Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Computer Communications 29 (2006) 490–501
computer communications
www.elsevier.com/locate/comcom

Relay node placement in large scale wireless sensor networks

Jian Tang*, Bin Hao, Arunabha Sen

Department of Computer Science and Engineering, Arizona State University, Tempe, AZ 85287, USA

Available online 19 February 2005

Abstract

Scalability and extended lifetime are two critical design goals of any large scale wireless sensor network. A two-tiered network model has been proposed recently for this purpose. This is a cluster-based network model composed of *relay* and *sensor* nodes. Relay nodes are placed in the *playing field* to serve as cluster heads and to form a connected network topology for information dissemination at the higher tier. The relay nodes are capable of aggregating data packets from the sensor nodes in their clusters and transmitting them to the sink node via wireless multi-hop paths. In this paper, we study the relay node placement problem in large scale wireless sensor networks. Our objective is to place the fewest number of relay nodes in the playing field of a sensor network such that (1) each sensor node can communicate with at least one relay node and (2) the network of relay nodes is connected.

However, placement strategies realizing goals (1) and (2) do not provide any fault-tolerance as the network may lose functionality after failure of some of the relay nodes. In order to incorporate fault-tolerance in such a network, we ensure that every sensor node is able to communicate with at least two relay nodes and the induced network topology is 2-connected. This strategy will ensure survivability of the network in the event of single fault, in lieu of higher relay node placement cost. We formulate the relay node placement in wireless sensor networks as two optimization problems: (i) Connected Relay Node Single Cover (CRNSC) problem and (ii) 2-Connected Relay Node Double Cover (2CRNDC) problem. We present two polynomial time approximation algorithms to solve the CRNSC problem. We prove that the ratio of the number of relay nodes needed by the approximation algorithm to the number of relay nodes needed by the optimal algorithm is bounded by 8 for the first algorithm and 4.5 for the second. In addition, for the 2CRNDC problem we provide two approximation algorithms with performance bounds 6 and 4.5, respectively.

Keywords: Relay node placement; Wireless sensor network; Fault-tolerance

1. Introduction

A wireless sensor network is composed of hundreds or even thousands of sensor nodes which use wireless links to perform distributed sensing tasks. Each sensor node includes a sensing module, a computing module, memory and a wireless communication module with a very limited communication range. Wireless sensor network has received intensive research attentions due to its enormous application potential in battlefield surveillance, environmental monitoring, biomedical observation and other fields [1]. The three basic requirements for designing efficient wireless sensor networks are scalability, fault-tolerance and energy efficiency. A sensor network, comprising of a number of sensor nodes, is usually required to cover a large geographic area. New sensor nodes may be added to the network and existing sensor nodes may become inoperative at any time. This large scale and frequently changing network requires scalable protocols and algorithms. Factors, such as energy depletion, harsh environmental conditions, and/or malicious attacks may result in node failures in a wireless sensor network. Therefore, survivability of sensor networks is a critical design goal. Moreover, energy is one of the most precious resource in wireless sensor networks. Sensor nodes are normally powered by batteries and can only last for a fairly short period of time if operated at high transmission power levels. As a consequence, energy efficient design is needed for prolonging network lifetime.

* Corresponding author. Tel.: +1 480 9664572; fax: +1 4809652751.
E-mail addresses: jian.tang@asu.edu (J. Tang), binhao@asu.edu (B. Hao), asen@asu.edu (A. Sen).

0140-3664/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2004.12.032

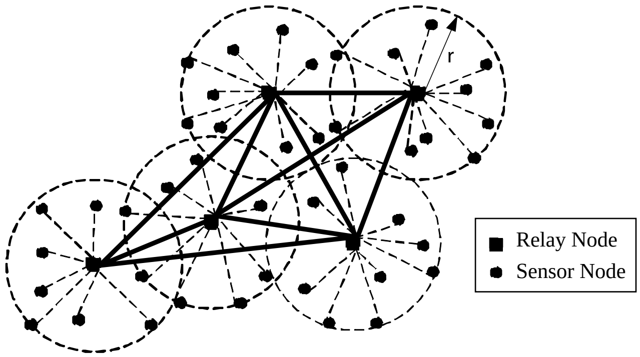

Fig. 1. A two-tiered wireless sensor network.

Information dissemination, i.e. sending out queries from sinks to sensor nodes and gathering results from sensor nodes back to sinks, is a basic function of wireless sensor networks. Direct transmission from data sources to sinks is usually not practical because sinks are generally far away from data sources and sensor nodes have very limited communication ranges. Hence, wireless multi-hop paths need to be established to route data. However, sensor nodes participating in data routing may consume a large volume of energy due to transmission. In order to relieve this burden from sensor nodes and to provide a scalable framework for sensor networks, a two-tiered network model is proposed [8,9,16]. Fig. 1 shows a two-tiered wireless sensor network, where the radii of dotted circles (centered at relay nodes) indicate the communication ranges of sensor nodes. From the figure, we can see that sensor nodes are grouped in clusters and relay nodes serve as cluster heads. It is well known that cluster-based architectures are scalable. In the two-tied network, the low tier consists of a large number of sensor nodes whose task is to sense the vicinity, to generate corresponding data packets and to transmit them directly to their cluster heads. Relay nodes are on the high tier. They are also called Gateway Nodes in [8] or *Application Nodes* (*AN*) in [16]. A relay node is more powerful than a sensor node in terms of energy storage, computing and communication capability. It can extract useful information and remove redundancy in data packets gathered from sensor nodes in its cluster. It can then generate outgoing packets with much smaller total size and send them to sinks via multi-hop paths composed of other relay nodes [16]. This in-network data fusion can dramatically decrease the traffic load thereby extend the network lifetime.

In this paper, we assume that sensor nodes are uniformly deployed in a playing field. In order to make the two-tiered sensor network work properly, a basic requirement is to guarantee that each sensor node can reach at least one relay node and that the relay nodes form a connected network topology. A sensor node is *covered* by a relay node if the relay node is within the communication range of the sensor node. We first study the problem of placing the minimum number of relay nodes in the playing field such that all sensor nodes are covered and relay nodes form a connected network topology. Relay nodes, just like sensor nodes, may fail due to various reasons. Failures of relay nodes may disconnect the network and prevent the collected data from reaching the sink node. In order to provide some measure of fault-tolerance in the network, it is preferable that each sensor node is covered by at least two relay nodes and there exist two node-disjoint paths between each pair of relay nodes in the network. This will ensure that in the event of a single relay node failure, the sensor data can still be delivered to the sink node. Moreover, this will also facilitate load balancing among the relay nodes, thereby extend network lifetime. The idea of covering each sensor node by multiple relay nodes was first considered by [9] and it provides an opportunity for load-balanced clustering. Our scheme provides fault-tolerance for the sensor network in lieu of higher cost for relay node placement.

We formulate the relay node placement in wireless sensor networks as two optimization problems: (i) *Connected Relay Node Single Cover* (*CRNSC*) problem and (ii) *2-Connected Relay Node Double Cover* (*2CRNDC*) problem. We present two polynomial time approximation algorithms to solve the CRNSC problem. We prove that the ratio of the number of relay nodes needed by the approximation algorithm to the number of relay nodes needed by the optimal algorithm is bounded by 8 for the first algorithm and 4.5 for the second. In addition, for the 2CRNDC problem we provide two approximation algorithms with performance bounds 6 and 4.5, respectively. To the best of our knowledge, this is the first paper to provide constant bound approximation algorithms for the relay node placement problems in the two-tiered large scale wireless sensor networks. Once the positions of the relay nodes are computed using the algorithms presented in this paper, the clustering schemes described in [8,9] and the data gathering protocols like *LEACH* [10] can be used.

The rest of the paper is organized as follows. We discuss related work in Section 2. We formally define problems and notations in Section 3. We describe our approximation algorithms, their correctness proofs and approximation bounds in Sections 4 and 5. We present and analyze our simulation results in Section 6. We conclude the paper in Section 7.

2. Related work

In recent times, several protocols and algorithms have been developed for two-tiered wireless sensor networks. The authors of [16] consider a two-tiered wireless sensor network consisting of sensor clusters deployed around strategic locations and Base Station (BS) whose locations are relatively flexible. They propose approaches to maximize the network lifetime by arranging the location of BSs and inter-Application Nodes (same as relay node in this paper). It may be noticed that their work does not address the issue of covering sensor nodes by relay nodes, which is one of the major goals of this research work. A fault-tolerant clustering scheme is proposed in [8]. The objective of this scheme is to detect failure and to recover sensor nodes from the failed gateway node. In addition, authors of [9] propose a clustering scheme whose objective is to balance the traffic load among all gateway nodes. The optimization problems considered in [8] are substantially different from the ones considered here as they do not address the placement problem of gateway/relay nodes.

The research that is closest to the results presented in this paper is due to [3] by Cheng et al. They assume that the sensor nodes are not connected initially and they study the placement problem of relay nodes to make the induced network topology globally connected. They formulate it as an optimization problem called *Steiner minimum tree with minimum number of Steiner points and bounded edge length*. This problem was originally proposed by Lin and Xue in [14]. They give two constant bound polynomial time approximation algorithms to solve this problem. Those approximation algorithms have performance ratios of 3 and 4, respectively. Their work is different from ours in two respects. First, their research is not based on the two-tiered network model. Second, the network survivability is not addressed by their work.

There may exist some redundancy, in case sensor nodes are densely packed in a geographic region (*playing field*). In some of the schemes, the redundant sensor nodes are put to sleep mode to save energy. Some recent work [6] analyzes sensing redundancy among neighboring sensor nodes. A coverage-based off-duty eligibility rule and node scheduling scheme is proposed in [20]. This scheme guarantees that the original sensing coverage is maintained after turning off redundant nodes. The authors of [7] propose an approximation algorithm to compute a connected sensor cover which includes minimum number of sensor nodes forming a connected topology. Information dissemination is also a fundamental issue in wireless sensor networks, and has been extensively studied in the literature. The authors of [10] present an energy-efficient cluster-based protocol, LEACH, for gathering data packets from all sensor nodes and delivering them to the Base Station (BS). In LEACH, only a fraction of nodes become *head nodes* in every round. Data reports from non-head sensor nodes are aggregated at the head nodes and sent directly to BS. In [13], authors present an improved scheme, called PEGASIS (Power-Efficient Gathering in Sensor Information Systems), which constructs a chain for data gathering. Nodes on the chain take turn to transmit aggregated packets to the BS. Directed diffusion, a flooding-based scheme, is presented in [12] for routing queries from sinks to all sensor nodes and for gathering result packets along the opposite direction. The sensor network considered in [12] is single-tiered and comprises of sensor nodes only (no relay nodes). The in-network data aggregation and packet relaying are carried out by the sensor nodes.

Since nodes in wireless sensor networks are vulnerable to failure, fault-tolerant design of sensor networks is important. The traditional way for tolerating node/link failures is to establish disjoint paths from the source node to the sink. Suurballe in [19] proposes an optimal algorithm to compute link disjoint paths with the minimum total cost in the network. The broadcasting feature of radio makes the fault-tolerant routing in the wireless network different from that in the traditional networks. Srinivas and Modiano in [18] propose elegant optimal algorithms for finding both node-disjoint and link-disjoint paths with minimum total energy in wireless networks.

The well known *Minimum Geometric Disk Cover*f (*MGDC*) problem [5] is similar to problems under study in this paper. The MGDC problem was proved to be NP-hard in [5]. A polynomial time approximation scheme capable of solving different kinds of geometric covering problem is given in [11]. It may be noted that, optimization problems studied in this paper are substantially different from the ones in [5]. This paper considers the issues of (1) connectivity and (2) double-cover, which are absent in [5].

3. Problem statements

In this section, we formally define the two problems addressed in this paper, i.e. *Connected Relay Node Single Cover* (*CRNSC*) problem, and *2-Connected Relay Node Double Cover* (*2CRNDC*) problem.

Consider a sensor network formed by a set of sensor nodes which are uniformly distributed in a rectangular region. We assume that all sensor nodes have the same communication range r. The network model is two-tiered, comprising of sensor and relay nodes. Each relay node has a communication range, R. Normally, R is much larger than r. In this paper, we assume that $R \geq 4r$. This assumption is consistent with the current communication ranges of sensor nodes and 802.11-based wireless ad hoc nodes. A relay node can communicate with any sensor node within a distance of r, the communication range of the sensor nodes.

To accomplish the task of data gathering, an intuitive objective may be to place the fewest number of relay nodes so that each sensor node is covered by at least one relay node. Since gathered data has to be transmitted to the sink node, which may be far away from the data sources, we also require that all relay nodes be able to communicate with each other through a multi-hop path.

We call a set of relay nodes *connected*, if any pair of them can communicate with each other through a multi-hop path of relay nodes. Now, we are ready to define the *Connected Relay Node Single Cover* problem.

Definition 1 [Connected Relay Node Single Cover (CRNSC) Problem]. Given a set of locations of uniformly distributed sensor nodes S, the communication range of sensor nodes r, and the communication range of relay nodes R, find the minimum number of relay nodes and their corresponding locations, so that each sensor node is covered by at least one relay node, and that the set of relay nodes is connected.

If the network of relay nodes are not required to be connected, then the problem becomes the Relay Node Single Cover (RNSC) problem.

We also study a fault-tolerant version of the above problem. In particular, we study a relay node placement problem which is resilient to single (relay) node failure. For this purpose, we require that each sensor node to be covered by at least two relay nodes and that the network induced by relay nodes to be *2-connected* [2]. The network of relay nodes is 2-connected, if there exists at least two node disjoint paths between every pair of relay nodes.

Definition 2 [2-Connected Relay Node Double Cover (2CRNDC) Problem]. Given a set of locations of uniformly distributed sensor nodes S, the communication range of sensor nodes r, and the communication range of relay nodes R, find the minimum number of relay nodes and their corresponding locations so that each sensor node is covered by at least two relay nodes, and that the network of relay nodes is 2-connected.

Remark. We make the assumption that no two relay nodes are placed at the same location. This is due to the fact that failure of a relay node is often caused by some event in a specific location. The same event may jeopardize two relays if they are placed at the same location.

If we ignore connectivity of the network of relay nodes, then 2CRNDC problem becomes the *Relay Node Double Cover* (*RNDC*) problem.

3.1. Complexity of CRNSC and 2CRNDC problems

If we consider a special case of the *Connected Relay Node Single Cover* (*CRNSC*) problem, where the relay node communication range R is large enough, then the set of relay nodes located at any points within a bounded region will automatically be connected. In this case, the CRNSC problem reduces to the RNSC Problem. In RNSC problem, relay nodes must be placed in locations such that each sensor node is able to reach at least one relay node. In other words, if the communication range of the sensor node is r, relay nodes must be placed in those locations such that if disks of radius r are placed in those locations, they should cover all the sensor nodes. The problem where one tries to find the minimum number of disks of radius r, needed to cover all the sensor nodes (or the points where they are located), is known as the *Minimum Geometric Disk Cover* problem [5]. It has been shown in [5], that the *Minimum Geometric Disk Cover* problem is NP-complete. Since the CRNSC problem is a special case of the RNSC problem, we conclude that it is also NP-Complete. Unfortunately, complexity result of the 2CRNDC problem is unknown at this time. Since it appears to be a harder problem in comparison with the CRNSC problem, we conjecture that the 2CRNDC problem is also NP-Complete. In Sections 4 and 5, we present polynomial time approximation algorithms with provable performance bounds for the CRNSC and the 2CRNDC problem, respectively.

4. Approximation algorithms for the CRNSC problem

In this section, we present two polynomial time approximation algorithms for the CRNSC problem. We first present an algorithm which is easy to understand. The performance guarantee of this algorithm is 8. We then refine the algorithm to get another polynomial time algorithm with a performance guarantee of 4.5.

Before describing our algorithms, we give some definitions and notations. We call a position p a *Possible-position* denoted by *P-position* for relay nodes, if there exist two sensor nodes s and s' such that $distance(s,p) = distance(s',p) = r$, where r is the communication range of sensor nodes. If a sensor node s is separated from other sensor nodes by a distance of more than $2r$, we take any two distinct points on the circle of center s and radius r as P-positions. One can easily see that corresponding to any pair of sensor nodes with distance less than $2r$, there are two P-positions for the relay nodes. Similarly, corresponding to any pair of sensor nodes with distance exactly $2r$, there is exactly one P-position for the relay node. For any P-position, p, of relay nodes, denote by $C(p)$ the set of sensor nodes which can be covered by a relay node located at position P. For a set H of P-positions, denote by $C(H)$ the set of sensor nodes which can be covered by some relay node located at a position in H. Ignoring connectivity, we claim that deploying relay nodes at all P-positions is sufficient to cover the set of sensor nodes. This can be verified as follows. Let H be any set of relay nodes, which covers the set of sensor nodes. Suppose that some of the relay nodes in H are not located at P-positions. For each relay node q in H which is not located at a P-position, we can find a P-position from where it can cover at least those sensor nodes, which can be covered from its original position. This is illustrated in Fig. 2. In the figure on the left side, the relay node is not located at a P-position. In the figure on the right side, the relay node has been moved to a P-position, from where it is able to cover all those sensor nodes it was covering before.

The main ideas of our algorithms are as follows: (i) divide the region, within which the sensor nodes are distributed, into small square boxes called *cells*; (ii) without considering the connectivity, find the optimal solution to cover (or double cover) the sensor nodes within each cell; (iii) make the network of relay nodes connected (or 2-connected), by adding in extra relay nodes, if necessary. One might think that the running time of the algorithms will be exponential, because in order to find the optimal solution

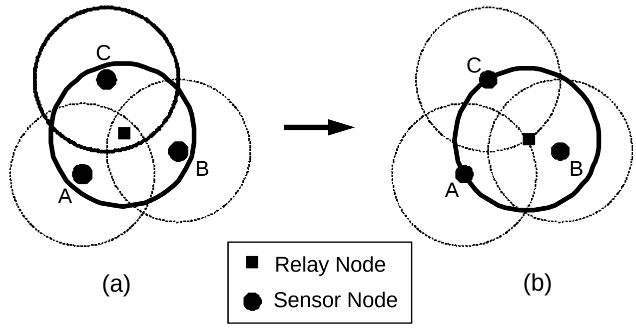

Fig. 2. (a) A relay node not located at a P-position covering sensor nodes A, B and C; (b) a relay node located at a P-position covering A, B and C.

for each cell, one has to exhaustively search all the subsets of P-positions. Fortunately, that is not the case, when the size of a cell is small. For example, suppose the size of a cell is $2r \times 2r$, where r is the communication range of sensor nodes. Then, one cell can always be covered by four disks of radius r, which implies that we only need to search the subsets of size less than or equal to 4. As a consequence, the algorithms run in polynomial time.

The idea of dividing the region into cells, and finding optimal solution for each cell was first introduced by Hochbaum and Maass [11]. Let B be a rectangular region in which a given set of points is placed. In order to cover those points by disks of radius r, divide B into strips of width $l \cdot 2r$, where l is an integer. Let A be an approximation algorithm for any strip, whose performance ratio is r_A. Here, the *Performance Ratio* of A is defined as the ratio of the size of the solution provided by A divided by the size of optimal solution. Let S_A be the algorithm for the whole region B, which combines the solutions of A on every strip.

Lemma 1. *Shifting Lemma* [11]

$$r_{S_A} \leq r_A \left(1 + \frac{1}{l}\right)$$

where r_{S_A} *is the performance ratio of algorithm* S_A.

Clearly, inside each strip, we can divide it into square boxes (cells) of side length $l \cdot 2r$, and apply Shifting Lemma again. Let A' be an approximation algorithm for one cell, with performance ratio $r_{A'}$. Let $S_{A'}$ be the algorithm for the whole region B, which combines the solutions of A' on every cell. By applying the Shifting Lemma twice, we can have the following corollary [11].

Corollary 1.

$$r_{S_{A'}} \leq r_{A'} \left(1 + \frac{1}{l}\right)^2$$

where $r_{S_{A'}}$ *is the performance ratio of algorithm* $S_{A'}$.

For the rest of the paper, we call l the *partition factor* of an algorithm.

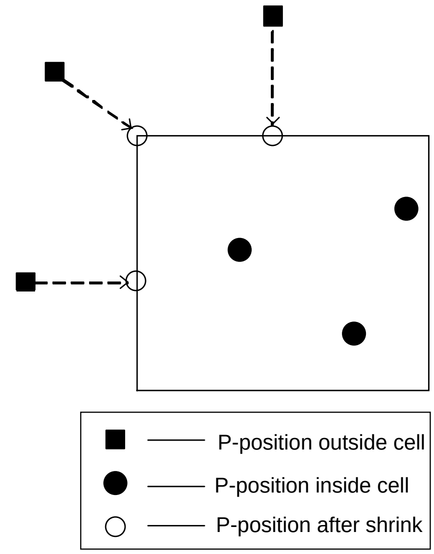

Fig. 3. Shrink operation.

As one may notice, given a cell B_i, some of the P-positions for sensor nodes inside B_i may not be in B_i. In [11], the authors do not consider any connectivity issue, so there is no need to require all P-positions to be inside the cell. But, in our case in order to make relay nodes connected (or 2-connected), we want the P-positions to be inside the cell. Hence, we define a *Shrink Operation* as follows (See Fig. 3): given a cell B_i and the set P of P-positions of relay nodes for the sensor nodes inside B_i. For all $p \in P$, if p is outside B_i, replace p by a point q on the border of B_i, such that q is the closest point from B_i to p. It is not hard to see that by applying the shrink operation, the new set of P-positions can cover at least the same set of sensor nodes inside B_i.

Now, we are ready to present the algorithms. We first give an approximation algorithm for CRNSC. We use K to represent the number of times a sensor node needs to be covered, e.g. $K=1$ means single covering, and $K=2$ means double covering.

Algorithm 1.1. (CRNSC, $l=1$)

Step 0. Divide the region into cells with side length $D=2r$, i.e. the partition factor $l=1$. For each cell, find all P-positions for relay nodes. Apply a shrink operation.

Step 1. Inside each cell, exhaustively search all 1,2,3,4-subsets of the P-positions inside (or on) the cell, to find a subset with smallest order which can cover all the sensor nodes in the cell.

Step 2. For each cell B_i, let H_i be the set of relay nodes found in Step 1 for B_i.

For all H_i, suppose H_{i+1} is the set to the right of H_i, and H_{i+x} is the set directly under H_i.

If either H_i and H_{i+1} are not connected or H_i and H_{i+x} are not connected, we add a relay node at the right bottom corner of B_i. We add a relay node at the right top corner of B_i, if B_i is in the bottom row (see Fig. 4).

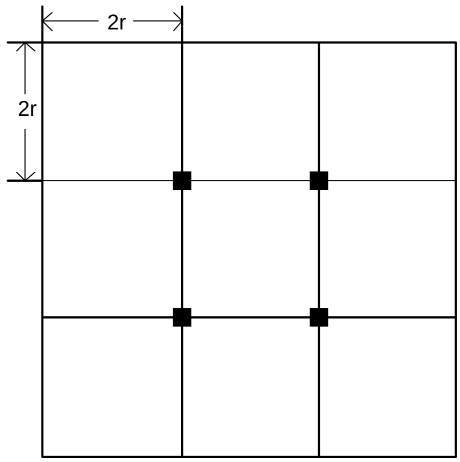

Fig. 4. Possible locations to add extra relay nodes for Step 2 of Algorithm 1.1.

Repeat Step 2, until all H_i's are connected.

Theorem 1. Algorithm 1.1 *can always give a solution for the CRNSC problem. And its performance ratio is bounded by 8.*

Proof. In order to prove the correctness of Algorithm 1.1, we need to prove the following:

I. All sensor nodes are covered by at least one relay node.
II. All relay nodes are connected to one another.

I. Every sensor node will be covered by at least one relay node, if Step 1 of the algorithm ends properly.

If any 1, 2, 3, or 4-subset of the set of P-positions can cover all the sensor nodes in a cell, then the subset will be found by Step 1.

Suppose no subset with size no more than 4 is found by Step 1 for some cell. We are going to derive a contradiction. We know that each cell of side length $2r$ can be covered by four disks of radius r, which means these four disks can cover all the sensor nodes inside the cell. If all the four disks are centered at P-positions, then the set of their centers is a 4-subset of the set of P-positions, and should be found by Step 1, a contradiction. So, we assume some of the centers are not at P-positions. For each disk C whose center is not at a P-position, we can move the disk towards the sensor nodes it covers, till there are two sensor nodes on the edge of the disk. Now, the distances between the center of the disk and the two sensor nodes are both r, i.e. the center of the disk is at a P-position. So, by moving the disks around, we get a 4-subset of the set of P-positions which covers all sensor nodes in the cell, a contradiction. Thus, we proved that after Step 1, every sensor node is covered by at least one relay node.

II. Clearly, since the side length of each cell is $2r$, and $R \geq 4r$, any two relay nodes in a cell are connected to each other. And by Step 2, relay nodes in different cells are connected to one another.

Now, we show that the performance ratio of Algorithm 1.1 is bounded by 8. Let H' be the set of relay nodes found in Step 1, i.e. $H' = \cup H_i$. Let OPT be the optimal solution of the corresponding RNSC problem. By Corollary 1, we have

$$\frac{|H'|}{|\mathrm{OPT}|} \leq (1 + 1/l)^2 = 4$$

Let H^* be the solution provided by Algorithm 1.1. By our assumption that each cell contains at least one sensor node, H' has at least one relay node for each cell.

By Step 2, at most one extra relay node is added for each cell. Hence,

$$\frac{|H^*|}{|H'|} \leq 2$$

Therefore,

$$\frac{|H^*|}{|\mathrm{OPT}|} \leq 8$$

□

As one may notice that the partition factor l plays an important role in the performance ratio of the algorithm. If everything else is fixed, then the bigger l gets, the smaller the performance ratio will be. So, next, we try $l=2$.

Algorithm 1.2. (CRNSC, $l=2$)

Step 0. Divide the region into cells with side length $D=4r$, i.e. $l=2$. For each cell, find all P-positions for relay nodes. Apply a shrink operation.

Step 1. Inside each cell, exhaustively search all 1 through 9-subsets of the P-positions inside (or on) the cell, to find a subset with smallest order which can cover all the sensor nodes in the cell.

Step 2. Try to connect the relay nodes found in Step 1.

Step 2.1. First connect the relay nodes inside each cell as follows.

$\forall H_i$, if it is not connected, add a relay node, q, at the center of B_i. Set $H_i = H_i \cup \{q\}$.

Step 2.2. Connect the H_i's.

Let H_{i+1} be the set of relay nodes to the right of H_i, and H_{i+x} be the set directly under H_i.

If H_i and H_{i+1} are not connected, try to add a relay node, q_i, at the center of B_i, or a relay node, q_{i+1}, at the center of B_{i+1}, or both, to make the new H_i and H_{i+1} connected.

Similarly, if H_i and H_{i+x} are not connected, try to add a relay node, q_i, at the center of B_i, or a relay node, q_{i+x}, at the center of B_{i+x}, or both, to make the new H_i and H_{i+x} connected.

Theorem 2. Algorithm 1.2 *can always give a solution for CRNSC problem. And its performance ratio is bounded by 4.5.*

Proof. Similar to Algorithm 1.1, in order to show the correctness we need to show the following:

I. All sensor nodes are covered by at least one relay node.
II. All relay nodes are connected to one another.

I. Based on the fact that each cell of side length $4r$ can be covered by 9 disks of radius r [15], apply the same argument for the correctness of Algorithm 1.1, we can have that every sensor node is covered by at least one relay node.

II. By Step 2, all relay nodes are connected to one another.

Now, we show the performance ratio of Algorithm 1.2 is bounded by 9/2. Let H' be the set of relay nodes found in Step 1, i.e. $H' = \cup H_i$. By Corollary 1, we have

$$\frac{|H'|}{|\mathrm{OPT}|} \leq (1 + 1/l)^2 = \frac{9}{4}$$

Let H^* be the solution provided by Algorithm 1.2. By our assumption that each cell contains at least one sensor node, H' has at least one relay node for each cell. By Step 2, at most one extra relay node is added for each cell. Hence,

$$\frac{|H^*|}{|H'|} \leq 2$$

Therefore,

$$\frac{|H^*|}{|\mathrm{OPT}|} \leq \frac{9}{2}$$

□

Be aware of the fact that the optimal solution of CRNSC problem is also a solution for the corresponding RNSC problem (relay nodes are not necessarily connected). Hence, the performance ratio of our approximation solution against the optimal solution of the corresponding RNSC problem is an upper bound of the actual performance ratio.

5. Approximation algorithms for the 2CRNDC problem

Now, we consider the *2-Connected Relay Node Double Cover* problem.

Remarks. As one may notice, both Shifting Lemma and Corollary 1 are for single cover scenario. But it is not hard to see, from the proof of the Shifting Lemma, that it does not depend on the number of times the points being covered. Hence, the Shifting Lemma can still be applied to the double cover scenario, and so can Corollary 1.

Algorithm 2.1. (2CRNDC, $l=1$)

Step 0. Divide the region into cells of side length $D=2r$, i.e. $l=1$. For each cell, find all P-positions for relay nodes. Apply a shrink operation.

Step 1. Inside each cell, exhaustively search all 1 through 8-subsets of the P-positions inside (or on) the cell, to find a subset with smallest order, which can cover all sensor nodes in the cell at least twice. For each cell B_i, let H_i be the set of relay nodes found for B_i (note that each H_i has at least two relay nodes).

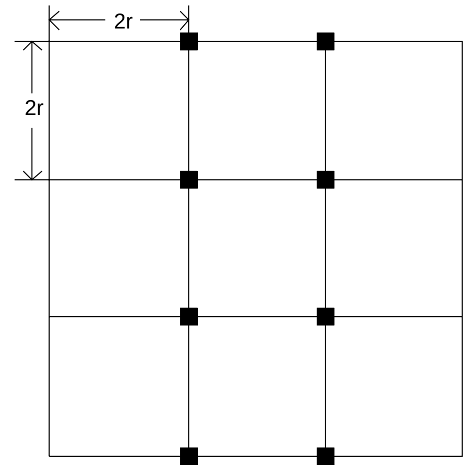

Fig. 5. Possible locations to add extra relay nodes for Step 2 and 3 of Algorithm 2.1.

Step 2. Clearly, since the side length is $2r$, and $R \geq 4r$, any H_i containing more than two relay nodes will be 2-connected. For all H_i's containing exactly two relay nodes, we put a relay node, q, at the right top corner of B_i, and set $H_i = H_i \cup \{q_i\}$. Add q at the left top corner of B_i, if there are no cells to the right of B_i (see Fig. 5).

Step 3. Now, we make the set of the chosen relay nodes to be 2-connected (see Fig. 5).

Step 3.1. Connect each row of the H_i's.

Let H_{i+1} be the set of relay nodes to the right of H_i. If H_i and H_{i+1} are not connected, add a relay node, q_i, at the right top corner of B_i, and set $H_i = H_i \cup \{q_i\}$. Add q_i at the right bottom corner of B_i, if B_i and B_{i+1} are in the bottom row. (Notice that we only make them to be connected here, instead of 2-connected.)

Repeat step 3.1, until every row of H_i's is connected.

Step 3.2. Connect each column of the H_i's.

For each H_i, denote LEFT_{B_i} the set of relay nodes in H_i, which are connected to some relay nodes in the set to the left of H_i, i.e. $\mathrm{LEFT}_{B_i} = \{q \in H_i | q$ is connected to a node in the set to the left of $H_i\}$; denote RIGHT_{B_i} the set of relay nodes in H_i, which are connected to some relay nodes in the set to the right of H_i, i.e. $\mathrm{RIGHT}_{B_i} = \{q \in H_i | q$ is connected to a node in the set to the left of $H_i\}$.

If $|\mathrm{LEFT}_{B_i} \cup \mathrm{RIGHT}_{B_i}| > 1$, then $H'_i = H_i$. Otherwise, $H'_{\mathrm{i}} = \mathrm{H}_i - \mathrm{LEFT}_{B_i} \cup \mathrm{RIGHT}_{B_i}$.

Let H_{i+x} be the set of relay nodes directly under H_i. If H'_i and H'_{i+x} are not connected, then add a relay node, q, at the right bottom corner of B_i, and set $H_i = H_i \cup \{q\}$.

Repeat Step 3.1, until every column of H_i's is connected.

Theorem 3. *Algorithm 2.1 can always give a solution for the 2CRNDC problem. And its performance ratio is bounded by 6.*

Proof. In order to show the correctness, we need to prove the following:

I. All sensor nodes are covered by at least two relay nodes.
II. All relay nodes are 2-connected. Let u, v be two distinct relay nodes.

I. Similar to the previous algorithms, this is based on the fact that a cell of side length $2r$ can be covered twice by 8 disks of radius r. As shown in Fig. 6, all eight disks are centered on the diagonals of the cell. Four of the centers are r units away from the left top, right top, left bottom and right bottom corners of the cell, respectively. And the other four are $(\sqrt{2}/2)r$ units away from the left top, right top, left bottom and right bottom corners of the cell, respectively.

II. Let q, q' be two relay nodes. We need to show that there are two node-disjoint q–q' paths. Consider the following cases.

Case 1. q, q' are in the same cell. By Step 2, they are 2-connected.

Case 2. q, q' are in different cells. Say q is in cell B, and q' is in cell B'. By the choice of the sets LEFT and RIGHT in Step 3.2, and by the way the cells are vertically connected, one can verify that no matter how B and B' are located, there are always two node-disjoint paths from q to q'.

To save space, we only discuss one subcase to illustrate how to connect q and q'. Suppose B is in ith row and jth column, and B' is in i'th row and j'th column; and $i<i'$, $j<j'$, i.e. B' is to the right of B and is blow B. Also suppose B use the same relay node to connect to its right cell and to the cell underneath. Then, by Step 3.2, we know B must have a different relay node connect to its left cell (for otherwise, $|\mathrm{LEFT}_B \cup \mathrm{RIGHT}_B|=1$). So, we can have one path going from q to the right cell of B, going horizontally then going down to q', and the other path going to the left cell of B, then turning down to the i'-row, and then going horizontally to q'. See Fig. 7.

Therefore, the set of relay nodes is 2-connected.

Now, we show the performance ratio of Algorithm 2.1 is 6. Let H' be the set of relay nodes found in Step 1, i.e. $H'=\cup H_i$ be the optimal solution of the corresponding RNDC

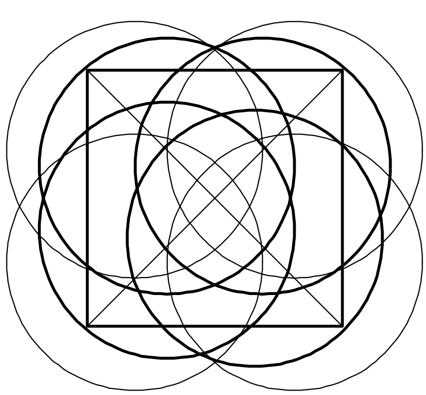

Fig. 6. Cover $2r\times 2r$ cell by eight disks.

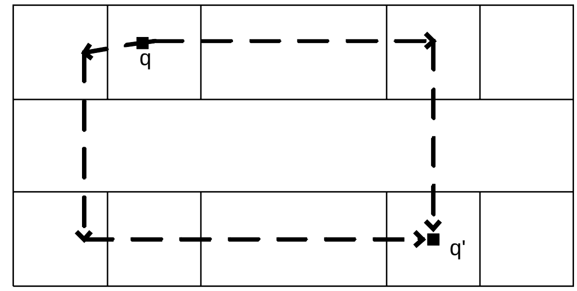

Fig. 7. Two node-disjoint paths between two relay nodes in different cells.

problem. By Corollary 1, we have

$$\frac{|H'|}{|\mathrm{OPT}|} \le (1+1/l)^2 = 4$$

Let H^* solution provided by Algorithm 2.1. By our assumption that each cell contains at least one sensor node, H' has at least two relay nodes for each cell. Suppose we have x rows and y columns of cells, then $|H'|\ge 2xy$ By Step 2, number of extra relay nodes added is at most $(x+1)(y-1)$. WLOG, we assume $x\ge y$, for otherwise, we can switch rows and columns in the algorithm. Hence,

$$\frac{|H^*|}{|H'|} \le \frac{3}{2}$$

Therefore,

$$\frac{|H^*|}{|\mathrm{OPT}|} \le 6$$

□

Algorithm 2.2. (2CRNDC, $l=2$)

Step 0. Divide the region into cells of side length $D=4r$, i.e. $l=2$. For each cell, find all P-positions for relay nodes. Apply a shrink operation.

Step 1. Inside each cell, exhaustively search all 1 through 21-subsets of the P-positions inside (or on) the cell, to find a subset with smallest order, which can cover all the sensor nodes in the cell at least twice.

For each cell B_i, let H_i be the set of relay nodes found for B_i (note that each H_i has at least two relay nodes).

Step 2. Inside each cell B_i, if the relay nodes in H_i are connected but not 2-connected, add a relay node on the horizontal mid-line of B_i, $4r$ units away from the left-top corner of B_i. If they are not connected, then add one more relay node on the vertical mid-line of B_i, $4r$ units away from the left-top corner of B_i (see Fig. 8).

Step 3. Now, we make the set of the chosen relay nodes to be 2-connected (see Fig. 8).

Step 3.1. Connect each row of the H_i's.

Let H_{i+1} be the set of relay nodes to the right of H_i. If H_i and H_{i+1} are not connected, add a relay node, q_i, on the horizontal mid-line of B_i, $4r$ units away from the left-top corner of B_i, and set $H_i=H_i\cup\{q_i\}$. If they are still not connected, add a relay node, q_{i+1}, on the

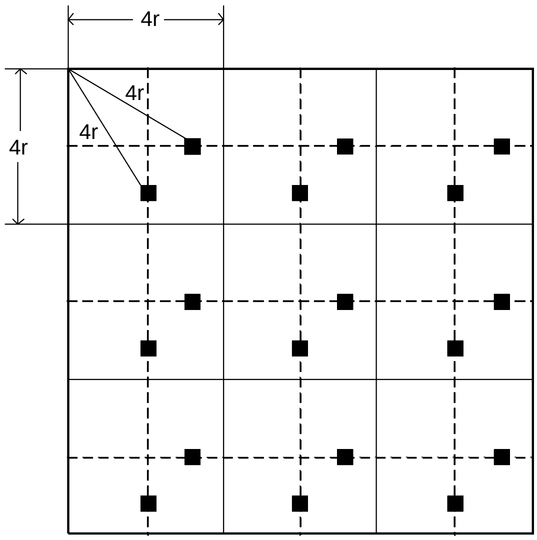

Fig. 8. Possible locations to add extra relay nodes for Steps 2 and 3 of Algorithm 2.2.

horizontal mid-line of B_{i+1}, $4r$ units away from the left-top corner of B_{i+1}, and set $H_{i+1}=H_{i+1}\cup\{q_{i+1}\}$ (notice that we only make them to be connected here, instead of 2-connected).

Repeat Step 3.1, until every row of H_i's is connected.

Step 3.2. Connect each column of the H_i's.

For each H_i, denote LEFT_{B_i} the set of relay nodes in H_i, which are connected to some relay nodes in the set to the left of H_i, i.e. $\mathrm{LEFT}_{B_i}=\{q\in H_i|q$ is connected to a node in the set to the left of $H_i\}$; denote RIGHT_{B_i} the set of relay nodes in H_i, which are connected to some relay nodes in the set to the right of H_i, i.e. $\mathrm{RIGHT}_{B_i}=\{q\in H_i|q$ is connected to a node in the set to the left of $H_i\}$.

If $|\mathrm{LEFT}_{B_i}\cup\mathrm{RIGHT}_{B_i}|>1$, then $H'_i=H_i$. Otherwise, $H_i'=H_i-\mathrm{LEFT}_{B_i}\cup\mathrm{RIGHT}_{B_i}$.

Let H_{i+x} be the set of relay nodes directly under H_i. If H'_i and H'_{i+x} are not connected, then add a relay node, q, on the vertical mid-line of Bi, $4r$ units away from the left-top corner of B_i, and set $H_i=H_i\cup\{q\}$. If they are still not connected, add a relay node, q', on the vertical mid-line of B_{i+x}, $4r$ units away from the left-top corner of B_{i+x}, and set $H_{i+x}=H_{i+x}\cup\{q'\}$.

Repeat Step 3.1, until every column of H_i's is connected.

Theorem 4. *Algorithm 2.2 can always give a solution for the 2CRNDC problem. And its performance ratio is bounded by 4.5.*

Proof. The following needs to be proved for the correctness:

I. All sensor nodes are covered by at least two relay nodes.
II. All relay nodes are 2-connected.

I. Similar to the previous cases, we only need to show that a cell with side-length $4r$ can be covered twice by 21 disks of radius r. By [15], the cell can be covered once by nine disks. We can get another nine disks by shifting the centers of these nine original disks to the right for a small amount, say $r/100$. Then, most part of the cell can be covered twice by these 18 disks, except a strip of width $r/100$ at the left end of the cell.

We use another three disks to cover this strip. Thus, the 21 disks cover the cell twice.

II. Similar to the argument for the correctness of Algorithm 2.1, the set of relay nodes is 2-connected.

Now, we show the performance ratio of Algorithm 2.2 is bounded 4.5. Let H' be the set of relay nodes found in Step 1, i.e. $H'=\cup H_i$. By Corollary 1, we have

$$\frac{|H'|}{|\mathrm{OPT}|}\le(1+1/l)^2=\frac{9}{4}$$

Let H^* be the solution provided by Algorithm 2.2. By our assumption that each cell contains at least one sensor node, H' has at least two relay nodes for each cell. And by the algorithm, we add at most two extra relay nodes for each cell. So,

$$\frac{|H^*|}{|H'|}\le 2$$

Therefore,

$$\frac{|H^*|}{|\mathrm{OPT}|}\le\frac{9}{2}$$

□

Similarly, the optimal solution of 2CRNDC problem is also a solution for the corresponding RNDC problem. Therefore, the performance ratio of our approximation solution against the optimal solution of the corresponding RNDC problem is an upper bound of the actual performance ratio.

Remarks. As mentioned above, the partition factor l plays an important role in the performance ratio of the algorithm. One may think to make l as big as possible to improve the performance ratio. We would like to point out two major tradeoffs for large values of l. First of all, each algorithm needs to perform an exhaustive search for every cell of sidelength $l\cdot 2r$. The greater l is, the more sensor nodes there will be inside each cell. And this means much longer execution time for the exhaustive search. Another tradeoff of larger l is that when the size of cells getting larger, more extra relay nodes need to be put in to guarantee the connectivity (or 2-connectivity), which will increase the performance ratio.

6. Performance evaluations

In this section, we evaluate the performance of our algorithm via simulations. In all simulation scenarios, we use the relay cover size as the performance metric because

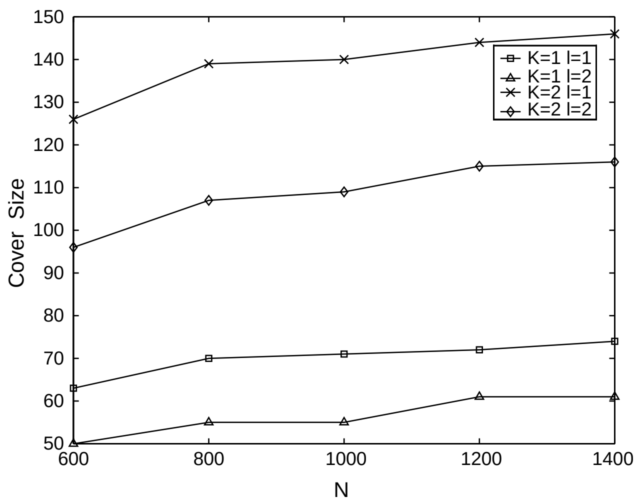

Fig. 9. Relay node cover size computed by different algorithms.

our objective is to minimize the network cost, i.e. the number of relay nodes in the cover, under the constraints that the network can work properly. We consider networks in which sensor nodes are uniformly distributed in a square-shaped playing field with the size of $480 \times 480\ \mathrm{m}^2$. In addition, we assume that all relay nodes' communication ranges are 200 m, because that it is possible that the relay node adopts the same 802.11-based wireless communication system as what is used by the general wireless ad hoc node and the typical communication range for such kind of node is between 200 and 250 m.

In the first simulation, we fix the communication range of each sensor node to be 40 meters since practically this value for the sensor node is between 20 and 40 [12,17].We adjust the network density by changing the number of nodes in the network. Simulations are run on networks with 600, 800, 1200 and 1400 sensor nodes, respectively. Fig. 9 shows the results.

In the figure, l is the partition factor and K stands for the times that each sensor node is covered. It is easy to understand that usually the relay node cover size increases with the increase of the number of sensor nodes. But, we can see that if the network is dense enough, this kind of increase becomes not very substantial. We can imagine that currently we have 1200 sensor nodes in the network and then we add 200 more sensor nodes into the network. It is not necessary to add too many new relay nodes for covering new added sensor nodes since most of them will be close enough to old sensor nodes and can be covered by the existing relay nodes. The most important observation about this figure is that when l is changed from 1 to 2, the cover size decreases dramatically. For example, for the double cover case of network with 1000 nodes, only 109 relay nodes are needed while the cover size is 140 if $l=1$, which involves a 22% improvement. Actually it has been shown that in the worst case, the performance ratio of our approximation algorithm for 2CRNDC problem with $l=2$ is 4.5, while the one with $l=1$ is 6. Simulation results show that in the average case, the approximation algorithm with $l=2$ is also much better than the one with $l=1$. Section 4 explains the corresponding reasons. However, increasing the partition factor will increase the time complexity of the algorithm since exhaustive search in one cell will become much longer. During the simulation, we find out that in the densely distributed large scale network, it will take a very long time to get the optimal solution for a cell if l is greater than 2. So we only consider cases with l no more than 2. In addition, we can see that the size for the double cover is always about twice as large as that for the single cover. So the gain for bringing survivability to the network is comparable to the extra cost needed to be paid.

In the second scenario, we still run the simulation on networks consisting of 600, 800, 1000, 1200 and 1400 sensor nodes. But the sensor's communication range is set to be different values, that is, 24, 30 and 40 m, respectively. Since through experiments in scenario one, we find out that the efficiency can be achieved by choosing bigger partition factor l. Considering the tradeoff between time complexity and performance ratio, $l=2$ should be the best choice. So in the following cases, we only consider algorithms with $l=2$. Actually, both the number and the communication range of sensor nodes are important factors influencing the size of relay node cover since the number of sensor nodes covered by a specific relay node will become less if the sensor nodes' communication range become smaller. According to Fig. 10, for the network having 1000 nodes, comparing to the double cover size of 109 with the communication range equal to 40 m, the double cover size is 252 if the communication range is 24 m, which is more than twice of the former one. Moreover, if the network has 600 sensor nodes and their communication ranges are 40 m, only 96 relay nodes are needed to cover each sensor node at least twice. However, we need 268 relay nodes to double-cover

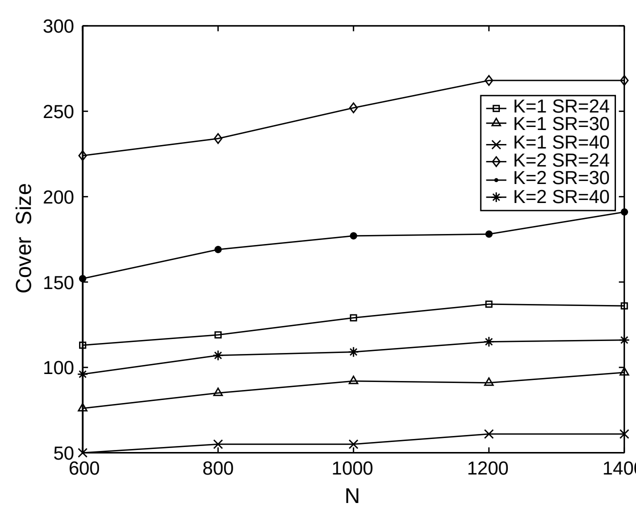

Fig. 10. The influence of sensor communication range (SR) on relay node cover size.

Table 1
Performance ratios of our algorithms

Cases		(500, 24)	(400, 30)	(400, 40)
$K=1$	Our Alg	105	67	48
	Optimal	84	57	39
	Ratio	1.250	1.75	1.231
$K=2$	Our Alg	209	132	94
	Optimal	166	113	75
	Ratio	1.259	1.168	1.253

1200 sensor nodes if their communication ranges are 24 m. So, we can conclude that the influence of the communication range to the cover size is more substantial than the change of the number of sensor nodes. Another interesting observation is that for a specific network, no matter what the communication range of the sensor node is, the double cover size is also always one time larger than that of the single cover.

Comparing proposed algorithms against optimal solutions should be the best way for evaluating their efficiency. But it may be noticed that it is not even possible to give an ILP formulation for our CRNSC and 2CRNDC problems, because it is impossible to identify finite number of possible positions for relay nodes. However, by removing connectivity constrains, we can formulate RNSC and RNDC problems by Integer Linear Programming (ILP). In the simulations, we use CPLEX to compute the optimal solutions for those formulations and employ them as lower bounds on the optimal solutions of the CRNSC and 2CRNDC problems since the size of optimal solutions for RNSC and RNDC problems must be less than those for the corresponding CRNSC and 2CRNDC problems. When the network is dense enough, it takes extremely long time for CPLEX to figure out the optimal solution. Hence, we can only work on relatively small cases in this scenario. The following table shows the performance ratios of our approximation algorithms (K has the same meaning as that in two previous scenarios). We consider three different instances. Firstly, we generate a network with 500 uniformly distributed nodes whose communication range is set to be 24 m. Then we uniformly generate two networks, both of which have 400 nodes. But, communication ranges are 30 and 40, respectively. We apply our approximation algorithms with partition factor equal to 2 to all of instances and then observe how close the sizes of our solutions could be to the optimal ones (Table 1).

Based on our theoretic analysis about performance ratios in the last section, our approximation algorithms with partition factor $l=2$ for both problems achieve a performance ratio of 4.5 in the worst cases. From the table, we can see that on the average cases, performance ratios of our algorithms are much better than the bound, 4.5, no matter for the single or double covering. They are only 1.26 or even less under different network instances. So, the sizes of solutions given by our approximation algorithms are fairly close to that of optimal solutions in average cases.

7. Conclusions

In this paper, we have formulated two optimization problems for relay node placement in large scale sensor networks, one is called *Connected Relay Node Single Cover* (*CRNSC*) problem and another is *2-Connected Relay Node Double Cover* (*2CRNDC*) problem. Two polynomial time approximation algorithms are presented to solve the CRNSC problem. We show that the size of the CRNSC given by the first approximation algorithm is bounded by eight times that of the optimal solution and the second one achieves a performance ratio of 4.5. Moreover, we propose two approximation algorithms to solve the 2CRNDC problem and we also prove that performance ratios associated with them are 6 and 4.5, respectively. In simulations, we find out the approximation algorithms with partition factor l equal to 2 tradeoff the time complexity and efficiency well. The cover sizes given by them are much smaller than algorithms with $l=1$ on any specific instance. We also study the influence of the network size and the communication range of the sensor node to the cover size given by our algorithms. In addition, our simulation results show that sizes of solutions provided by our approximation algorithms are very close to that of optimal solutions under different instances.

In the future, We will extend our work to the more general K-Connected Relay Node K Cover (KCRNKC) problem, i.e. covering each sensor node at least K times and the network composed of relay nodes forms a K-connected topology, where K can be any positive integer.

Acknowledgements

Jian Tang's research was supported in part by ARO grant W911NF-04-1-0385 and NSF grant CCF-0431167.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Computer Networks Journal 38 (2002) 393–422.
[2] J.A. Bondy, U.S.R. Murthy, Graph Theory with Applications, North Holland, New York, 1976.
[3] X. Cheng, D.Z. Du, L. Wang, B. Xu, Relay sensor placement in wireless sensor networks, ACM Wireless networks submitted for publication. Also available at http://citeseer.nj.nec.com/cheng01relay.html
[5] R.J. Fowler, M.S. Paterson, S.L. Tanimoto, Optimal packing and covering in the plane are NP-complete, Information Processing Letter 12 (1981) 133–137.

[6] Y. Gao, K. Wu, F. Li, Analysis on the redundancy of wireless sensor networks, Proceedings of ACM WSNA'2003.

[7] H. Gupta, S.R. Das, Q. Gu, Connected sensor cover: self-organization of sensor networks for efficient query execution, Proceedings of ACM MOBIHOC'2003, pp. 189–200.

[8] G. Gupta, M. Younis, Fault-tolerant clustering of wireless sensor networks, Proceedings of IEEE WCNC'2003, pp. 1579–1584.

[9] G. Gupta, M. Younis, Load-balanced clustering of wireless sensor networks, Proceedings of IEEE ICC'2003, pp. 1848–1852.

[10] W. Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy efficient communication protocols for wireless microsensor networks, Proceedings of HICSS'2000, pp. 3005–3014.

[11] D.S. Hochbaum, W. Maass, Approximation schemes for covering and packing problems in image processing and VLSI, Journal of ACM 32 (1985) 130–136.

[12] C. Intanagonwiwat, R. Govindan, D. Estrin, Directed diffusion: a scalable and robust communication paradigm for sensor networks, Proceedings of ACM MOBICOM'2000, pp. 56–67.

[13] S. Lindsey, C. Raghavendra, K.M. Sivalingam, Data gathering algorithm in sensor networks using energy metrics, IEEE Transactions on Parallel and Distributed System 13 (2002) 924–935.

[14] G. Lin, G. Xue, Steiner tree problem with minimum number of steiner points and bounded edge-length, Information Processing Letters 69 (1999) 53–57.

[15] K.J. Nurmela, P.R.J. Ostergard, Covering a square with up to 30 equal circles, Research Report A62, Helsinki University of Technology, 2000. Also available at http://www.tcs.hut.fi/Publications/info/bibdb.HUT-TCS-A62.shtml

[16] J. Pan, Y.T. Hou, L. Cai, Y. Shi, S.X. Shen, Topology control for wireless sensor networks, Proceedings of ACM MOBICOM'2003, pp. 286–299.

[17] C. Schurgers, V. Tsiatsis, S. Ganeriwal, M. Srivastava, Topology management for sensor networks: exploiting latency and density, Proceedings of ACM MOBIHOC'2002, pp. 135–145.

[18] A. Srinivas, E. Modiano, Minimum energy disjoint path routing in wireless ad-hoc network, Proceedings of ACM MOBICOM'2003, pp. 122–133.

[19] J.W. Suurballe, Disjoint paths in a network, Networks 4 (1974) 125–145.

[20] D. Tian, N.D. Georganas, A coverage-preserving node scheduling scheme for large wireless sensor networks, Proceedings of ACM WSNA'2002, pp. 32–41.