# Brief Contributions

## Relay Node Placement in Wireless Sensor Networks

Errol L. Lloyd and
Guoliang Xue, *Senior Member, IEEE*

**Abstract**—A wireless sensor network consists of many low-cost, low-power *sensor nodes*, which can perform sensing, simple computation, and transmission of sensed information. Long distance transmission by sensor nodes is not energy efficient since energy consumption is a superlinear function of the transmission distance. One approach to prolonging network lifetime while preserving network connectivity is to deploy a small number of costly, but more powerful, *relay nodes* whose main task is communication with other sensor or relay nodes. In this paper, we assume that sensor nodes have communication range $r > 0$, while relay nodes have communication range $R \geq r$, and we study two versions of *relay node placement* problems. In the first version, we want to deploy the minimum number of relay nodes so that, between each pair of sensor nodes, there is a connecting path consisting of relay and/or sensor nodes. In the second version, we want to deploy the minimum number of relay nodes so that, between each pair of sensor nodes, there is a connecting path consisting solely of relay nodes. We present a polynomial time 7-approximation algorithm for the first problem and a polynomial time $(5+\epsilon)$-approximation algorithm for the second problem, where $\epsilon > 0$ can be any given constant.

**Index Terms**—Relay node placement, wireless sensor networks, approximation algorithms.

## 1 Introduction

A wireless sensor network (WSN) consists of a large number of low-cost, low-power *sensor nodes*, which can perform sensing, simple computation, and communication over short distances [1], [6]. Since sensors are powered by batteries and are usually deployed outdoors in harsh environments, extensive research has been focused on energy aware routing [4], [11], [13], network lifetime [10], [16], and survivability [9], [15], [18].

With the current technology, long distance transmission in WSNs is very costly since energy consumption is proportional to $d^\kappa$ for transmitting over distance $d$, where $\kappa$ is a constant in the interval $[2, 4]$, depending on the media. One approach to prolonging the network lifetime while preserving network connectivity is to deploy a small number of costly, but more powerful *relay nodes* whose main task is communication with the sensor nodes and with other relay nodes. This is the general topic addressed in this paper.

### 1.1 Earlier Work on Relay Node Placement

Cheng et al. [3] proposed deploying a minimum number of relay nodes in a WSN so that, between every pair of sensor nodes, there is a connecting path consisting of relay and/or sensor nodes and such that each hop of the path is no longer than the *common* transmission range of the sensor nodes and the relay nodes. This problem is exactly the *Steiner minimum tree with minimum number of Steiner points and bounded edge length* problem (SMT-MSPBEL) defined by Lin and Xue in the study of amplifier placement in wide area optical networks [14]. Lin and Xue [14] proved that the SMT-MSPBEL problem is NP-hard and presented a simple minimum spanning tree (MST) based 5-approximation algorithm.[1] In [2], Chen et al. proved that the Lin-Xue algorithm is actually a 4-approximation algorithm. They also presented a 3-approximation algorithm for this problem. In [3], Cheng et al. presented a faster 3-approximation algorithm, and a randomized algorithm[2] with an approximation ratio of 2.5.

In [16], Pan et al. studied a *two-tiered* network model where the sensor nodes are grouped into clusters, each covered by an *application node*. The sensor nodes transmit sensed information to the application node, which then processes the received information and sends the processed information to the base station. In [9], [19], relay node placement was studied in a two-tiered WSN under the assumption that the sensor nodes have a communication range $r > 0$ and the relay nodes have a communication range $R \geq 4r$. They studied two problems. For the *connected relay node single cover* (CRNSC) problem, they aimed to deploy a minimum number of relay nodes so that 1) every sensor node is within distance $r$ of a relay node and 2) between every pair of relay nodes, there is a connecting path consisting of relay nodes such that each hop of the path is not longer than $R$. For the *2-connected relay node double cover* problem, they aimed to deploy a minimum number of relay nodes so that 1) every sensor node is within distance $r$ of two relay nodes and 2) between every pair of relay nodes, there are two node-disjoint connecting paths consisting of relay nodes such that each hop of the paths is not longer than $R$. Under the assumption that *the sensor nodes are uniformly distributed* in the playing field and that $R \geq 4r$, the authors of [19] presented 4.5-approximation algorithms for both problems.

### 1.2 The Problems We Study

In this paper, we study two versions of relay node placement when sensor and relay nodes have different communication ranges.

- The *single-tiered relay node placement problem* is a generalization of the SMT-MSPBEL problem where the sensor nodes have communication range $r$ and the relay nodes have communication range $R \geq r$. That is, we seek to deploy a minimum number of relay nodes such that, between every pair of sensor nodes, there is a path consisting of relay and/or sensor nodes, where consecutive nodes on that path are within distance $R$ if both are relay nodes and within distance $r$ otherwise. For this NP-hard problem [8], [14], we present a polynomial time 7-approximation algorithm.
- The *two-tiered relay node placement problem* is the general case of the CRNSC problem (i.e., without the sensor distribution and the $R \geq 4r$ constraints). That is, we seek to deploy a minimum number of relay nodes such that, between every pair of sensor nodes, there is a path consisting solely of relay nodes, where the sensor nodes on either end of that path are within distance $r$ of the adjacent relay node on the path and successive relay nodes on the path are within distance $R$ of one another. For this problem, we present a general framework which combines any $\alpha$-approximation algorithm for the

- E.L. Lloyd is with the Department of Computer and Information Sciences, University of Delaware, Newark, DE 19716. E-mail: elloyd@udel.edu.
- G. Xue is with the Department of Computer Science and Engineering, Arizona State University, Tempe, AZ 85287-8809. E-mail: xue@asu.edu.

*Manuscript received 31 Dec. 2005; revised 8 May 2006; accepted 30 May 2006; published online 22 Nov. 2006.*
*For information on obtaining reprints of this article, please send e-mail to: tc@computer.org, and reference IEEECS Log Number TC-0462-1205.*

1. An algorithm solving a minimization problem is an $\alpha$-approximation algorithm (or has approximation ratio $\alpha$) if the solution provided by the algorithm is no more than $\alpha$ times the optimal solution [8]. Additional details are given in Section 2.

2. A randomized algorithm with an approximation ratio of $\alpha$ for a minimization problem is an algorithm that provides a solution no more than $\alpha$ times the optimal solution with a positive probability. Additional details are given in Section 2.

*minimum geometric disk cover* problem [12] and any $\beta$-approximation algorithm for the SMT-MSPBEL problem to obtain a $(2\alpha+\beta)$-approximation algorithm for this two-tiered relay node placement problem. Using the currently best-known values for $\alpha$ and $\beta$, our framework provides a $(5+\epsilon)$-approximation algorithm and a randomized $(4.5+\epsilon)$-approximation algorithm, where $\epsilon$ is any given positive constant.

Our work on single-tiered relay node placement is different from previous works because the problem studied in previous works [14], [2], [3] is a special case $(R=r)$ of the problem studied in this paper. Our work on two-tiered relay node placement is different from previous works because we do not make any assumption on sensor node distribution and do not require the condition $R \geq 4r$ as in [9], [19].

The remainder of this paper is organized as follows: In Section 2, we formally define the problems to be studied, as well as some related problems and notations that will be used in this paper. In Section 3, we present our approximation algorithm for the single-tiered relay node placement problem. In Section 4, we present our approximation framework and subsequent approximation algorithm for the two-tiered relay node placement problem. We conclude this paper in Section 5 with some future research directions.

## 2 Problem Formulations and Background

In this section, we formally define the problems and notations that will be used throughout the paper. We refer readers to [5] for graph theoretic notations not defined here and to [8], [5] for definitions such as "NP-hard" and other concepts in complexity theory that are not defined here.

### 2.1 Complexity Related Definitions

*A polynomial time $\alpha$-approximation algorithm* for a minimization problem is an algorithm $\mathcal{A}$ that, for any instance of the problem, computes a solution that is at most $\alpha$ times the optimal solution of the instance, in time bounded by a polynomial in the input size of the instance [5]. In this case, we also say that *$\mathcal{A}$ has an approximation ratio of $\alpha$*. $\mathcal{A}_\epsilon$ is a *polynomial time approximation scheme* (PTAS) for a minimization problem if, for any fixed $\epsilon > 0$, $\mathcal{A}_\epsilon$ is a polynomial time $(1+\epsilon)$-approximation algorithm with $\epsilon$ treated as a constant.

*A randomized $\alpha$-approximation algorithm* for a minimization problem is an algorithm $\mathcal{A}$ that, for any instance of the problem, *with probability greater than a positive constant*, computes a solution that is at most $\alpha$ times the optimal solution of the instance. Although the algorithm does not necessarily guarantee providing such a solution, by repeated application the probability can be made arbitrarily close to 1.

### 2.2 Relay Node Terminology and Preliminaries

The general class of problems that we consider is as follows: We are given two positive real constants, $r > 0$ and $R \geq r$, where $r$ is the *communication range* of a *sensor node* and $R$ is the communication range of a *relay node*. In addition, we are given a set $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ of $n$ sensor nodes on the Euclidean plane and the goal is to deploy sufficiently many relay nodes so as to ensure network connectivity. Without confusion, we will also use $x_i$ to denote the *location* of sensor node $x_i$, $i = 1, 2, \ldots, n$.

We utilize the following notation and terminology. Let $p$ and $q$ be two points in the plane, then $[p, q]$ denotes the *line segment* connecting $p$ and $q$ and $\|p\, q\|$ denotes the *Euclidean distance* between $p$ and $q$. Two sensor nodes $x_i$ and $x_j$ can communicate with each other if $\|x_i x_j\| \leq r$. A sensor node $x_i$ and a relay node $y_j$ can communicate with each other if $\|x_i\, y_j\| \leq r$. Two relay nodes, $y_i$ and $y_j$, can communicate with each other if $\|y_i\, y_j\| \leq R$. Two nodes are said to be *neighbors* if they can communicate with each other.

We study relay node placement in two kinds of WSNs. In the first, named *single-tiered WSN*, both sensor nodes and relay nodes can receive packets from a neighbor node and forward packets to a neighbor node. In the second, named *two-tiered WSN*, relay nodes can receive and forward packets, while sensor nodes do not forward packets they receive (they only transmit sensed information to the relay nodes). Correspondingly, we study two kinds of relay node placement problems. Before defining these relay node placement problems, we will first define a variant of the well-known Euclidean Steiner tree problem [8]. This problem and related terminology will be extensively used in our definitions and proofs.

**Definition 2.1 (Constrained Steiner Trees).** *Let $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ be a set of $n$* target points *in the Euclidean plane. Let $R \geq r > 0$ be two positive real constants. An $(r, R)$-*constrained Steiner tree of *$\mathcal{X}$ is a tree $T$ spanning the set $\mathcal{X}$ of target points and an additional set* $\mathrm{Steiner}(T)$ *of* Steiner points *such that 1) each edge connecting a target point with a target/Steiner point has length no more than $r$ and 2) each edge connecting a Steiner point with another Steiner point has length no more than $R$. For each Steiner point $u$ of $T$, the* Steiner degree *of $u$ in $T$ (denoted by $d_s(u)$) is the number of Steiner points incident with $u$ and the* target degree *of $u$ in $T$ (denoted by $d_t(u)$) is the number of target points incident with $u$. The* size *of an $(r, R)$-constrained Steiner tree $T$ is the number of Steiner points in $T$, namely, $|\mathrm{Steiner}(T)|$. The* Steiner length *of an $(r, R)$-constrained Steiner tree $T$ is the sum of edge lengths over all Steiner-target and Steiner-Steiner edges. Note that target-target edges do not contribute to the Steiner length. An $(r, R)$-constrained Steiner tree of $\mathcal{X}$ with the minimum size is called a* minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$ *(denoted by $\mathrm{MCST}(\mathcal{X}, r, R)$). A* shortest minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$ *is a minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$ with the shortest Steiner length. Note that finding a minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$ is an extension of the SMT-MSPBEL problem to two edge lengths.*

### 2.3 Single and Two Tiered RNPs

Now, we are ready to define the two kinds of relay node placement problems studied in this paper.

**Definition 2.2 (Single-Tiered Relay Node Placement).** *Let $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ be a set of sensor nodes with known locations. Let $r > 0$ and $R \geq r$ be the communication ranges for sensor nodes and relay nodes, respectively. A set of relay nodes $\mathcal{Y} = \{y_1, y_2, \ldots, y_m\}$ is said to be a* feasible single-tiered relay node placement *(F1tRNP) for $(\mathcal{X}, r, R)$ if $\mathcal{Y}$ is the set of Steiner points of an $(r, R)$-constrained Steiner tree $T$ of $\mathcal{X}$, (i.e., $\mathcal{Y} = \mathrm{Steiner}(T)$). The* size *of the corresponding F1tRNP is $|\mathcal{Y}|$. An F1tRNP is said to be a* minimum single-tiered relay node placement *(M1tRNP) for $(\mathcal{X}, r, R)$ (also denoted by $\mathrm{M1tRNP}(\mathcal{X}, r, R)$) if it has the minimum size among all F1tRNPs for $(\mathcal{X}, r, R)$. The* single-tiered relay node placement problem *for $(\mathcal{X}, r, R)$, denoted by $\mathrm{1tRNP}(\mathcal{X}, r, R)$, seeks a minimum single-tiered relay node placement for $(\mathcal{X}, r, R)$.*

**Definition 2.3 (Two-Tiered Relay Node Placement).** *Let $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ be a set of sensor nodes with known locations. Let $r > 0$ and $R \geq r$ be the communication ranges for sensor nodes and relay nodes, respectively. A set of relay nodes $\mathcal{Y} = \{y_1, y_2, \ldots, y_m\}$ is said to be a* feasible two-tiered relay node placement *(F2tRNP) for $(\mathcal{X}, r, R)$ if*

1. *for each sensor node $x_i \in \mathcal{X}$, there is a relay node $y_j \in \mathcal{Y}$ such that $\|x_i\, y_j\| \leq r$;*
2. *the undirected graph $G(\mathcal{Y}, R)$ is connected, where the vertex set of $G$ is $V = \mathcal{Y}$ and the edge set of $G$ is $E = \{(y_i, y_j) | y_i, y_j \in \mathcal{Y}, \|y_i\, y_j\| \leq R\}$.*

*We call $|\mathcal{Y}|$ the* size *of the corresponding F2tRNP. An F2tRNP is said to be a* minimum two-tiered relay node placement

*(**M2tRNP**) for $(\mathcal{X}, r, R)$ (also denoted by $\mathrm{M2tRNP}(\mathcal{X}, r, R)$) if it has the* minimum size *among all **F2tRNPs** for $(\mathcal{X}, r, R)$. The* two-tiered relay node placement problem *for $(\mathcal{X}, r, R)$, denoted by $\mathrm{2tRNP}(\mathcal{X}, r, R)$, seeks a minimum two-tiered relay node placement for $(\mathcal{X}, r, R)$.*

### 2.4 The Complexity of Relay Node Placement

The problem $\mathrm{1tRNP}(\mathcal{X}, r, R)$ is easily seen to be NP-hard by noting that, when $r = R$, $\mathrm{1tRNP}(\mathcal{X}, r, R)$ is identical to the SMT-MSPBEL problem [14], which is a known NP-hard problem [14]. The best-known polynomial time approximation algorithm for SMT-MSPBEL has an approximation ratio of 3 [2], [3]. Cheng et al. [3] presented a *randomized algorithm* for SMT-MSPBEL with an approximation ratio of 2.5. Obviously, those same bounds hold for $\mathrm{1tRNP}(\mathcal{X}, R, R)$. However, to the best of our knowledge, there is no constant ratio approximation algorithm for the general $\mathrm{1tRNP}(\mathcal{X}, r, R)$ problem.

The NP-hardness of $\mathrm{2tRNP}(\mathcal{X}, r, R)$ can be established using the following problem, which was shown to be NP-hard in [7]:

**Definition 2.4 (Geometric Disk Cover).** *Let $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ be a set of points in the Euclidean plane and let $r > 0$ be a positive constant. The* minimum geometric disk cover *problem for $(\mathcal{X}, r)$ (denoted by $\mathrm{DCover}(\mathcal{X}, r)$) seeks a minimum cardinality set of points $\mathcal{C} = \{c_1, c_2, \ldots, c_m\}$ with the property that, for each point $x_i \in \mathcal{X}$, there exists a point $c_j \in \mathcal{C}$ such that $\|x_i\ c_j\| \le r$. We call $\mathcal{C}$ an* optimal solution *to $\mathrm{DCover}(\mathcal{X}, r)$ and use $\mathrm{MDCover}(\mathcal{X}, r)$ to denote $\|\mathcal{C}\|$. A set of points $\mathcal{C} = \{c_1, c_2, \ldots, c_k\}$ is said to be a* feasible solution *to $\mathrm{DCover}(\mathcal{X}, r)$ if, for each point $x_i \in \mathcal{X}$, there exists a point $c_j \in \mathcal{C}$ such that $\|x_i\ c_j\| \le r$.*

With this definition in hand, the following is a straightforward reduction from $\mathrm{DCover}(\mathcal{X}, r)$ to $\mathrm{2tRNP}(\mathcal{X}, r, R)$. Let an instance $I_1$ of $\mathrm{DCover}(\mathcal{X}, r)$ be given by $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ and $r > 0$. Construct an instance $I_2$ of $\mathrm{2tRNP}(\mathcal{X}, r, R)$ by setting $R = 2r + \max_{1 \le i < j \le n} \|x_i\ x_j\|$. It is easily seen that a set of points in the plane is an optimal solution to $I_1$ if and only if it is an optimal solution to $I_2$. Therefore, $\mathrm{2tRNP}(\mathcal{X}, r, R)$ is NP-hard. Tang et al. in [19] presented a 4.5-approximation algorithm for $\mathrm{2tRNP}(\mathcal{X}, r, R)$ under the constraints that $R \ge 4r$ and that the sensor nodes $\mathcal{X}$ are uniformly distributed. To the best of our knowledge, there is no constant ratio approximation algorithm for the general $\mathrm{2tRNP}(\mathcal{X}, r, R)$ problem. For the $\mathrm{DCover}(\mathcal{X}, r)$ problem, there exists a polynomial time approximation scheme (PTAS) due to Hochbaum and Maass [12].

## 3 Single-Tiered Relay Node Placement

In this section, we present a simple minimum spanning tree (MST) based approximation algorithm for 1tRNP and prove that the number of relay nodes generated by this algorithm is no more than $7 \times |\mathrm{M1tRNP}(\mathcal{X}, r, R)|$, where $\mathrm{M1tRNP}(\mathcal{X}, r, R)$ is any minimum single-tiered relay node placement. Therefore, we have a 7-approximation algorithm for 1tRNP.

Given a set of sensor nodes, $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$, and constants, $r > 0$ and $R \ge r$, as the sensor node communication range and relay node communication range, respectively, our MST-based algorithm first computes an MST of $\mathcal{X}$, denoted by $T_{mst}$. It then *steinerizes* [14] that MST to obtain an $(r, R)$-constrained Steiner tree of $\mathcal{X}$ (hence, an F1tRNP for $(\mathcal{X}, r, R)$) by placing relay nodes on the line segment $[x_i, x_j]$ for each edge $e = (x_i, x_j)$ in $T_{mst}$. The complete algorithm is formally presented as Algorithm 1.

**Algorithm 1.** MST-1tRNP

INPUT: A set of $n$ sensor nodes $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ in the Euclidean plane and positive real constants $r > 0$ and $R \ge r$ representing the communication ranges of sensor nodes and relay nodes, respectively.

OUTPUT: A set $\mathcal{Y} = \{y_1, \ldots, y_k\}$ of relay nodes.

step_1 Let $V = \{x_1, x_2, \ldots, x_n\}$. Let $G = G(V, E, w)$ be an undirected edge-weighted complete graph on the vertices in $V$, where, for $1 \le i < j \le n$, the weight of edge $(x_i, x_j)$ is $w(x_i, x_j) = \|x_i\ x_j\|$.

step_2 Compute a minimum spanning tree $T_{mst}$ of $G$. Set $k := 0$.

step_3 **for** each edge $(x_i, x_j) \in T_{mst}$ **do**
  **if** $(r < \|x_i\ x_j\| \le 2r)$ **then**
    $k := k + 1$. Place a relay node $y_k$ at the midpoint of the line segment $[x_i, x_j]$.
  **elseif** $(2r < \|x_i\ x_j\|)$ **then**
    Place two relay nodes on the line segment $[x_i, x_j] : y_{k+1}$ with distance $r$ to $x_i$, and $y_{k+2}$ with distance $r$ to $x_j$.
    Place $\lceil \frac{\|x_i\ x_j\| - 2r}{R} \rceil - 1$ relay nodes $y_{k+3}, y_{k+4}, \ldots, y_{k+\lceil \frac{\|x_i\ x_j\| - 2r}{R} \rceil + 1}$ on the line segment $[y_{k+1}, y_{k+2}]$, separating the line segment $[y_{k+1}, y_{k+2}]$ into $\lceil \frac{\|x_i\ x_j\| - 2r}{R} \rceil$ equal parts.
    $k := k + 2 + \lceil \frac{\|x_i\ x_j\| - 2r}{R} \rceil - 1$.
  **endif**
**endfor**

step_4 Output $\mathcal{Y} = \{y_1, y_2, \ldots, y_k\}$

Note that, in Algorithm 1, since $T_{mst}$ is an MST of $G$ and the set $\mathcal{Y}$ is obtained by steinerizing $T_{mst}$, it follows that, in step_3, the algorithm implicitly computes an $(r, R)$-constrained Steiner tree $\mathcal{T}^A$ of $\mathcal{X}$, with $\mathcal{Y} = \mathrm{Steiner}(\mathcal{T}^A)$. In the proof that follows, we will find it convenient to refer directly to $\mathcal{T}^A$ and to $\mathrm{Steiner}(\mathcal{T}^A)$, rather than to $\mathcal{Y}$.

**Theorem 3.1.** *The $(r, R)$-constrained Steiner tree $\mathcal{T}^A$ computed in Algorithm 1 is such that*

$$|Y| = |\mathrm{Steiner}(\mathcal{T}^A)| \le 7 \times |\mathrm{MCST}(\mathcal{X}, r, R)| = 7 \times |\mathrm{M1tRNP}|.$$

We need to prove a sequence of lemmas before proving Theorem 3.1. Note that, in $\mathcal{T}^A$, *every Steiner point $u$ has degree (measured as $d_t(u) + d_s(u)$) exactly* 2.

**Lemma 3.1.** *Let $T$ be a $(r, R)$-constrained Steiner tree of $\mathcal{X}$ such that every Steiner point has degree exactly equal to 2. Then, $|\mathrm{Steiner}(T)| \ge |\mathrm{Steiner}(\mathcal{T}^A)|$.*

**Proof.** A path $\pi$ in $T$ or $\mathcal{T}^A$ is called a *superedge* if the two end nodes of $\pi$ are both target points and every interior node of $\pi$ (if any) is a Steiner point. From the definition of $(r, R)$-constrained Steiner tree we have the following facts: If a superedge contains $\eta \le 1$ Steiner points, the Euclidean distance between its two end nodes is no more than $(\eta + 1)r$. If a superedge contains $\eta \ge 2$ Steiner points, the Euclidean distance between its two end nodes is no more than $2r + (\eta - 1)R$. Note that *the minimum number of Steiner points required on a superedge is a nondecreasing function of the Euclidean distance between the two end nodes of the superedge*.

Let $\underline{T}$ ($\underline{\mathcal{T}^A}$) be a spanning tree of $\mathcal{X}$ containing exactly those edges $(x_i, x_j)$ such that there is a path in $T$ ($\mathcal{T}^A$) from $x_i$ to $x_j$, where all of the interior nodes are Steiner points. Note that $\underline{\mathcal{T}^A} = T_{mst}$ is a minimum spanning tree of $\mathcal{X}$. From the (matroid) property of minimum spanning trees, we know that $\underline{\mathcal{T}^A}$ can be obtained from $\underline{T}$ by a sequence of operations in each of which we replace an edge $e$ in $\underline{T}$ by an edge $e^A$ in $\underline{\mathcal{T}^A}$ such that the length of $e$ (the Euclidean distance between its two end nodes) is no less than the length of $e^A$. Since the minimum number of Steiner points on a superedge is monotonically nondecreasing in the length of a superedge, we have $|\mathrm{Steiner}(T)| \ge |\mathrm{Steiner}(\mathcal{T}^A)|$. □

**Lemma 3.2.** *Let $T$ be any $(r, R)$-constrained Steiner tree of $\mathcal{X}$. Then, the sum of Steiner degrees over all Steiner points of $T$ is no more than $2k - 2$, where $k = |\mathrm{Steiner}(T)|$.*

**Proof.** Since there are $k$ Steiner points in $T$, there are at most $k - 1$ Steiner-Steiner edges (edges connecting two Steiner points) in $T$. Each Steiner edge contributes exactly two Steiner degrees. Therefore, the sum of Steiner degrees over all Steiner points of $T$ is no more than $2k - 2$. □

Next, we say that an M1tRNP is a *shortest minimum single-tiered relay node placement* for $(\mathcal{X}, r, R)$ if it is the set of Steiner points of a shortest minimum $(r, R)$-constrained Steiner tree $T$ of $\mathcal{X}$.

**Lemma 3.3.** *Let $T$ be any shortest minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$. Then,*

1. *Let $u$ be a Steiner point of $T$. Let $x_i$ and $x_j$ be two target points incident with $u$ in $T$. Then, $\angle x_i u x_j > 60°$.*
2. *Every Steiner point $u$ of $T$ has target degree no more than 5.*

**Proof.** To prove 1, note that, if $\angle x_i u x_j \leq 60°$, we have $\|x_i\, x_j\| \leq \max\{\|u\, x_i\|, \|u\, x_j\|\} \leq r$. Therefore, we can replace $(u, x_i)$ by $(x_i, x_j)$ to obtain a minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$ with a *shorter Steiner length*, contradicting the assumption of $T$.

That every Steiner point $u$ of $T$ has target degree no more than 5 follows from 1. □

**Lemma 3.4.** *There exists an $(r, R)$-constrained Steiner tree $T$ of $\mathcal{X}$ such that*

1. $|\mathrm{Steiner}(T)| \leq 7 \times |\mathrm{MCST}(\mathcal{X}, r, R)|$.
2. *Each Steiner point of $T$ has degree exactly 2.*

**Proof.** Let $T^{opt}$ be a shortest minimum $(r, R)$-constrained Steiner tree of $\mathcal{X}$ such that each Steiner point has target degree at most 5. Since every tree is a planar graph, we consider a layout of $T^{opt}$ in the plane. Starting from a leaf node of the tree and taking a clockwise walk of the tree, we obtain an Eulerian tour. Note that each Steiner point $u$ of $T^{opt}$ is used exactly $d_t(u) + d_s(u)$ times by this Eulerian tour. Define a graph whose vertices are the target points and where there is an edge between vertices $x_i$ and $x_j$ if there is a superedge between $x_i$ and $x_j$ in the Eulerian tour. Since the graph is induced from the Eulerian tour, it is connected. Therefore, we can obtain a spanning tree of this graph which spans all of the target points. By replacing each edge of this tree with the corresponding superedge from the Eulerian tour, we obtain an $(r, R)$-constrained Steiner tree of $\mathcal{X}$, with no more than $\sum_{u \in \mathrm{Steiner}(T^{opt})}(d_t(u) + d_s(u))$ Steiner points. It follows from Lemma 3.2 that

$$\sum_{u \in \mathrm{Steiner}(T^{opt})} d_s(u) \leq 2 \times |\mathrm{Steiner}(T^{opt})|.$$

It follows from Lemma 3.3 that

$$\sum_{u \in \mathrm{Steiner}(T^{opt})} d_t(u) \leq 5 \times |\mathrm{Steiner}(T^{opt})|.$$

Together, these two facts prove the lemma. □

**Proof of Theorem 3.1.** Combining Lemma 3.1 and Lemma 3.4, it follows directly that $|\mathrm{Steiner}(\mathcal{T}^A)| \leq 7 \times |\mathrm{MCST}(\mathcal{X}, r, R)|$, from which the theorem follows. □

Theorem 3.1 shows that 7 is an upper bound of the approximation ratio of Algorithm 1. The following example shows that 6 is a lower bound of the approximation ratio of Algorithm 1. Assume that $R \geq 5r$. Let $K$ be any positive integer; for each value of $k = 1, \ldots, K$, we place five sensor nodes evenly distributed on the circle of radius $r$ with center $(kR, 0)$. There is a feasible solution with $K$ relay nodes, by placing the $k$th relay node at $(kR, 0)$. Using Algorithm 1, we would place $4K + 2(K - 1) = 6K - 2$ relay nodes. This shows that 6 is a lower bound of the approximation ratio of our algorithm.

In regard to the running time of Algorithm 1, we note that step_2 takes $O(n \log n)$ time to compute the MST $T_{mst}$ using the method of Shamos and Hoey [17], while step_3 takes $O(|\mathrm{Steiner}(\mathcal{T}^A)|)$ time to steinerize $T_{mst}$. Therefore, the worst-case time complexity of Algorithm 1 is $O(n \log n + |\mathrm{MCST}(\mathcal{X}, r, R)|)$.

## 4 Two-Tiered Relay Node Placement

In this section, we present a general framework that combines any approximation algorithm for DCover and any approximation algorithm for SMT-MSPBEL to obtain an approximation for 2tRNP. The general framework is presented as Algorithm 2.

**Algorithm 2** A General Framework for 2tRNP

INPUT: A set of $n$ sensor nodes $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ in the Euclidean plane and positive real constants $r > 0$ and $R \geq r$ representing the communication ranges of sensor nodes and relay nodes, respectively, an approximation algorithm $\mathcal{A}$ for DCover, and an approximation algorithm $\mathcal{B}$ for SMT-MSPBEL.

OUTPUT: A set $\mathcal{R} = \{r_1, \ldots, r_k\}$ of relay nodes.

step_1 Apply algorithm $\mathcal{A}$ to $\mathcal{X}$ to obtain a set of points $\mathcal{C} = \{c_1, c_2, \ldots, c_m\}$ that is a feasible solution to $\mathrm{DCover}(\mathcal{X}, r)$. Without loss of generality, we assume that $\mathcal{C}$ is *minimal*, meaning that no proper subset of $\mathcal{C}$ is a feasible solution for $\mathrm{DCover}(\mathcal{X}, r)$.

step_2 Construct a set $\mathcal{D} \subseteq \mathcal{X}$ such that for each $c_i \in \mathcal{C}$, there is exactly one $d_j \in \mathcal{D}$ such that $\|d_j\, c_i\| \leq r$. and, for each point $d_j \in \mathcal{D}$, there is exactly one point $c_i \in \mathcal{C}$ such that $\|d_j\, c_i\| \leq r$.

step_3 Apply $\mathcal{B}$ to obtain a set of relay nodes $\mathcal{Y} = \{y_1, y_2, \ldots, y_l\}$ that is an F1tRNP for $(\mathcal{D}, R, R)$.

step_4 Output $\mathcal{R} = \mathcal{C} \cup \mathcal{D} \cup \mathcal{Y}$.

**Theorem 4.1.** *The set of relay nodes $\mathcal{R}$ produced by Algorithm 2 is an* F2tRNP *for $(\mathcal{X}, r, R)$. In addition, $|\mathcal{R}| \leq (2\alpha + \beta) \times |\mathrm{M2tRNP}(\mathcal{X}, r, R)|$, where $\mathrm{M2tRNP}(\mathcal{X}, r, R)$ is any minimum two-tiered relay node placement for $(\mathcal{X}, r, R)$, $\alpha$ is the approximation ratio of $\mathcal{A}$ for* DCover*, and $\beta$ is the approximation ratio of $\mathcal{B}$ for* SMT-MSPBEL.

**Proof.** We begin by noting that each step in Algorithm 2 is well-defined, other than possibly step_2. In that step, to see that the set $\mathcal{D}$ must exist, consider any $c_i \in \mathcal{C}$ and suppose to the contrary of the second condition that, for every point $d_j \in \mathcal{X}$ with $\|d_j\, c_i\| \leq r$, that there exists a second point in $\mathcal{C}$ within distance $r$ of $d_j$. In that case, $c_i$ could be removed from $\mathcal{C}$ and the resulting set would also be a feasible solution to $\mathrm{DCover}(\mathcal{X}, r)$. This would contradict the assumption that $\mathcal{C}$ is minimal.

To see that $\mathcal{R}$ is an F2tRNP for $(\mathcal{X}, r, R)$, note first that, from step_1 of Algorithm 2, each sensor node is within distance $r$ of some relay node in $\mathcal{C}$ and that, from step_2, each relay node in $\mathcal{C}$ is within distance $R \geq r$ of some relay node in $\mathcal{D}$. Further, from step_3, for any two relay nodes in $\mathcal{D}$, there is a path consisting of relay nodes in $\mathcal{Y}$ such that each hop of the path is not longer than $R$. It follows that the set of relay nodes $\mathcal{R}$ produced by Algorithm 2 is an F2tRNP for $(\mathcal{X}, r, R)$.

To establish the approximation ratio, we begin by letting M2tRNP be a minimum two-tiered relay node placement for $(\mathcal{X}, r, R)$. We will prove a two part *lower bound* on

$|\mathrm{M2tRNP}(\mathcal{X}, r, R)|$. First, note that each sensor node in $\mathcal{X}$ must be within distance $r$ of some relay node in $\mathrm{M2tRNP}(\mathcal{X}, r, R)$. Therefore, $|\mathrm{M2tRNP}(\mathcal{X}, r, R)| \geq |\mathbf{C}|$, where $\mathbf{C}$ is any optimal solution to $\mathrm{DCover}(\mathcal{X}, r)$. Since the approximation ratio of algorithm $\mathcal{A}$ for DCover is $\alpha$, we have that $|\mathcal{C}| \leq \alpha|\mathbf{C}|$; hence,

$$|\mathrm{M2tRNP}(\mathcal{X}, r, R)| \geq \frac{|\mathcal{C}|}{\alpha}. \tag{4.1}$$

Second, since $\mathrm{M2tRNP}(\mathcal{X}, r, R)$ is a feasible solution for $2\mathrm{tRNP}(\mathcal{X}, r, R)$ and since $\mathcal{D} \subseteq \mathcal{X}$, it follows that $\mathrm{M2tRNP}(\mathcal{X}, r, R)$ is a feasible solution to $2\mathrm{tRNP}(\mathcal{D}, r, R)$ and, therefore, a feasible solution to $2\mathrm{tRNP}(\mathcal{D}, R, R)$, which is a feasible solution to $1\mathrm{tRNP}(\mathcal{D}, R, R)$. This implies that

$$|\mathrm{M2tRNP}(\mathcal{X}, r, R)| \geq |W|, \tag{4.2}$$

where $W$ is an M1tRNP to $1\mathrm{tRNP}(\mathcal{D}, R, R)$. Combining (4.1) and (4.2), we obtain the lower bound

$$|\mathrm{M2tRNP}(\mathcal{X}, r, R)| \geq \max\left\{\frac{|\mathcal{C}|}{\alpha}, |W|\right\}. \tag{4.3}$$

Since $|\mathcal{D}| = |\mathcal{C}|$ and since $|\mathcal{Y}| \leq \beta \times |W|$, where $\beta$ is the approximation ratio of algorithm $\mathcal{B}$ for SMT-MSPBEL, we have

$$\frac{|\mathcal{R}|}{|\mathrm{M2tRNP}(\mathcal{X}, r, R)|} = \frac{|\mathcal{C}| + |\mathcal{D}| + |\mathcal{Y}|}{|\mathrm{M2tRNP}(\mathcal{X}, r, R)|} \leq \frac{2 \times |\mathcal{C}| + \beta \times W}{\max\left\{\frac{|\mathcal{C}|}{\alpha}, W\right\}}. \tag{4.4}$$

When $\frac{|\mathcal{C}|}{\alpha} \geq W$, we have

$$\frac{2 \times |\mathcal{C}| + \beta \times W}{\max\left\{\frac{|\mathcal{C}|}{\alpha}, W\right\}} \leq \frac{2 \times |\mathcal{C}| + \beta \times \frac{|\mathcal{C}|}{\alpha}}{\frac{|\mathcal{C}|}{\alpha}} \leq 2 \times \alpha + \beta. \tag{4.5}$$

When $\frac{|\mathcal{C}|}{\alpha} \leq W$, we have

$$\frac{2 \times |\mathcal{C}| + \beta \times W}{\max\left\{\frac{|\mathcal{C}|}{\alpha}, W\right\}} \leq \frac{2 \times \alpha \times W + \beta \times W}{W} \leq 2 \times \alpha + \beta. \tag{4.6}$$

Therefore, in both cases, we have

$$\frac{|\mathcal{R}|}{|\mathrm{M2tRNP}(\mathcal{X}, r, R)|} \leq 2 \times \alpha + \beta. \tag{4.7}$$

This completes the proof of the theorem. □

The best approximation ratio emerging from this framework combines the best approximation algorithms for the DCover and SMT-MSPBEL problems. Specifically, for the DCover problem, there is a polynomial time approximation scheme $\mathcal{A}_\epsilon$, which, for any given positive constant $\epsilon > 0$, produces a $(1 + \epsilon)$-approximation of the optimal solution [12]. For the SMT-MSPBEL problem, there is a polynomial time 3-approximation algorithm $\mathcal{B}$ [2], [3], as well as a randomized 2.5-approximation algorithm [3]. Therefore, we have the following theorem:

**Theorem 4.2.** *Let $\epsilon > 0$ be any given positive constant. There is a polynomial time approximation algorithm for 2tRNP with an approximation ratio of $5 + \epsilon$. There is also a polynomial time randomized approximation algorithm for 2tRNP with an approximation ratio of $4.5 + \epsilon$.*

## 5 CONCLUSIONS

In this paper, we have studied the single-tiered relay node placement problem and the two-tiered relay node placement problem in a wireless sensor network. For the first problem, we have presented a polynomial time approximation algorithm whose approximation ratio is between 6 and 7. For the second problem, we have presented a general framework combining an approximation algorithm for the minimum geometric disk cover problem and an approximation algorithm for the Steiner minimum tree with minimum number of Steiner points and bounded edge length problem. Using the best-known algorithm for each of the problems, the framework gives a $(5 + \epsilon)$-approximation algorithm, where $\epsilon$ can be any positive constant, and a randomized $(4.5 + \epsilon)$-approximation algorithm. Future research directions include tighter analysis of the algorithms presented here and design of better algorithms for these problems.

## ACKNOWLEDGMENTS

The authors thank the associate editor and the anonymous reviewers whose valuable comments on an earlier version of this paper helped to greatly improve the presentation of the paper. The work described here was initiated while Errol L. Lloyd was visiting Arizona State University and prepared through collaborative participation in the Communications and Networks Consortium sponsored by the US Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-001. The research of Guoliang Xue was supported in part by US Army Research Office grant W911NF-04-1-0385 and US National Science Foundation grant CCF-0431167. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory, the Army Research Office, or the US Government.

## REFERENCES

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer Networks J.*, vol. 38, pp. 393-422, 2002.

[2] D. Chen, D.Z. Du, X.D. Hu, G. Lin, L. Wang, and G. Xue, "Approximations for Steiner Trees with Minimum Number of Steiner Points," *J. Global Optimization*, vol. 18, pp. 17-33, 2000.

[3] X. Cheng, D.Z. Du, L. Wang, and B. Xu, "Relay Sensor Placement in Wireless Sensor Networks," *ACM/Springer J. Wireless Networks*, to appear, http://www.seas.gwu.edu/~cheng/Publication/relay.pdf.

[4] X. Cheng, B. Narahari, R. Simha, M.X. Cheng, and D. Liu, "Strong Minimum Energy Topology in Wireless Sensor Networks: NP-Completeness and Heuristics," *IEEE Trans. Mobile Computing*, vol. 2, pp. 248-256, 2003.

[5] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, second ed. MIT Press and McGraw-Hill, 2001.

[6] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next Century Challenges: Scalable Coordination in Sensor Networks," *Proc. ACM MobiCom '99*, pp. 263-270, 1999.

[7] R.J. Fowler, M.S. Paterson, and S.L. Tanimoto, "Optimal Packing and Covering in the Plane Are NP-Complete," *Information Processing Letters*, vol. 12, pp. 133-137, 1981.

[8] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W.H. Freeman and Co., 1979.

[9] B. Hao, J. Tang, and G. Xue, "Fault-Tolerant Relay Node Placement in Wireless Sensor Networks: Formulation and Approximation," *Proc. IEEE Workshop High Performance Switching and Routing (HPSR '04)*, pp. 246-250, 2004.

[10] Y.T. Hou, Y. Shi, and H.D. Sherali, "Rate Allocation in Wireless Sensor Networks with Network Lifetime Requirement," *Proc. ACM MobiHoc '04*, pp. 67-77, 2004.

[11] B. Karp and H. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Networks," *Proc. ACM MobiCom '00*, pp. 243-254, 2000.

[12] D.S. Hochbaum and W. Maass, "Approximation Schemes for Covering and Packing Problems in Image Processing and VLSI," *J. ACM*, vol. 32, pp. 130-136, 1985.

[13] Q. Li, J. Aslam, and D. Rus, "Online Power-Aware Routing in Wireless Ad-Hoc Networks," *Proc. ACM MobiCom '01*, pp. 97-107, 2001.

[14] G. Lin and G. Xue, "Steiner Tree Problem with Minimum Number of Steiner Points and Bounded Edge-Length," *Information Processing Letters*, vol. 69, pp. 53-57, 1999.

[15] E. Lloyd, R. Liu, M. Marathe, R. Ramanathan, and S.S. Ravi, "Algorithmic Aspects of Topology Control Problems for Ad-Hoc Networks," *Proc. ACM MobiHoc '02*, pp. 123-134, 2002.

[16] J. Pan, Y.T. Hou, L. Cai, Y. Shi, and S.X. Shen, "Topology Control for Wireless Sensor Networks," *Proc. ACM MobiCom '03*, pp. 286-299, 2003.

[17] M.I. Shamos and D. Hoey, "Closest-Point Problems," *Proc. Ann. Symp. IEEE Foundation of Computer Science (FOCS '75)*, pp. 151-162, 1975.

[18] A. Srinivas and E. Modiano, "Minimum Energy Disjoint Path Routing in Wireless Ad-Hoc Networks," *Proc. ACM MobiCom '03*, pp. 122-133, 2003.

[19] J. Tang, B. Hao, and A. Sen, "Relay Node Placement in Large Scale Wireless Sensor Networks," *Computer Comm.*, vol. 29, pp. 490-501, 2006.