover, it would directly affect the network lifetime due to inaccuracy in energy availability prediction caused by the volatile relationship between energy harvesting and depletion rates discussed in Section 3.1. Energy harvesting from ambient sources is unreliable and uncontrollable (*e.g.,* the sun, wind, heat, vibration, etc.). Therefore, energy availability could be modeled as spatial and temporal variable in a stochastic process [30]. On the other hand, energy sources demonstrate partial patterns or periodic properties such as high energy harvesting from solar sources during day time and very low to nil harvesting at night time. Thus, it can be modeled as periodic process [17] as well. Such contradicting inference leads to erroneous and mercurial prediction and estimation results for energy availability. Hence, any node placement algorithm depending on the energy availability of the *relay* nodes would be highly unstable. Therefore, the backup *relay* node placement algorithm should be independent of the ambient energy harvesting rate or future energy availability of the *relay* nodes. In this paper, we develop an efficient redundant *relay* node placement strategy where each *source sensor* has k relays in its neighborhood and an intelligent scheduling of these k *relays* would enable the network to sustain perpetually. The least number of *relay* nodes that can ensure network connectivity, even when any arbitrary $k-1$ *relay* nodes from a *source sensor* neighborhood doze off to harvest energy (when energy level drops below $\mathbb{M}$) is the optimal number of *relay* nodes for ***k*-connected SMRMC**.

4. Proposed model

We begin this section by introducing our basic node placement strategy. Next, we formulate an optimization framework for ***k*-connected SMRMC** problem using mixed integer linear programming (MILP), which is proven to be an NP-hard problem. Then, a greedy alternate solution to the **1-connected SMRMC** problem has been presented. Finally, we discuss on a generalized light weight solution to the SMRMC problem, i.e., for any $k \geq 2$, so as to achieve a sustainable EH-WSN.

Table 1
Notations for the optimization framework.

Symbol	Description
R	Set of *relay* node locations
$\mathbb{R}$	Family set of k *relay* node sets
S	Set of all *source sensor* node locations
$\mathbb{Q}$	$\mathbb{R} \setminus R$
Y	Location of the sink node
E	Activation energy of the *relay* nodes
$E_q'^Q$	Residual energy of *relay* $q \in Q$
$dist(i, j)$	Euclidean distance between nodes i and j
T	Transmission range of *relay* nodes
δ	A small distance for coverage safety margin
$\mathcal{X}_i^R$	Binary variable, contains 1 if $i \in R$ and 0, otherwise
$\mathcal{Z}_s^i$	Binary variable, contains 1 if $dist(i, s) \le T$ and 0, otherwise
γ_r^R	Energy harvesting rate of *relay* node r of *relay* set R
λ_r^R	Energy depletion rate of *relay* node r of *relay* set R

4.1. Node placement strategy

Classically, this kind of problem is dealt with *Max-min* approaches (*i.e.,* minimizing number of *relay* nodes while maximizing network lifetime). However, in this work, we develop an algorithm to obtain sustainability in EH-WSNs, *i.e.,* the network would have infinite lifetime (theoretically) by facilitating each *source sensor* node at least one live path toward the sink. Hence, we create sets of connected *relay* nodes until the state of sustainability for the network is achieved. Moreover, our aim is to optimize (minimize) the number of *relay* nodes (in each set) required to serve every *source sensor* node in the vicinity. Intuitively, we can achieve this goal by deploying *relay* nodes at locations where it could discover maximum number of unattended *source sensor* nodes. Then again, this will create isolated clusters of *source sensor* nodes connected with a *relay* node. Moreover, it is quite likely that, there will be no path to the sink from some of these clusters. To eradicate this problem, *relay* node deployment should start by assuring and maintaining connectivity with the sink. Therefore, the first *relay* node must be within the transmission range of the sink. And from there on, a *relay* node must be within the transmission range of another *relay* node or the sink. The assignment will be further optimized if the *relay* nodes have maximal non-overlapping transmission areas. These two observations are contradictory and thus a trade-off between amount of overlapped transmission ranges of the *relay* nodes and coverage of *source sensor* node must be made. The distance between two neighbor *relay* nodes i and j follows the condition $dist(i, j) \le T - \delta$, as stated in Section 3.1.

4.2. Optimal relay node placement framework

Given the set of all *source sensor* node locations, S, and position of the sink node Y, we formulate an optimization framework for our **k-connected SMRMC** problem. It is a mixed integer linear program (MILP), designed for k coverage, *i.e.,* k independent sets of *relay* nodes are deployed to carry the sensed data packets from the *source sensor* nodes to the sink. Table 1 shows the notations used in the MILP formulation. Here, $\mathbb{R}$ represents the family set for independent *relay* sets, $\{R_1, R_2, \cdots, R_k\}$ and $|\mathbb{R}| = k$. The objective function and the constraints of the MILP are formulated as follows:

$$minimize: \sum_{p=1}^{k} | R_p | \tag{1}$$

$$subject\ to:$$

$$dist(s, i) \le T - \delta; \quad \forall s \in S,\ \exists\, i \in R,\ \forall R \in \mathbb{R} \tag{2}$$

$$path(i \overset{j}{\rightsquigarrow} Y) = true; \quad \forall i \in R,\ \exists j \in R \cup \emptyset \tag{3}$$

$$\frac{E}{\gamma_r^R} \le \sum_{Q \in \mathbb{Q}} \frac{E_q'^Q}{\lambda_q^Q - \gamma_q^Q}; \quad r = \operatorname*{argmin}_r \left(\frac{E_i'^R}{\lambda_i^R - \gamma_i^R} \right),\ \forall i \in R, \forall R \in \mathbb{R}, \quad \forall Q \in \mathbb{Q} \tag{4}$$

$$\sum_{\forall R \in \mathbb{R}} \sum_{\forall i \in R} \mathcal{X}_i^R = 1 \tag{5}$$

$$\sum_{\forall R \in \mathbb{R}} \sum_{\forall i \in R} \mathcal{Z}_s^i \ge k; \quad \forall s \in S \tag{6}$$

Here, Eq. (1) is the objective function and Eq. (2)– (6) are the constraints. The optimization function minimizes the number of *relay* nodes to be deployed in the vicinity so as to achieve k-coverage for each *source sensor* node, *i.e.,*$\forall s \in S$. The **coverage constraint** in Eq. (2) indicates every *source sensor* node is monitored by at least one *relay* node from each of the k sets. The Eq. (3) is the **connectivity constraint**, which ensures that there exists at least one path from any *relay* i to the sink Y, either directly or via a forwarding *relay(s)* j. The **energy sustainability constraint** of Eq. (4) says that the time required to harvest enough energy to be active for the shortest lived *relay* of any set must be less than or equal to the aggregated service time of the shortest lived *relays* from all the other *relay* sets. That is, the shortest lived *relay* $r \in R$ should harvest as minimum as E amount of energy (activation energy) by the time all the shortest lived *relays* $q \in Q$ of all the other *relay* sets $Q \in \mathbb{Q}$, deplete of their residual energies. The shortest lived *relay* $r \in R$ is derived by $r = \operatorname{argmin}_r(\frac{E_i'^R}{\lambda_i^R - \gamma_i^R}),\ \forall i \in R, \forall R \in \mathbb{R}$; similarly, the shortest lived *relays* for other sets $Q \in \mathbb{Q}$ can be derived. Thus, the **energy sustainability constraint** guarantees that each *source sensor* node would always have a relay node for data transfer. The Eq. (5) ensures that the set of *relay* nodes in $\mathbb{R}$ are independent, i.e., any relay node i can be the member of one and only one set in $\mathbb{R}$; otherwise, our SMRMC solution would fail to provide sustainability. The constraint in Eq. (6) imposes that each of the *source sensors* in the network is covered by at least k *relay* nodes.

Of course, the above linear program would generate many feasible solution sets (feasible region) as there is no constraint imposed on physical locations of the *relay* nodes. That is, small changes in locations of some *relay* nodes might produce a new feasible solution. However, for every feasible solution, the number of *relay* nodes will remain the same (minimized). Furthermore, if the number of the *source sensor* nodes increases to a large number, the linear program will suffer from huge time and space complexities. The same is true when the area of the network is expanded. In other words, the above MILP is not scalable, i.e., as $|S| \to \infty$, it becomes an NP-hard problem.

4.2.1. SMRMC is NP-hard

The SMRMC or sustainable minimum-relay maximum-connectivity problem could be reduced into the classical *RNP* (relay node positioning problem) [22]. In [22], authors proposed a deployment strategy for placing minimum number of *relay* nodes (RNs) in a *WSN* vicinity so that between every pair of *source sensor* nodes, there is a path consisting of *relays* or *sources*. And, the maximum hop length of the path is no longer than the common transmission range $r > 0$ of the *source* nodes. This *RNP* problem is equivalent to the Steiner minimum tree with minimum number of

Table 2
Notations for the greedy algorithm.

Symbol	Description
Y	Location of the sink
θ	Coverage factor measurement interval
T	Transmission range of the *relay* nodes
δ	A short distance for coverage safety margin
R'	Set of candidate *relay* node locations
$C_{i'}$	Coverage factor at *relay* node location i'
ζ_{Ψ_v}	Centroid of convex polytope Ψ_v

Steiner points and bounded edge length problem ($SMT - MSPBEL$) [19], which is proven to be an **NP-hard** problem.

As *RNP* is a subproblem of SMRMC and *RNP* is equivalent to a **NP-hard** problem (Steiner minimum tree), a sub-problem of SMRMC is equivalent to Steiner minimum tree. Hence, SMRMC is **NP-hard**. In other words, *RNP* is a subproblem of SMRMC; *RNP* $\equiv$ **NP-hard** and therefore, SMRMC $\equiv$ **NP-hard**.

To overcome the NP-Hardness of the optimal solution (MILP), we develop an alternative greedy solution which will produce a good enough *relay* node deployment plan for the SMRMC problem.

4.3. Greedy solution

Now, we develop a greedy iterative algorithm which produces a near optimal *relay* node placement plan for a large scale WSN. This greedy node placement strategy creates a **1-connected** *relay* node network covering all the *source sensor* nodes. We denote this situation as $k = 1$ coverage. Intuitively, if the number of *source sensor* nodes is n and they are distributed over an area of A m^2, then the *source sensor* node density, $\Delta = \frac{n}{A}$ nodes/m^2. And if the transmission range of a *relay* node is T, then approximately $\pi T^2 \times \Delta$ *source sensor* nodes are covered by one *relay* node. Hence, the number of *relay* nodes required to cover the whole area A would be $\frac{n}{\pi T^2 \times \Delta}$. However, applying this naive mathematics for *relay* assignment might create holes or gaps in the wireless vicinity. In practical WSNs, the transmission area is more comparable to an amoeba shaped region, rather being a circular one [31]. Therefore, instead of considering the transmission area of the *relay* nodes as disks of radius T, we consider the maximally inscribed hexagons in that disk. This hexagon would have the area of $\frac{3\times\sqrt{3}}{2}T^2$ m^2. Hence, $|R'| = \frac{A}{\frac{3\times\sqrt{3}}{2}T^2}$ *relay* nodes will be required to monitor the entire area and to keep connectivity among them it would require at most another $|R'| - 1$ *relay* nodes. Thus, the maximum number of *relay* nodes required to deploy is theoretically limited by, $2 \times R' - 1$.

However, practical *relay* node placement might differ significantly due to *source sensor* node distribution across the network. The philosophy behind node deployment policy of our work is to land *relay* nodes at places where they can contribute the most, while maintaining association with the sink either directly or via another *relay* node. We present our near optimal greedy algorithm in Algorithm 1 and Table 2 contains the notations used in the algorithm. The algorithm is node density and connectivity aware. What follows, we describe our SMRMC (k=1) algorithm in detail.

We use the set of *source sensor* node locations S as input to our algorithm along with the physical location of the sink Y. Network topology model based on circular transmission range is unrealistic due to anisotropic path loss [32]. Hence, a transmission area with a radii equal to the actual transmission range T is unrealistic as well. To address this issue, we use a transmission range of $(T - \delta)$, where δ is a small distance for coverage safety margin. Even though this consideration increases the number of *relay* nodes to be deployed in the network, it helps much in achieving

Algorithm 1 1-connectivity SMRMC algorithm.

```
Input: S, Y, θ, T, δ
Output: R
initialize R ← Y
while S ≠ ∅ do
  for each (x, y) location in R do
    generate R′ using Eq. (8) and (9);   ∀ a ∈ Θ
  end for
  R′=R′ \ R
  for each candidate relay location i′ ∈ R′ do
    generate coverage factor C_i′ using Eq. (7)
    C ← {C ∪ C_i′}
  end for
  L = AssignRelay(C)
  R ← {R ∪ L}
  S ← S \ C_L,  C_L is the coverage factor of locations in L
end while
procedure: AssignRelay
  Input: C
  Output: a sorted list of candidate relay locations L
    while C ≠ ∅ AND | L |≤ m do
      C_sort ={C_g |  | C_sort[1] |>| C_sort[2] | ··· >| C_sort[| C |] |  And C_g ∈ C}
      L ← {L ∪ location of C_sort[1]}
      C ← {C_g \ C_sort[1] | C_g ∈ C_sort}
    end while
```

good connectivity throughout the network consequently increasing the network lifetime. Other than δ we have another design parameter, the coverage factor measurement interval θ (an angle). We start node deployment at $T - \delta$ distant locations from the sink so as to maintain connectivity with the sink. We measure *source sensor* coverage factor (C) for every candidate *relay* location $i' \in R'$.

Source sensor node coverage factor C would be the number of *source sensor* nodes that could be monitored if a *relay* node is placed on at any arbitrary location (x, y). In other words, the number of *source sensor* nodes that are within the transmission range of location (x, y) is its coverage factor, measured as follows,

$$C_{i'} = \{s \mid dist(s, i') < T - \delta\} \quad \forall\, s \in S,\; i' \in R'. \tag{7}$$

We measure C value for every *candidate relay location*, $i' \in R'$. Initially, the set of *candidate relay locations* R' is the set of physical locations that are at $T - \delta$ distance from Y (pole) and are $\theta°$ apart. As a result, there would be $\frac{2\pi}{\theta}$ *candidate relay* locations. And, they are derived as follows:

$$R'(t, \theta) = ((T - \delta), (\theta \times a)); \qquad a \in \Theta = \{0, 1, 2, \cdots, \frac{2\pi}{\theta}\}, \tag{8}$$

$$R'(x, y) = (t\cos\theta, t\sin\theta); \quad t = T - \delta. \tag{9}$$

Fig. 2 has a pictorial representation of initial *candidate relay location* generation and *coverage factor* calculation. It shows the highest coverage factor bearing location becomes an actual location for $k = 1$ *relay* node. As the algorithm proceeds, we update R' with respect to the set of *relay* nodes R. *Candidate relay locations* are calculated around all the existing members of R. Every time we update R', we choose m most contributory *candidate relay locations* to be actual locations for *relay* nodes. We continue this process until all the *source sensor* nodes have been covered, *i.e.*, when the algorithm ends, all the *source sensor* nodes have at least one path to the sink. Algorithm 1 generates sub-optimal results since our greedy choice may cause overlapping coverage of some of the *source sensor* nodes. However, **1-connected SMRMC** algorithm will be able to provide a good enough solution for large scale deployments; whereas, the MILP in Eq. (1) becomes $NP - hard$.

Although the network is connected by $|R|$ *relay* nodes, when any of these *relay* nodes run out of energy some of the *source sensor* nodes will be detached from the sink. Hence, with $k = 1$, maximum life expectancy of the EH-WSN is when the first *relay* node depletes of its energy. To prevent this death we employ $k \geq 2$ connectivity algorithm so that when a *relay* node dozes off to harvest energy another one takes over the responsibility, details are presented in Section 4.4.

4.4. Generalized SMRMC algorithm

As we are dealing with energy harvester *relays*, dozing off to accumulate energy is a common phenomenon. For SMRMC (k=1), no other *relay* node can monitor the *source sensor* nodes that have been left unsupervised; thus, they will not be able to deliver their data packets to the sink. The EH-WSNs will not sustain for long in such scenarios. For sustainability of the network, we need to cover each *source sensor* by more than one *relays*. In this way, the network will survive even if one or more *relay* nodes doze off to energy harvesting mode. As there will be another or a few other *relay* nodes to administer their communication to the sink. Henceforth, we formulate this problem as a k-coverage problem, where, every *source sensor* node is monitored by at least k *relays* and removing $k-1$ *relays* from the vicinity of a *source sensor* does not affect the connectivity of the network. Now, we discuss our proposed generalized SMRMC ($k \geq 2$).

4.4.1. SMRMC (k $\geq$ 2) using convex polytopes

In this section, we construct *convex polytopes* among the nodes to find optimal solution to the problem of covering each *source sensor* nodes by k number of *relay* nodes. Our **1-connectivity SMRMC** algorithm creates a density profile for the entire network, i.e., *relay* nodes are placed on the highest density locations in the vicinity. Thus, if we start building convex polytopes using the existing *relay* nodes, a new *relay* node at the centroid of those polytopes will be a lucrative position for a backup *relay* node. This will not only enhance connectivity of the *source sensor* nodes but also work for hop distance minimization of the network. And, if we plant these new *relay* nodes maintaining connectivity as well, it would enable us to establish an alternative path toward the sink for each *source sensor* node in the transmission range of the newly placed *relay* node. Now, maintaining connectivity is only a matter of having sink or any of the newly placed *relay* nodes as a member of the convex polytopes.

We write a function *BuildConvexPolytopes(n)* that returns all possible convex polytopes consisting of n nodes. That is, the convex hull of size n is returned. We build convex polytopes using the gift wrapping [33] algorithm because of its simple implementation and moderate computational complexity compared to other algorithms. The algorithm treats an arbitrary point as reference and iteratively builds a polytope with the nearest node which has all the other nodes to its right.

The motivation behind using convex polytopes for our $k-$*coverage* is simple. We intend to keep the number of *relay* nodes minimum. That will only happen if a newly placed *relay* node contributes the most in *source sensor coverage*. From our first set of *relay* nodes R, we get the idea where *source sensor coverage* is dense furthermore, if a few $i \in R$ are gathered around then the centroid of that gathering must have good *coverage factor*. Therefore, the centroid of a convex polytope seems like an ideal candidate. We can trim out those polytopes whose member nodes are not within transmission range of the centroid. We iteratively reduce the size of the convex polytopes until all the *source sensor* nodes are covered.

Now, we discuss the generalized SMRMC (k=2) in detail. We use the *relay* nodes deployed by **1-connectivity SMRMC** algorithm to choose the *relay* nodes for **k(=2)-connectivity SMRMC** algorithm. First, we generate the convex polytopes of size $|R|$(if any). For network configuration of Fig. 3(a) we can see that, it's not possible to create a polytope of size $|R|$. Hence, we create the next largest one possible. In Fig. 3(a) we can see such a polytope (red). Next, we check whether the centroid of the new convex polytope is a viable location for k=2 *relay* node or not. If all the members of the polytope are within T distance from the centroid and minimal overlap is maintained for *source sensor coverage* with the nearest *relay* node, we can place a *relay* node at the centroid. However, the red polytope in Fig. 3(a) fails this test. Our next step is to repeatedly decrease $|R|$ by 1 unit ($|R|-1, |R|-2, \cdots, 2$) and choose a feasible polytope. The green polytope in Fig. 3(a) is such a polytope. We continue this process until all the *source sensor* nodes are covered by k=2 *relay* nodes, as shown in Fig. 3(b). However, convex polytope building approach would only guarantee the connectivity and coverage of *source sensor* nodes that are inside the biggest polytope created by the *1-connectivity relay* nodes. To eradicate this problem, we use a few dummy *relay* nodes (yellow triangles in Fig. 3(b)). We assume they are placed at the boundary of the wireless vicinity at every T meter intervals. Similarly, we can generate set of *relay* nodes for $k = 3$ by using the *relay* node set for $k = 2$ by constructing convex polytopes.

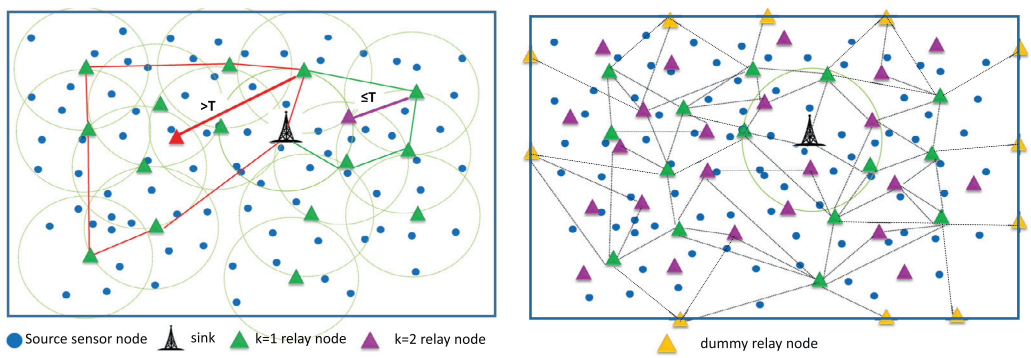
For the EH-WSN to be sustainable, we need to estimate the value of k that depends on the charging and discharging rates of the *relay* nodes. Let γ and λ be the average charging and discharging rates. First, we consider homogeneous charging and discharging cycles that is all the nodes have the same charging and discharging rates. If $\gamma \geq \lambda$, then $k = 1$ will suffice to maintain connectivity and sustainability of the EH-WSN. However, if $\gamma < \lambda$, we will require $k \geq 2$; for homogeneous *relay sensor* nodes, $k = \lceil \frac{\lambda}{\gamma} \rceil + 1$ would suffice the network sustainability requirement. However, homogeneous charging and discharging rates for EH-WSNs is an impractical assumption. In case of heterogeneous charging and discharging rates, $\gamma < \lambda$ is the crucial case to handle. In such condition, the set of *relay* nodes $\mathbb{R}$ must satisfy the inequality (4) so that the time it takes for the shortest lived *relay* node of a particular set to harvest enough energy to get back to active state, should be smaller than or equal to the aggregated time for all the shortest lived *relay* nodes in other sets to deplete of their energies. Hence, there will always be at least one set of active *relay* nodes monitoring the wireless vicinity.

4.4.2. Sustainable switching of the relay sets

After successful derivation of k sets of *relay* nodes to maintain connectivity as well as sustainability, these k sets need an efficient switching mechanism to share *relay* responsibility among them. Though such switching mechanism in EH-WSNs stems a challenging research domain by itself, we provide a generic solution to the problem.

In $k-connected$ generalized SMRMC, each set has higher number of *relays* than its previous one, *i.e.,* $|R1| \leq |R2| \leq |R3|$. Hence, when a *relay* set dozes off, the *relay* responsibility switches to the next set having the least number of *relays* and sufficient residual energy. In other words, the switching mechanism follows a cyclic order over the sets of *relay* nodes. Given that $R1$ is the first set to perform *relay* duties, $R2$ will pick up when $R1$ dozes off. Similarly, when it is the time for $R2$ to go to harvest energy, $R3$ will take over the responsibility, not $R1$. Even though $R1$ has the least number of *relay* nodes, theoretically it would not have sufficient residual energy.

Note that the above switching strategy is not an optimal one, which requires instantaneous monitoring of the energy harvesting and depletion rates and dynamic switching schemes. Development of such a scheme for dynamically switching *relay* responsibility among the available *relay* sets is left as a future work.



(a) Convex polytope (b) **2-connected** deployment

Fig. 3. Deployment process of *relay* nodes in **2-connected** SMRMC. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.4.3. Determination of control parameters

In the greedy solution of the SMRMC, we have three control parameters, namely the candidate *relay* selection interval θ, coverage safety margin δ and number of *relay* nodes deployed per iteration *m*. These are tuning parameters and can be tuned according to fault tolerant requirements of the network. The first pass of the algorithm deals with the candidate region of the sink *Y* only. The placement of *relay* nodes around the sink node is of crucial importance as they would carry traffic from the entire network. We are considering largest hexagon inscribed in the disk of radius equal to the transmission range *T* of a *relay* node, we can ensure coverage range of $\sqrt{3}\times T$ if we place one *relay* node at each peak point of the hexagon. Hence, if we place $m = 6$ equidistant *relay* nodes, our greedy algorithm can guarantee coverage range of $\sqrt{3}\times T$ from the sink in the first pass. In this particular case, $\theta = 60^\circ$ will suffice. To determine the value of δ, we need to monitor the *SNR* from the deployed *relay* node. Thus, this is strictly a design parameter.

4.5. Complexity and scalability of SMRMC

Both the proposed optimization framework and the greedy generalized SMRMC are centralized algorithms, which raise concerns for scalability of the proposed algorithms. We have analyzed the complexities of these algorithms and present the asymptotic upper bound analysis here. The integer part of the MILP implies that its upper bound computational complexity is $O\binom{n}{R,}$ which is the number of combinations of *R relay* nodes over *n source sensor* nodes. Thus, as $n\rightarrow\infty$ the problem becomes infeasible. The framework also suggests *R* and *k* values are increased with *n*; hence, $R\rightarrow\infty$ when $n\rightarrow\infty$.

On the other hand, for **1-connected SMRMC**, the upper bound of asymptotic analysis is $O(RR'n)$. Here, R' is the number of candidate *relay* nodes. It is evident that $O(RR'n) < < O\binom{n}{R}$. Moving on to our last algorithm for generalized SMRMC ($k\geq 2$), the computational complexity is analyzed as $O(RR'n + n^2 h\log h)$, where *h* is the average number of nodes in the polytopes. In practical cases, $h < < n$, as the nodes building the polytopes has to be within each other's transmission ranges. From the discussion above, we can conclude that $O(RR'n) < O(RR'n + n^2 h\log h) << O\binom{n}{R}$. Our optimization framework becomes infeasible as the number of *source sensor* nodes is increased; however, **1-connected SMRMC** and generalized SMRMC can find a solution in polynomial time. Thus, the optimal solution is not scalable; however, the greedy solutions are able to find good enough solution in polynomial time.

5. Performance evaluation

In this section, we implement two versions of SMRMC (*i.e.*, $1-connected$ and $k-connected$) along with two state-of-the-art approaches $OPT-PRIM$ [16] and α-lifetime maximization [15] using a commercial software MATLAB [18] and present the comparative performance results. The topology of the *source sensor* nodes, energy harvesting and depletion rates, initial energy and transmission range of the nodes, energy thresholds, etc. are kept same for the studied approaches in support of fair comparison. For the implementation of α-lifetime maximization [15], we have considered $\alpha = 80\%$, *i.e.,* 20% nodes take the *relay* responsibility and for $OPT-PRIM$ [16], a greedy deployment process is run iteratively starting from sink so as to increase the coverage factor.

5.1. Simulation environment

We have deployed [500, 3500] *source sensor* nodes with uniform random distribution over an area of [300, 2000] m^2 and a sink node is placed at a random location. All the *source sensor* nodes disburse their data to the sink. The transmission range of the *source sensor* and *relay* nodes are assumed to be the same. We use coverage factor measurement interval, $\theta = 36$ that generates 10 initial candidate locations and set coverage safety margin, $\delta = 10$m to mitigate anisotropic path loss and for better simulation scalability [32]. We place $m = 1$ *relay* node after each iteration. For energy harvesting sensor nodes, we are using random renewable energy distribution following the data provided in technical note of EH-Link [34], so as to model the inconstant and discrete characteristics of renewable energy supplies [35]. The amount of energy of a *relay* is chosen from the range [0.0005, 0.08] mW with uniform random distribution. Furthermore, we divide this range in three sub-ranges [0.0005, 0.02] for night time (8.00 pm –4.00 am) or dark hours, [0.02, 0.05] for afternoon time (3.00 pm–8.00 pm) or cloudy weather and [0.05,0.08] for day time (4.00 am–3.00 pm) or sunny weather. This helps the simulator to generate more realistic scenarios. We are considering a 1000 μF capacitor for the solar cell and the sensor can start communication once it has reached $E = 5.4$ V; and, the sensor stops all communications once its energy level goes below $\mathbb{M}$= 4.0 V. Following the data sheet, in 480lx light intensity (typical for office lightings), a sensor takes approximately 25 seconds to be operational and 132 seconds to be fully charged. Energy depletion is also considered to be a random variable as the relays near the sink will deplete of their energy faster than distant ones. We also consider each source node transmits in a rate

Table 3
Simulation parameters.

	Parameter	Value
Deployment	Deployment Area	1000 m × 1000 m
	Number of *source sensor* node	1500
	Source Sensor node distribution	Uniform random
Node description	Transmission range (T)	100 m
	Coverage safety margin (δ)	10 m
	Data rate	250 kbps
	Packet size	1024 bytes
	Transmission energy E_{tx}	92 mJ/packet
	Reception energy E_{tx}	45 mJ/packet
	Sleep state energy consumption E_{sleep}	15 μJ
	Time required to charge fully	132s (480lx)
	Generated voltage every 5s	1.2 V
SMRMC	Candidate location selection interval θ	36
	Renewable energy supply	Uniform random
	Activation energy for *relay* node *E*	5.4 V
	Relay node inactive below $\mathbb{M}$	4.0 V

randomly chosen from [0.05,1] packets per second. For power consumption profile of the *relay* nodes, we use the data sheet of EH-Link mote, a self-powered wireless sensor node harvesting ambient energy. The energy consumption for startup is 12 mJ, energy required for transmission is 92 mJ/data packet, 8 mJ for only processing and 15 μJ for idle state. All the simulation parameters are summarized in Table 3. Each graph data point is the average of the results from 50 simulation runs with different random seed values. We run all simulation experiments on a machine having 2.8 GHz Intel CPU and 4GB memory.

5.2. Performance metrics

We analyze the performances of the proposed SMRMC system on the following metrics.

- *Number of relay nodes deployed:* It is defined as the total number of *relay* nodes required to deploy by an algorithm to maintain network connectivity and sustainability.
- *Achievable network lifetime:* We define this metric by the total time from the initiation of the network until a *source sensor* node fails to locate a *relay* node within its communication range so as to communicate its sensed data with the sink.
- *Average hop distance to the sink:* We calculate the minimum hop distance of each *source sensor* node in the network to the sink and take the average hop distance in the network.
- *Execution time:* For each of the studied algorithms, we record the time required to produce the set(s) of *relay* node locations.

5.3. Results and discussion

In this section, we discuss the impacts of the number of *source sensor* nodes, node deployment density, network size, constant and random nature of renewable energy supplies on the performances of the studied deployment approaches.

5.3.1. Impacts of increasing node deployment density

We change the network density by incrementally placing *source sensor* nodes per unit area. Employing the proposed algorithm in various node deployment densities would shed light to the robustness of the SMRMC algorithm. Fig. 4 shows the behavior of the algorithms for increasing *source sensor* node density.

In Fig. 4(a) (with 90% confidence interval), we observe SMRMC for $k = 1$ as well as for $k = 2$, the number of *relay* nodes deployed increases gradually as the *source sensor* node density increases, which is coherent with our initial assumption. However, the number of *relay* nodes deployed reaches a steady value after certain node density (0.002); this is due to the fact that the entire area of the network has been covered by the deployed *relay* nodes and adding more *relay* nodes does not favor connectivity. The $OPT - PRIM$ and α-lifetime maximization approaches require more *relay* nodes than SMRMC for $k = 1$. The reason is two folds, *firstly*, both the $OPT - PRIM$ and α-lifetime maximization are density based that is, they cover the highest density regions first and then connect the isolated regions. Therefore, they require additional *relays* to ensure connectivity. *Secondly*, the hop optimization in $OPT - PRIM$ forces the algorithm to select the set of *relay* nodes that minimizes hop count to the furthest node in the network rather than one with minimum number of *relays* that minimizes average hop count of the network. Hence, even with the relaxed conditions on *source sensor* coverage, $OPT - PRIM$ and α-lifetime maximization performs poorly. The confidence interval on SMRMC is quite steady compared to other algorithms. This observation proves SMRMC is superior than other state-of-the art approaches in terms of varying *source sensor* densities.

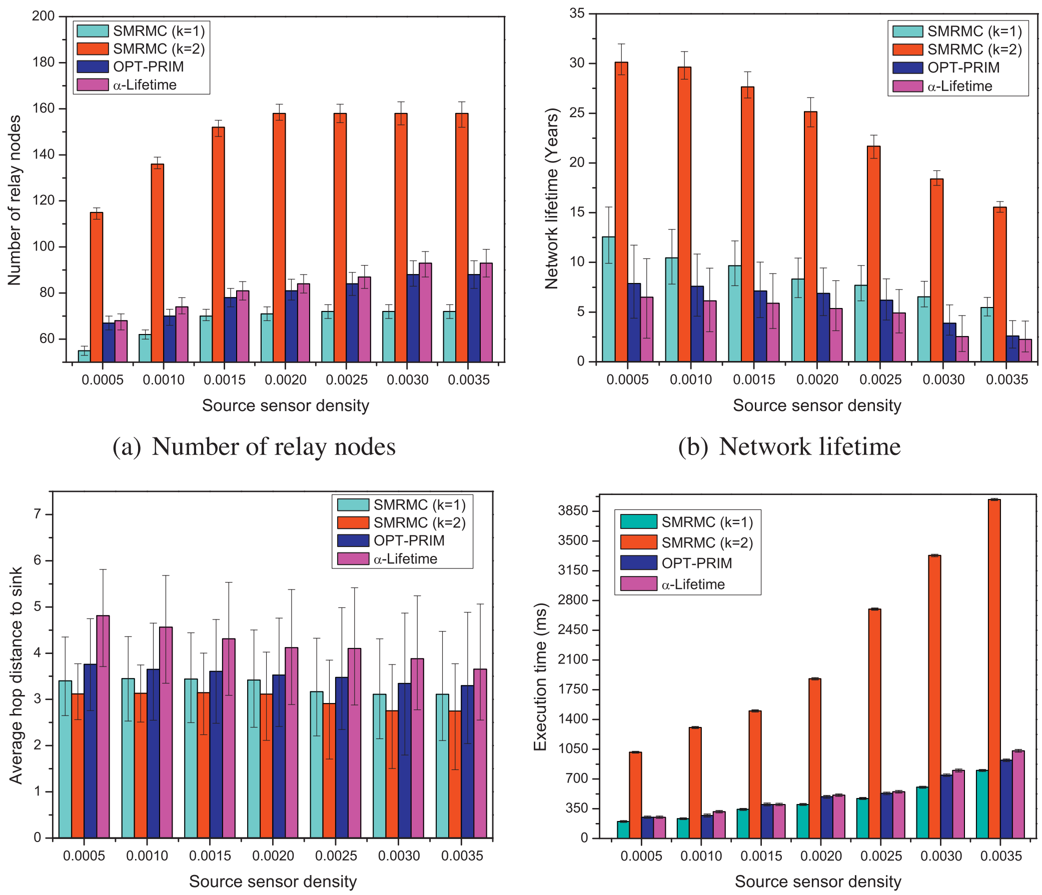
Fig. 4(b) presents the network lifetime as a function of *source sensor* node density (with 90% confidence interval). As the node density increases the algorithms convey reduction in network lifetime. This reduction is rationalized by the fact that, increased number of sources in the network results in increased network activity (packet communications, synchronizations, retransmissions etc.) hence, excelling the energy consumption and reducing the network lifetime. The SMRMC for $k = 2$ conveys highest lifetime (more than 2 times of SMRMC for $k = 1$). Since the former has the capability of load sharing among the *relays* and thereby decreases the void region. For the same number of *relay* nodes, SMRMC for $k = 1$ out performs both $OPT - PRIM$ and α-lifetime maximization. The $OPT - PRIM$ does not ensure each and every *source sensor* is covered by a *relay*. And, it prompts the *source sensors* to carry out relay responsibilities if *relay* nodes are not available. Consequently, the *source sensors* deplete of their energy and the network dies out. Situation is even worse if a *relay* near the sink goes to harvest energy. Our algorithm guarantees connectivity and thus it never imposes *relay* responsibilities to the *source sensors*, this contributes to extending network lifetime. The confidence interval for SMRMC ($k = 2$) shows that it gives more stable lifetime. Therefore, it is beneficial to have $k > 1$.

Fig. 4(c) presents the experimental results for average hop distance to the sink from the *source sensors* (with 90% confidence interval). For varying node densities the average hop distance remains steady in our algorithm. The *source sensors* are distributed with random uniform distribution, so the average hop distance depends on the *relay* nodes deployment rather than node density in the network. We observe that, for equal number of *relay* nodes our algorithm produces better average hop distance than $OPT - PRIM$. This is due to the fact that, the SMRMC can ensure connectivity with less number of *relays* than $OPT - PRIM$ and the excess nodes required by $OPT - PRIM$ is an added advantage for SMRMC.

Fig. 4(d) shows the execution time for the proposed algorithms along with $OPT - PRIM$ and α-lifetime maximization as a function of *source sensor* density over the network. For varying node densities, the average execution time of the proposed $1 - connected$ SMRMC excels others (lower execution time), whereas $k - connected$ SMRMC suffers as expected theoretically. The prime reason of consuming more time by $k - connected$ SMRMC is the requirement of extra computation cycles for polytope building, viability test of the polytopes, etc. However, in exchange, it offers greater degree of network connectivity.

5.3.2. Impacts of incremental network size

We gradually increase the dimension of the network and study its impact on to the performance of the studied node deployment



(a) Number of relay nodes (b) Network lifetime

(c) Hop distance (d) Execution time

Fig. 4. Impacts of increasing node deployment density.

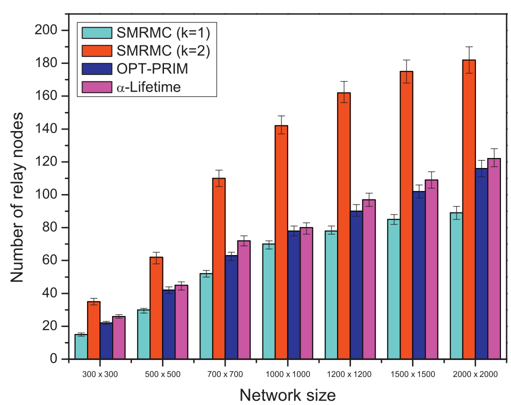
algorithms. This would help us to analyze the scalability of SMRMC system.

In Fig. 5(a), we present number of *relay* nodes that are required to deploy by each of the studied algorithms as a function of network size (with 90% confidence interval). The deployed number of *relay* nodes increases with the size of the network in all the algorithms, as expected theoretically. The SMRMC (k=1) deploys the least number of *relay* nodes and the increment is gradual. This shows SMRMC has a high tolerance for scalability than $OPT-PRIM$ and α-lifetime maximization algorithms. The reason is straight forward, SMRMC expands from the sink and deploys *relays* to the highest density in the observable region; thus, with increasing network size the observable region resembles dilation process and a gradual increase in number of *relay* nodes is encountered. However, larger the network size more scattered the *relay* node deployment scheme will be for $OPT-PRIM$ and α-lifetime maximization. Therefore, the connectivity maintenance overheard will increase with the network size as well. In Fig. 5(a), an interesting phenomenon is observed, the gap between the **1-connected SMRMC** and **2-connected SMRMC** graphs increases with the network size. This is due to the fact that the boundary area of the network remains isolated in our convex polytope approach and requires special attention to be k-connected. Larger the network, larger the isolated boundary region and more number of *relays* are required to achieve k-connectivity.

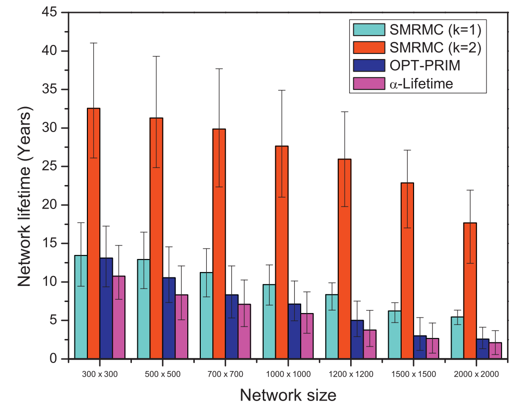
Fig. 5(b) portrays the relationship between the network size and network lifetime (with 90% confidence interval). Lifetime of the network gradually decreases as the network size increases in all the studied algorithms. In a large network, the *relays* around the sink incurs greater loads than the *relays* in a small network. To be more specific, *relays* placed near the sink carry data packets from all the other nodes of the network and when the network is large, those *relays* are burdened with huge number of data packets. Hence, they deplete of their energy quickly, causing the network to die out earlier. However, the noteworthy observation is that, as the network size increases SMRMC ($k=2$) excels in performance in comparison with SMRMC ($k=1$). With small networks (*i.e, 300* $\times$ *300*) SMRMC ($k=2$) conveys almost twice network lifetime than SMRMC ($k=1$) whereas, with larger network size (*i.e*, 2000 $\times$ 2000) SMRMC ($k=2$) delivers more than 3 times lifetime compared to SMRMC ($k=1$). This highlights the scalability of our proposed SMRMC algorithm.

Fig. 5(c) shows that the average hop count to the sink node increases as the network size increases. For fair comparison, we used the optimal number of *relay* nodes from $OPT-PRIM$ for SMRMC ($k=1$) and α-lifetime. And for SMRMC ($k=2$), we use optimal *relay* node deployment from Algorithm 2. We observe that the SMRMC ($k=2$) achieves the shortest average hop distance as there are many alternative paths, some of which will be shorter than those achieved with $k=1$. In SMRMC ($k=1$), the average hop distance is less than both the $OPT-PRIM$ and α-lifetime. This is caused by the fact that the SMRMC ($k=1$) can achieve connectivity with fewer number of nodes than the number of nodes required by $OPT-PRIM$; so the additional nodes take part in minimization of average hop distance to the sink.

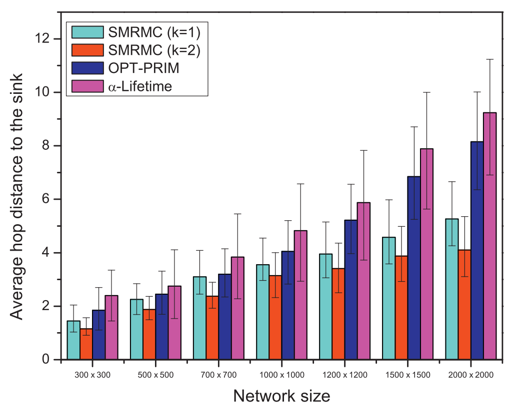
Fig. 5(d) presents the average execution time of the proposed algorithms compared to $OPT-PRIM$ and α-lifetime maximization while network size is the controlled parameter. Network size has



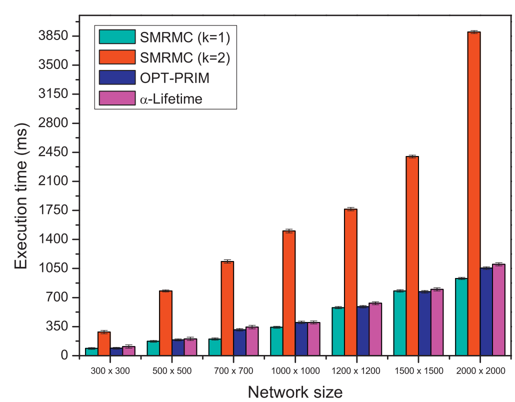
(a) Number of relay nodes



(b) Network lifetime



(c) Hop distance



(d) Execution time

Fig. 5. Impacts of increasing network size.

Algorithm 2 Generalized SMRMC ($k \geq 2$) using convex polytopes

Input: Y, S, R
Output: a list of locations, ψ for $k = 2$ *relay* nodes
Initialize ψ ={Y}, $h = 0$
while $S \neq \emptyset$ **do**
  P = GenerateNodes (ψ, R, h)
  Ψ = BuildConvexPolytopes($P_g \mid P_g \in P$)
  $h \leftarrow h + 1$
  for each $\Psi_v \in \Psi$ **do**
    if $max(dist(\{p \mid \forall p \in \Psi_v\}, \zeta_{\Psi_v})) \leq T - \delta$ **then**
      $C_v \leftarrow C_{\zeta_{\Psi_v}}$; $C_{\zeta_{\Psi_v}}$ is the coverage factor of the centroid
      $C \leftarrow \{C_v\}$
    end if
  end for
  L=AssignRelay(C)
  $S \leftarrow S \setminus C_L$; C_L is the coverage factor of *relay* locations L
  $\psi = \psi \cup L$
end while
  procedure: GenerateNodes
  Input: ψ, R, h
  Output:P
    $P_h = \{p \mid \forall\ p \subset \mathbb{P}(R)$ and $\mid p \mid = \mid R \mid - h\}$
    $P = \{P_h \times \psi\}$

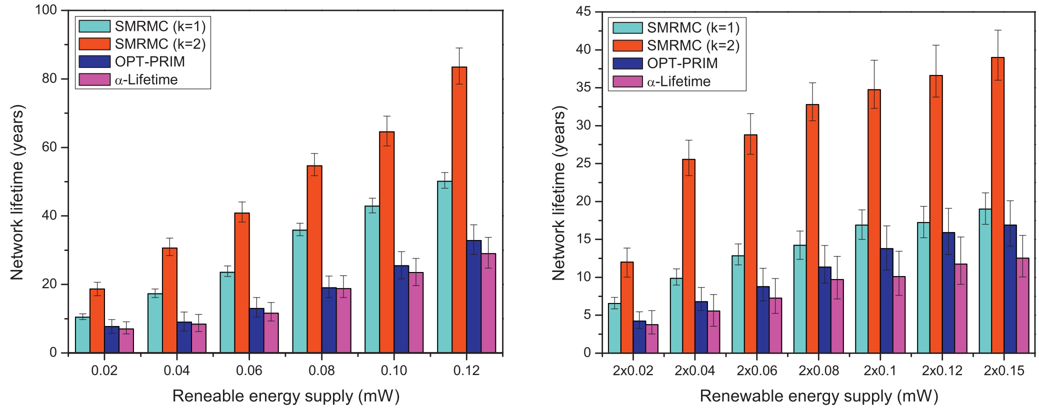
a significant impact on the execution time of the algorithms. As it increases execution time of the algorithms also increase. However, $1 - connected$ SMRMC is the fastest among all. The $k - connected$ SMRMC executes for the highest duration as it produces a $k - connected$ network of *relays* as opposed to 1-connected SMRMC. This higher degree of connectivity comes at a price of longer execution time.

5.3.3. Impacts of constant light intensity

We run our simulation for constant supply of incremental renewable energy and present the results in Fig. 6(a) (with 90% confidence interval). We allocate same amount of initial energy as well as constant energy harvesting rate. Although the energy depletion rates are variable. Energy depletion rates are depended on the amount of traffic being carried by the individual *relay* nodes. We gradually increase the constant energy supply to all the *relay* nodes in the network. We can see the increase in the network lifetime is linear. After a certain amount of constant renewable energy supply (0.12 mW/sec) the network reaches an unbounded lifetime state, *i.e.*, network lifetime goes beyond 100 years.

5.3.4. Impacts of random light intensity

We run our simulation on random supply of renewable energy and observe the results of Fig. 6(b) (with 90% confidence interval). All the *relay* nodes start with randomly assigned initial energies and they can harvest random amount of energy within a range with mean equal to the constant energy supplies of Fig. 6(a) for fair comparison. We increase the range of random renewable energy supply and observe the resultant data. Firstly, the graph depict that the network lifetime achievement is less than the lifetime achieved while the energy supplies were constant. Next, the difference is even more prominent when approaching infinite lifetime



(a) Constant energy harvesting

(b) Random energy harvesting

Fig. 6. Impacts of renewable energy harvesting on network lifetime.

(100 years) for SMRMC ($k = 2$), which is achieved with constant energy supply after (0.12 mW); however, even with random energy supply [0.005,2 × 0.15] mW, it barely achieves 40 years of lifetime. This result is more realistic than the one in 6(a) as in practice the renewable energy supplies are random, not constant.

The performance analysis of our proposed SMRMC algorithms renders the conclusion that **1-connected SMRMC** moderately minimizes the number of required *relay* nodes compared to $OPT - PRIM$ and α-lifetime minimization, which was one of the targets of this work. We analyzed the lifetime of the EH-WSNs using heterogeneous energy harvesting and expenditure behaviors of practical ambient energy harvesting nodes. The outcome shows significant increase in network lifetime and moderate decrease in average hop distance to the sink. We also present the effect of random energy harvesting on the network lifetime for real life situations. The proposed **1-connected SMRMC** has the computational complexity of $O(RR'n)$ [upper bound], which is lower than the MILP $O\binom{n}{R}$ and **2-connected SMRMC** bears $O(RR'n + n^2 h \log h)$, where, h is the number of nodes in the polytope. Here, the convex polytope building algorithm has $O(n\log h)$ computational complexity. For the substantial gain in network lifetime in **2-connected SMRMC**, such computational complexity is endurable by the pre-compilation engine.

6. Conclusion

In this work, we have investigated the problem of optimal *relay* node placement in EH-WSNs. Careful node placement can be a very effective optimization means for achieving enhanced, even perpetual lifetime in EH-WSN. We have studied the SMRMC (sustainable minimum-relay maximum-connectivity) problem of EH-WSNs and analyzed the state-of-the-art researches on optimal node deployments. The problem of optimal deployment of *relay* nodes has been formulated as an MILP so as to achieve perpetual sensor networks. The MILP optimization framework is applicable irrespective of locations of the sink and *source* nodes; however, its complexity raises to very high as the network size increases. We have also investigated greedy solutions to the k-connected SMRMC problem using convex polytopes. Through computer simulations, we show that our greedy SMRMC algorithms can provide upto two times better performance in terms of network lifetime. Thus, our solution would provide network designers and researchers with a lucrative deployment strategy for *relay* nodes in EH-WSNs.

Acknowledgments

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding of this research through the Research Group Project no. RGP-281. Special thanks to the ICT Division of the Government of Bangladesh for student fellowship.

References

[1] N. Rahman, M. Wright, D. Liu, Fast and energy-efficient technique for jammed region mapping in wireless sensor networks, ACM Computing and Research Repository (CoRR) 2014 (2014).

[2] A.Z. Abbasi, N. Islam, Z.A. Shaikh, et al., A review of wireless sensors and networks' applications in agriculture, Comput. Standards Interfaces 36 (2) (2014) 263–270.

[3] G. Xu, W. Shen, X. Wang, Applications of wireless sensor networks in marine environment monitoring: A survey, Sensors 14 (9) (2014).

[4] G. Hackmann, W. Guo, G. Yan, Z. Sun, C. Lu, S. Dyke, Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks, IEEE Trans. Parallel Distr. Syst. 25 (1) (2014) 63–72.

[5] B.A. Bakr, L.T. Lilien, Comparison by simulation of energy consumption and wsn lifetime for leach and leach-sm, Proc. Comput. Sci. 34 (2014) 180–187.

[6] W. Dron, S. Duquennoy, T. Voigt, K. Hachicha, P. Garda, An emulation-based method for lifetime estimation of wireless sensor networks, in: IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), 2014, pp. 241–248.

[7] M. Magno, D. Boyle, D. Brunelli, B. O'Flynn, E. Popovici, L. Benini, Extended wireless monitoring through intelligent hybrid energy supply, IEEE Transa. Ind. Electr. 61 (4) (2014) 1871–1881.

[8] Z.A. Eu, H.-P. Tan, W.K. Seah, Design and performance analysis of mac schemes for wireless sensor networks powered by ambient energy harvesting, Ad Hoc Netw. 9 (3) (2011) 300–323.

[9] A. Hande, T. Polk, W. Walker, D. Bhatia, Indoor solar energy harvesting for sensor network router nodes, Microprocess. Microsyst. 31 (6) (2007) 420–432.

[10] M.A. Razzaque, M.H.U. Ahmed, C.S. Hong, S. Lee, Qos-aware distributed adaptive cooperative routing in wireless sensor networks, Ad Hoc Netw. 19 (0) (2014) 28–42.

[11] C. Renner, S. Unterschütz, V. Turau, K. Römer, Perpetual data collection with energy-harvesting sensor networks, ACM Trans. Sen. Netw. 11 (1) (2014) 12:1–12:45.

[12] S. Misra, N. Majd, H. Huang, Approximation algorithms for constrained relay node placement in energy harvesting wireless sensor networks, IEEE Trans. Comput. 63 (12) (2014) 2933–2947.

[13] Z.A. Eu, H.-P. Tan, Adaptive opportunistic routing protocol for energy harvesting wireless sensor networks, in: IEEE International Conference on Communications (ICC), IEEE, 2012, pp. 318–322.

[14] D. Yang, S. Misra, X. Fang, G. Xue, J. Zhang, Two-tiered constrained relay node placement in wireless sensor networks: computational complexity and efficient approximations, IEEE Trans. Mobile Comput. 11 (8) (2012) 1399–1411.

[15] A. Abu-Baker, H. Huang, E. Johnson, S. Misra, R. Asorey-Cacheda, M. Balakrishnan, Maximizing α-lifetime of wireless sensor networks with solar energy sources, in: Military Communications Conference, MILCOM, IEEE, 2010, pp. 125–129.

[16] R. Asorey-Cacheda, A.J. García-Sánchez, F. García-Sánchez, J. García-Haro, F.J. González-Castano, On maximizing the lifetime of wireless sensor networks by optimally assigning energy supplies, Sensors 13 (8) (2013) 10219–10244.

[17] S. Liu, J. Lu, Q. Wu, Q. Qiu, Harvesting-aware power management for real-time systems with renewable energy, IEEE Trans. Very Large Scale Integr. (VLSI) Syst. 20 (8) (2012) 1473–1486.
[18] MATLAB, version 7.10.0 (R2010a), The MathWorks Inc., Natick, Massachusetts, 2010.
[19] G.-H. Lin, G. Xue, Steiner tree problem with minimum number of steiner points and bounded edge-length, Inf. Process. Lett. 69 (2) (1999) 53–57.
[20] X. Han, X. Cao, E.L. Lloyd, C.-C. Shen, Fault-tolerant relay node placement in heterogeneous wireless sensor networks, IEEE Trans. Mobile Comput. 9 (5) (2010) 643–656.
[21] A. Kashyap, S. Khuller, M.A. Shayman, Relay placement for higher order connectivity in wireless sensor networks., in: INFOCOM, 17, 2006, pp. 1–12.
[22] X. Cheng, D.-Z. Du, L. Wang, B. Xu, Relay sensor placement in wireless sensor networks, Wireless Netw. 14 (3) (2008) 347–355.
[23] J.L. Bredin, E.D. Demaine, M. Hajiaghayi, D. Rus, Deploying sensor networks with guaranteed capacity and fault tolerance, in: Proceedings of the 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing, ACM, 2005, pp. 309–319.
[24] S. Misra, S.D. Hong, G. Xue, J. Tang, Constrained relay node placement in wireless sensor networks: Formulation and approximations, IEEE/ACM Trans. Netw. (TON) 18 (2) (2010) 434–447.
[25] T. La Porta, C. Petrioli, D. Spenza, Sensor-mission assignment in wireless sensor networks with energy harvesting, in: 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), IEEE, 2011, pp. 413–421.
[26] Z.A. Eu, H.-P. Tan, W.K. Seah, Opportunistic routing in wireless sensor networks powered by ambient energy harvesting, Comput. Netw. 54 (17) (2010) 2943–2966.
[27] M.H. bin Omar, M.H. Lim, Ambient energy harvesting for wireless sensor nodes: a review, in: Advanced Materials Research, 1025, Trans Tech Publ, 2014, pp. 1093–1098.
[28] W. Seah, Z.A. Eu, H. Tan, Wireless sensor networks powered by ambient energy harvesting (wsn-heap) - survey and challenges, in: 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology. Wireless VITAE, 2009, pp. 1–5.
[29] M.A. Razzaque, C.S. Hong, Analysis of energy-tax for multipath routing in wireless sensor networks, Ann. Telecommun. 65 (1-2) (2010) 117–127.
[30] V. Raghunathan, P.H. Chou, Design and power management of energy harvesting embedded systems, in: Proceedings of the 2006 International Symposium on Low Power Electronics and Design, ACM, 2006, pp. 369–374.
[31] G. Zhou, T. He, S. Krishnamurthy, J.A. Stankovic, Models and solutions for radio irregularity in wireless sensor networks, ACM Trans. Sen. Netw. 2 (2) (2006) 221–262.
[32] B. Dezfouli, M. Radi, S.A. Razak, T. Hwee-Pink, K.A. Bakar, Modeling low-power wireless communications, J. Netw. Comput. Appl. 51 (2015) 102–126.
[33] R. Jarvis, On the identification of the convex hull of a finite set of points in the plane, Inf. Process. Lett. 2 (1) (1973) 18–21.
[34] EH-Link technical note, MicroStrain, 2010. Rev. 01.
[35] V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, M. Srivastava, Design considerations for solar energy harvesting wireless embedded systems, in: Proceedings of the 4th International Symposium on Information Processing in Sensor Networks, in: IPSN '05, IEEE Press, Piscataway, NJ, USA, 2005.

Nusrat Mehajabin completed her B.S. and M.S. from the Department of Computer Science and Engineering, University of Dhaka, Bangladesh in 2013 and 2015, respectively. Her areas of research interest include Wireless Sensor Networks, Energy Optimization, Wi-Fi Networks, modeling and analysis of communication algorithms, etc. She is a student member of IEEE.

Md. Abdur Razzaque received his B.S. and M.S. degrees from the University of Dhaka, Bangladesh in 1997 and 1999, respectively. He obtained his Ph.D. degree in Wireless Networking from the Department of the Computer Engineering, School of Electronics and Information, Kyung Hee University, South Korea in 2009. He was a research professor, in the same university during 2010–2011. He is now working as a Professor in the Department of Computer Science and Engineering, University of Dhaka, Bangladesh. He is the group leader of Green Networking Research Group (http://gnr.cse.univdhaka.edu) of the same department. His research interest is in the area of modeling, analysis and optimization of wireless networking protocols and architectures, Wireless Sensor Networks, Body Sensor Networks, Cooperative Communications, Sensor Data Clouds, Internet of Things, Cognitive Radio Networks, etc. He has published a good number of research papers in international conferences and journals. He is an editorial board member of International Journal of Distributed Sensor Networks, TPC member of IEEE HPCC 2013–2015, ICOIN 2010–2015, ADM 2014, NSysS 2015. He is a senior member of IEEE, member of IEEE Communications Society, IEEE Computer Society, Internet Society (ISOC), Pacific Telecommunications Council (PTC) and KIPS.

Mohammad Mehedi Hassan is an Assistant Professor of Information Systems Department in the College of Computer and Information Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia. He received his Ph.D. degree in Computer Engineering from Kyung Hee University, South Korea in February 2011. He was a Research Professor at Computer Engineering department, Kyung Hee University, South Korea from March, 2011 to October, 2011. He has authored and co-authored more than 50 publications including refereed IEEE/ACM/Springer journals, conference papers, books, and book chapters. He has served as, chair, and Technical Program Committee member in numerous international conferences/workshops like IEEE HPCC, IEEE ICME, ACM Multimedia, ICA3PP, IEEE ICC, TPMC, IDCS, etc. His research interests include Cloud collaboration, multimedia Cloud, sensor-Cloud, mobile Cloud, Thin-Client, Grid computing, IPTV, virtual network, sensor network, and publish/subscribe system.

Ahmad Almogren is an Associate Professor of Information Systems Department in the College of Computer and Information Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia. He received his M.S. in 1995 and Ph.D. in 2002 from Computer Science, Southern Methodist University, Dallas, Texas, USA. He completed his B.S. from King Saud University, Kingdom of Saudi Arabia in 1992. He has authored and co-authored more than 70 publications including refereed IEEE/ACM/Springer journals, conference papers, books, and book chapters. He has served as, chair, and Technical Program Committee member in numerous international conferences/workshops. His research interests include mobile, cloud and pervasive computing, wireless sensor network, computer security, and data mining.

Atif Alamri is an Assistant Professor of Information Systems Department in the College of Computer and Information Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia. He received his Ph.D. degree from School of Information Technology and Engineering, University of Ottawa, Canada in 2010. He completed his B.S. and M.S. degrees from College of Computer and Information Science in King Saud University, Kingdom of Saudi Arabia. He has authored and co-authored more than 40 publications including refereed IEEE/ACM/Springer journals, conference papers, books, and book chapters. He has served as, chair, and Technical Program Committee member in numerous international conferences/workshops. His research interests include Cloud computing, Sensor networks, multimedia Cloud, sensor-Cloud, mobile Cloud, etc.