Fault-Tolerant Relay Node Placement in Wireless Sensor Networks

Hai Liu, Peng-Jun Wan, and Xiaohua Jia

Department of Computer Science, City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong, P.R. China
{liuhai,pwan,jia}@cs.cityu.edu.hk

Abstract. The paper addresses the relay node placement problem in two-tiered wireless sensor networks. Given a set of sensor nodes in an Euclidean plane, our objective is to place minimum number of relay nodes to forward data packets from sensor nodes to the sink, such that: 1) the network is connected, 2) the network is 2-connected. For case one, we propose a $(6 + \varepsilon)$-approximation algorithm for any $\varepsilon > 0$ with polynomial running time when ε is fixed. For case two, we propose two approximation algorithms with $(24+\varepsilon)$ and $(6/T + 12+\varepsilon)$, respectively, where T is the ratio of the number of relay nodes placed in case one to the number of sensors.

Keywords: sensor networks, fault-tolerant, relay node placement.

1 Introduction and Related Work

A sensor network is composed of a large number of sensor nodes that can be deployed on the ground, in the air, in vehicles, inside buildings or even on bodies. Sensor networks are widely deployed in environment monitoring, biomedical observation, surveillance, security and so on [4, 5]. Unlike the cellular networks and MANETs where there is unlimited energy supply in base stations or by batteries that can be replaced as needed, nodes in sensor networks have very limited energy supply and their batteries cannot usually be replaced due to special environments [6]. Since sensors' energy cannot support long range communication to reach a sink which is generally far away from the data source, multi-hop wireless connectivity is required to forward data to the remote sink. It is a key problem regarding how to gather data packets from sensor nodes to the sink in applications.

The two-tiered network architecture was proposed in [1, 2]. They employed relay nodes as gateways that are more powerful than sensor nodes in terms of energy storage, computing and communication capabilities. The network is partitioned into a set of clusters. The relay nodes act as cluster heads and they are connected with each other to perform the data forwarding task. (In the following, relay node, gateway and cluster-head refer to the same thing in the two-tiered sensor networks.) Each cluster has only one cluster-head and each sensor belongs to at least one cluster, such that sensor nodes can switch to

L. Wang (Ed.): COCOON 2005, LNCS 3595, pp. 230–239, 2005.

backup cluster heads when current communication fails. In each cluster, sensor nodes collect raw data and report to the cluster-head. The cluster-head analyzes the raw data, extracts useful information, and then generates outgoing packets with much smaller total size to the sink through multi-hop path [3, 12]. Topology control in two-tiered wireless sensor networks has been discussed in [12]. The objective of the paper is to maximize the topological lifetime of the network, where the topological lifetime is defined as the lifetime of a network with regard to a given mission and the placement of sensor nodes, application nodes and base-stations. The authors proposed approaches to maximize the topological lifetime by arranging the base-stations location and the relay allocation. Similar to the works in [1, 2], two-tiered wireless sensor network architecture was proposed in [11] as a solution for structural health monitoring. The communication protocols used in the lower-tier (in clusters) and in the upper-tier (inter-clusters) were both described. Analysis in [11] showed that the maximum total number of sensor nodes that a network can handle is about 2000 4000 under current wireless data rates of 10Mbps.

Since a large number of nodes cooperate with each other in the network, fault-tolerance is one of important issues in wireless sensor networks. Communication faults in sensor networks can be caused by hardware damage, energy depletion, harsh environment conditions and malicious attacks. A fault in transmitter can cause the relay nodes to stop transmitting tasks to the sensors as well as relaying the data to the sink. Data sent by the sensors will be lost if the receiver of a relay node fails. So, a communication link failure to a sensor requires the sensor to be re-allocated to other cluster-head within communication range. If faults occur in inter-cluster-heads, the two corresponding cluster-head should be re-connected by another multi-hop path. That is, in order to handle general communication faults, there should be at least two node-disjoint paths between each pair of relay nodes in the network.

An intuitive objective of relay node placement is to place the minimum number of relay nodes to make the network connected, such that each sensor is covered by some relay nodes and all data packets can be gathered to the sink through these relay nodes (node a is covered by node b means that node a is within the communication range of node b). E. Biagioni et al. investigated placement problem of sensors in [8]. Three parameters were considered in the placement of sensors: resilience to single node failure, coverage of area of interest and minimizing the number of sensors. The authors showed that the choice of placement depends on sampling distance and communication radius. Different from the placement of sensor nodes, the placement problem of relay nodes in two-tiered sensor networks was discussed in [7]. The problem is to place the minimum number of relay nodes such that 1) each sensor node can communicate with at least two relay nodes and 2) the network of the relay nodes is 2-connected. An approximation algorithm was proposed and ratio was proved within $O(D \log n)$, where n is the number of sensor nodes in the network, and D the diameter of the network. Obviously, the proposed ratio is not a constant, which is a function of the size of input.

In this paper, we propose three approximation algorithms for the minimum relay-node placement problem (**MRP** in short). We discuss two cases of the MRP: 1) the network is connected and 2) the network is 2-connected. We proposed a $(6 + \varepsilon)$-approximation algorithm for case one. We further proposed a $(24+\varepsilon)$-approximation algorithm and a $(6/T+12+\varepsilon)$-approximation algorithm for case two, respectively, for any $\varepsilon > 0$, where T is the ratio of the number of relay nodes placed to the number of sensors in case one.

2 System Model and Problem Specification

We first describe the MRP in two tiered wireless sensor networks, and then give a formal formulation of the problem.

Given a set of sensor nodes that are randomly distributed in a region and their location, in order to gather data packets from sensor nodes to the sink, we need to place some relay nodes to forward the data packets, such that each sensor node is covered by at least one relay node. Since sensor nodes have limited computing and communication capability, especially very limited energy resource, we assume that sensor nodes only report data packets to relay nodes within their communication range but not participate in data forwarding. That is, there is no direct link between any pair of sensor nodes. We further assume that all sensor nodes and all relay nodes have the same communication radius. The problem is to

1) place the minimum number of relay nodes in the region, such that the network (including sensor nodes and relay nodes) is connected;

2) place the minimum number of relay nodes in the region, such that there exists at least two node-disjoint paths between any pair of nodes (sensor nodes or relay nodes). That is, the whole network is 2-connected.

There are several major differences between our problem and the problem addressed in [7]. Firstly, the network of sensor nodes was assumed to be 2-connected in [7], we do not need this assumption. Secondly, sensor nodes also participate forwarding of data packets in [7] while they are not in our problem. Note that sensor nodes have very limited energy resource and computing capability. If they participate in data forwarding, it may cause early depletion of sensors, and make the network disconnected. Finally, placement of relay nodes in [7] only makes the network of relay nodes 2-connected, but we make the whole network 2-connected. Since our problem is more general than that in [7], our solution is also applicable to the problem in [7].

Now, we formally give the definition of the MRP problem.

Definition. Minimum Relay-node Placement problem (MRP): Given a set of sensor nodes S in a region and a uniform communication radius d, the problem is to place a set of relay nodes R, such that 1) the whole network G is connected and 2) G is 2-connected. The objective of the problem is to:

$$Minimize|R|$$

where $|R|$ denotes the number of relay nodes in R.

We first give a $(6+\varepsilon)$-approximation solution for the case one of MRP (MRP-1 in short), and then propose a $(24+\varepsilon)$-approximation and a $(6/T+12+\varepsilon)$-approximation solutions for MRP-2 by adding some relay nodes to make the network 2-connected.

3 Solution to MRP-1

Our solution is based on two foundational works. The first is the covering with disks problem. Given a set of points in the plane, the problem is to identify the minimum set of disks with prescribed radius to cover all the points. In [10], a polynomial time approximation scheme (PTAS) for this problem was proposed. That is, for any given error ε, the ratio of the solution found by the scheme to the optimal solution is not larger than $(1+\varepsilon)$. The running time is polynomial when ε is fixed. We call the scheme ***min-disk-cover scheme***. Different from covering with disks problem, MRP requires not only covering all sensor nodes, but also requires connection of the network, such that data packets can be gathered to the sink.

The other problem is the Steiner tree problem with minimum number of Steiner points (STP-MSP). Given a set of terminals in the Euclidean plane, the problem is to find a Steiner tree such that each edge in the tree has length at most d and the number of Steiner points is minimized. Du *et al.* proposed a 2.5-approximation algorithm for the STP-MSP [9]. We call the algorithm ***STP-MSP algorithm***. In our problem, sensor nodes do not relay messages for other nodes. So STP-MSP algorithm cannot be used to MRP problem directly.

Based on the min-disk-cover scheme and the STP-MSP algorithm, our solution for MRP-1 is composed in two steps. In step one, for any give error ε, we use min-disk-cover scheme to find a set of relay nodes that cover all sensor nodes. Note that the network of these relay nodes may not be connected if distance between them is larger than d. In order to connect these relay nodes, more relay nodes are needed. In step two, we run the STP-MSP algorithm and place additional relay nodes to get a Steiner tree. Finally, the Steiner tree and the sensor nodes form the connected network we desire. The algorithm is formally presented as follows.

$(6+\varepsilon)$-approximation algorithm for MRP-1

1. Place a set of relay nodes R_1 by using the min-disk-cover scheme, such that for $\forall s \in S, \exists r \in R_1$, and r cover s.
2. Place another set of relay nodes R_2 by running STP-MSP algorithm with input R_1.
3. Output $R_1 + R_2$.

Let R denote the set of all relay nodes we place in the above algorithm. Without loss of generality, we assume $R_1 \bigcup R_2 = \Phi$, i.e., $R = R_1 + R_2$ (or $R_1 \bigcup R_2$). Then the total number of relay nodes we place is $|R_1 + R_2|$. In the following theorem, we show that the algorithm for MRP-1 has ratio $(6+\varepsilon)$ for any $\varepsilon > 0$.

Theorem 1. Let R be our solution to MRP-1 and R^{opt} be the optimal solution to MRP-1. Then $\frac{|R|}{|R^{opt}|} \leq (6+\varepsilon)$.

Proof. Let R_1^{opt} denote the minimum set of relay nodes that cover S. Obviously, we have $|R_1^{opt}| \leq |R^{opt}|$. Since R_1 is the solution of PTAS to covering with disks problem, we have

$$|R_1| \leq (1+\varepsilon)|R_1^{opt}|. \tag{1}$$

Let R_2^{opt} denote the minimum set of relay nodes that makes R_1 connected. Since R_2 is the 2.5-approximation solution to STP-MSP problem, we have

$$|R_2| \leq 2.5|R_2^{opt}|. \tag{2}$$

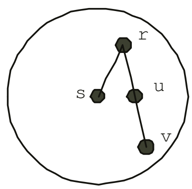

Fig. 1. Communication circle of sensor s.

In order to find out the relationship between R_2^{opt} and R^{opt}, we consider each relay node $r \in R_1$. For any $r \in R_1$, there must be at least one sensor node s which is covered by r (Otherwise, we can remove this useless r from R_1). We consider the communication circle of s (See Fig. 1). Note that for $\forall s \in S, \exists v \in R^{opt}$ such that both r and v can cover s. That is, relay nodes r and v are both in the communication circle of sensor s. So we have $d(r,v) \leq 2d$, where $d(r,v)$ is the Euclidean distance between r and v. If we place another relay node u in the middle point of edge (v,r), we have $d(u,v) = d(r,u) \leq d$. That is, node v can communicate with node r via node u.

For any relay node $r \in R_1$, we place another relay node according to the above description, such that R^{opt} can communicate with every relay node in R_1. That is, R^{opt} and these added relay nodes make R_1 connected. Note that R_2^{opt} is the minimum set of relay nodes that makes R_1 connected, and the number of added relay nodes is equal to $|R_1|$. So we have

$$|R_2^{opt}| \leq |R^{opt}| + |R_1|. \tag{3}$$

According to (1),(2),(3), the total number of relay nodes placed by the algorithm to MRP-1 is:

$$\begin{aligned} |R_1| + |R_2| &\leq (1+\varepsilon)|R_1^{opt}| + 2.5(|R^{opt}| + |R_1|) \\ &\leq (1+\varepsilon)|R^{opt}| + 2.5(|R^{opt}| + (1+\varepsilon)|R^{opt}|) \\ &\leq 6|R^{opt}| + 3.5\varepsilon|R^{opt}| \end{aligned}$$

Note that ε is an arbitrary positive number.
That is, $\frac{|R|}{|R^{opt}|} = \frac{|R_1|+|R_2|}{|R^{opt}|} \leq (6+\varepsilon)$

4 Solution to MRP-2

MRP-2 requires the network is 2-connected. It means that when one relay node fails, the network is still connected and data packet gathering operation can be carried out. Let T denote the ratio of the number of relay nodes place in MRP-1 to the number of sensors. That is, $T = |R|/|S|$. We first propose a $(24 + \varepsilon)$-approximation algorithm for general case of MRP-2, then improve the ratio to $(6/T + 12 + \varepsilon)$ if $T > 1/2$.

4.1 $(24 + \varepsilon)$-Approximation

Based on the solution R to MRP-1, our main idea is to add some backup nodes to the communication circle of each relay node r in R, such that the whole network is 2-connected. We first study the features these backup nodes should have. For$\forall r \in R$,

C_1. The backup nodes should cover all nodes in the communication circle of r. The first purpose of this condition is to make any node in the circle can switch to one of the backup nodes when r is out of service. The other purpose is to make at least one of the backup nodes communicate with outside when r is out of service. The first is for collecting messages purpose and the second is for forwarding messages purpose.

C_2. The backup nodes in the communication circle of r should communicate with each other. This is because that condition 1 only guarantees there exist at least one backup node can communicate with outside, but some other nodes in the circle may not communicate with outside if r is out of service. That is, the purpose of this condition is to make all nodes in the circle can communicate with outside.

We call the above two features ***2-connected conditions***. Are these two conditions sufficient? The following lemma shows that for each relay node r in R, if the backup nodes in the communication circle of r satisfy the 2-connected conditions, the network is 2-connected.

Lemma 1. Let $R^{'}$ denote the set of backup nodes in all communication circles of R. For $\forall r \in R$, if the backup nodes in the communication circle of r satisfy the 2-connected conditions, then the resulting network $S + R + R^{'}$ is 2-connected.

Proof. For $\forall v \in S + R + R^{'}$, we assume v fails. There are three cases: $v \in S, v \in R$ or $v \in R^{'}$. We prove the lemma case by case.

If $v \in S$, note that sensor nodes only report data packets to its cluster-head but not participate the data forwarding task, the network is connected.

If $v \in R^{'}$, since R is the solution to MRP-1 and the network $S + R$ is connected, there is no need to use backup nodes in this case.

If $v \in R$, according to the 2-connected conditions, all nodes in the communication circle of v are covered by backup nodes and can send their data packets to the outside. So the network is connected.

Since v is an arbitrary node in , Lemma 1 is proved.

According to Lemma 1, our algorithm is to add minimum number of backup nodes to the communication circle of each $r \in R$, such that the backup nodes

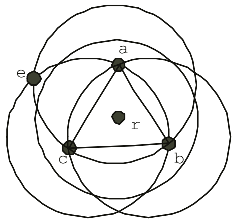

Fig. 2. Adding backup nodes to the communication circle of r.

satisfy the 2-connected conditions. For the condition C_1, obviously, adding two backup nodes can not cover the communication circle of r. That is, at least three backup nodes are needed in each communication circle. Based on R, our algorithm for MRP-2 is to add three backup nodes in the communication circle of each relay node in R, such that the network is 2-connected. The algorithm is formally presented as follows.

$(24 + \varepsilon)$-approximation algorithm for MRP-2

1. Place a set of relay nodes R by running $(6 + \varepsilon)$-approximation algorithm for MRP-1, such that $S + R$ is connected.
2. Place three backup nodes in the communication circle of each $r \in R$. The three backup nodes are placed on the three vertices of an equilateral triangle with length d. The center of the equilateral triangle is in r. See Fig. 2, where a, b, c are backup nodes. We denote the set of all backup nodes in this step by $R^{'}$.
3. Output $R + R^{'}$.

The following theorem claims the correctness of the algorithm.

Theorem 2. The set of backup nodes $R^{'}$ in the algorithm satisfy the 2-connected conditions and the final network $S + R + R^{'}$ is 2-connected.

Proof. In the following proof, we use *circlex* denote the communication circle of relay node x. We will prove the 2-connected conditions are satisfied one by one.

First, it is not difficult to see that the three new circles can cover the circle r. That is to say, any node in circle r is also in one of the three circles. See Fig. 2, since triangle abc is an equilateral triangle, we have

$$\angle ear = \angle eac + \angle car = 60^\circ + 30^\circ = 90^\circ$$
$$d(e, a) = d,$$
$$\text{then } d(e, r) > d.$$

It means that the circle a crosses the circle c on point e, which is outside the circle r. Likewise, points of intersection of circle r and circle b, circle c are both outside the circle r. So all nodes in the circle r can be covered by the union of circle a, circle b and circle c. The first condition C_1 is satisfied.

Second, since edge length of each equilateral triangle is d, which equals the communication radius of relay nodes. That means the three backup nodes can communicate with each other. It satisfies the condition C_2.

Associated with the Lemma 1, Theorem 2 is proved.

The following theorem states that the ratio of the above algorithm to MRP-2 is at most $(24 + \varepsilon)$.

Theorem 3. Let $R + R^{'}$: our solution to MRP-2, R^{2-opt}: the optimal solution to MRP-2, we have $\frac{|R+R^{'}|}{|R^{2-opt}|} \leq (24 + \varepsilon)$.

Proof. Note that for each communication circle of relay node in R, we add three backup nodes. We have $|R^{'}| = 3|R|$. So the total number of relay nodes we used to make the network 2-connected is:

$$\begin{aligned}|R + R^{'}| &= |R| + |R^{'}| = 4|R| \\ &\leq 4(6 + \varepsilon)|R^{opt}| \\ &\leq 4(6 + \varepsilon)|R^{2-opt}| \\ &= 24|R^{2-opt}| + 4\varepsilon|R^{2-opt}|\end{aligned}$$

Note that ε is an arbitrary positive number.

That is, $\frac{|R+R^{'}|}{|R^{2-opt}|} \leq (24 + \varepsilon)$

In the next section, we will show that the ratio can be improved if $T > 1/2$.

4.2 $(6/T + 12 + \varepsilon)$-Approximation

After placing relay nodes by running the algorithm for MRP-1, the network is formed as a connected tree. Our main idea is to place backup nodes on each link of the tree to make the network 2-connected. The algorithm is formally presented as follows.

$(6/T + 12 + \varepsilon)$**-approximation algorithm for MRP-2**

1. Place a set of relay nodes R by running the $(6 + \varepsilon)$-approximation algorithm for MRP-1, such that the tree $S + R$ is connected.
2. Place one backup node in the middle point of each link in the tree. We denote the set of all backup nodes in this step by $R^{''}$.
3. Output $R + R^{''}$.

The following theorem claims the correctness of the algorithm.

Theorem 4. The network $S + R + R^{''}$ is 2-connected.

Proof. To prove the theorem, we need to prove for $\forall v \in S + R + R^{''}$, if v fails, the resulting network is still connected. There are three cases: $v \in S, v \in R$ or $v \in R^{''}$. We prove the theorem case by case.

If $v \in S$, note that sensor nodes only report data packets to its cluster-head but not participate the data forwarding task, the network is still connected.

If $v \in R^{''}$, since R is the solution to MRP-1 and the network $S + R$ is connected, there is no need to use backup relay nodes in this case.

If $v \in R$, since v is a relay node, without loss of generality, we assume v forwards data packets from node u to node w (see Fig. 3, a, b are backup nodes). According to the above approximation algorithm, we have

$$d(a, v) = d(b, v) = 0.5d,$$

$$\text{then, } d(a, b) \leq d.$$

That is, the backup nodes a and b can communicate with each other. So the data packets can be forwarded to w via $u \to a \to b \to w$. The network is connected.

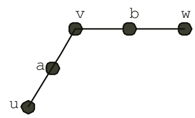

Fig. 3. Adding backup nodes a, b on the links of v.

Since v is an arbitrary node in $S + R + R^{''}$, Theorem 4 is proved.

The following theorem states that the ratio of the solution to MRP-2 can be improved to $(6/T + 12 + \varepsilon)$ if $T > 1/2$.

Theorem 5. Let $R + R^{''}$: our solution to MRP-2, R^{2-opt}: the optimal solution to MRP-2, we have $\frac{|R+R^{''}|}{|R^{2-opt}|} \leq (6/T + 12 + \varepsilon)$.

Proof. Since we place a node on each link in the tree which contains $|R + S|$ nodes in total, we have

$$|R^{''}| \leq |R + S| \leq |R| + |R|/T,$$

$$\text{then } |R + R^{''}| \leq 2|R| + |R|/T \leq (2 + 1/T)(6 + \varepsilon)|R^{opt}|$$
$$\leq (6/T + 12 + (2 + 1/T)\varepsilon)|R^{2-opt}|.$$

Note that ε is an arbitrary positive number.

That is, $\frac{|R+R^{''}|}{|R^{2-opt}|} \leq (6/T + 12 + \varepsilon)$.

5 Conclusions

The fault-tolerant relay node placement problem in two-tiered wireless sensor networks was studied in the paper. Given a set of sensor nodes in a Euclidean plane, our objective is to place minimum number of relay nodes, such that 1) the network is connected and 2) the network is 2-connected.

Given an arbitrary positive number ε, we proposed a $(6 + \varepsilon)$-approximation algorithm for the problem one. The running time is polynomial when ε is fixed. We further proposed a $(24 + \varepsilon)$-approximation algorithm and a $(6/T + 12 + \varepsilon)$-approximation algorithm for problem two, where T is the ratio of the number of relay nodes placed in problem one to the number of sensors.

References

1. G. Gupta, M. Younis, "Fault-Tolerant clustering of wireless sensor networks",*Proceeding of IEEE WCNC'2003*, pp. 1579-1584.
2. G. Gupta, M. Younis, "Load-Balanced clustering of wireless sensor networks",*Proceeding of IEEE ICC'2003*, pp. 1848-1852.

3. J. Pan, Y. T. Hou, L. Cai, Y. Shi, S. X. Shen, "Topology control for wireless sensor networks",*Proceeding of ACM MOBICOM'2003*, pp. 286-299.
4. C.-Y. Chong and S. P. Kumar, "Sensor Networks: Evolution, Opportunities, and Challenges",*Proc. of the IEEE*, Vol 91, No. 8, Aug. 2003.
5. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A Survey on Sensor Networks",*IEEE Communications Magazine*, Aug. 2002.
6. P. Rentala, R. Musunuri, S. Gandham and U. Saxena, "Survey on Sensor Networks",http://citeseer.nj.nec.com/479874.html.
7. B. Hao, J. Tang and G. Xue, "Fault-tolerant relay node placement in wireless sensor networks: formulation and approximation", accepted by HPSR'2004.
8. E. Biagioni and G. Sasaki, "Wireless sensor placement for reliable and efficient data collection",*In Proceeding of the Hawaii International Conference on Systems Sciences*, Jan. 2003.
9. D. Du, L. Wang and B. Xu, "The Euclidean Bottleneck Steiner Tree and Steiner Tree with Minimum Number of Steiner Points", Computing and Combinatorics,*7th Annual International Conference COCOON 2001*, Guilin, China, Aug. 2001 Proceedings.
10. D. S. Hochbaum and W. Maass, "Approximation Schemes for Covering and Packing in Image Processing and VLSI",*Journal of the ACM (JACM)*, Vol 32, Issue 1, pp 130-136, Jan. 1985.
11. V. A. Kottapalli, A. S. Kiremidjian, J. P. Lynch, E. Carryer, T. W. Kenny, K. H. Law and Y. Lei, "Two Tiered Wireless Sensor Network Architecture for Structural Health Monitoring",*10th Annual International Symposium on Smart Structures and Materials*, San Diego, CA, USA, March 2-6, 2003.
12. Jianping Pan, Thomas Hou, Lin Cai, Yi Shi and Sherman X. Shen, "Topology Control for Wireless Sensor Networks", *MobiCom'03*, September 14-19, 2003, San Diego, California, USA.