2006

Optimization strategies for two-tiered sensor networks.

Ataul Bari
University of Windsor

Optimization Strategies for Two-tiered Sensor Networks

by

Ataul Bari

A Thesis

Submitted to the Faculty of Graduate Studies and Research

through Computer Science

in Partial Fulfillment of the Requirements for

the Degree of Master of Science at the

University of Windsor

Windsor, Ontario, Canada

Your file Votre référence
ISBN: 978-0-494-17074-8
Our file Notre référence
ISBN: 978-0-494-17074-8

Abstract

Sensor nodes are tiny, low-powered and multi-functional devices operated by lightweight batteries. Replacing or recharging batteries of sensor nodes in a network is usually not feasible so that a sensor network fails when the battery power in critical node(s) is depleted. The limited transmission range and the battery power of sensor nodes affect the scalability and the lifetime of sensor networks. Recently, relay nodes, acting as cluster heads, have been proposed in hierarchical sensor networks. The placement of relay nodes in a sensor network, such that all the sensor nodes are covered using a minimum number of relay nodes is a NP-hard problem. We propose a simple strategy for the placement of relay nodes in a two-tiered network that ensures connectivity and fault tolerance. We also propose two ILP formulations for finding the routing strategy so that the lifetime of any relay node network may be maximized.

Dedication

To my parents, Salam and Salma, my wife, Limi, my daughters, Chhoaa and Chhuti, my son, Prashna, my sister, Lovely, her husband, Lenin, and my nephews, Shaon and Ayon.

Acknowledgements

I like to take this opportunity to thank my thesis supervisors, Dr. Subir Bandyopadhyay and Dr. Arunita Jaekel, for their guidance and cooperation throughout my graduate studies. This work could not have been achieved without their help and support. I would also like to thank Dr. Yash P. Aneja and Dr. Akshaia Aggarwal for their valuable time, cooperation and thoughtful suggestions.

I wish to express my sincerest gratitude to my wife and kids for their understanding and cooperation all the way. My special thanks to my nephew, K. Raiyan Kamal, whose active support and interest in my work has always been a source of encouragement to me.

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Sensor Networks

A *sensor network* is an interconnection of tiny, lightweight, energy-constrained devices, known as *sensor nodes*, and is usually deployed to monitor some kind of physical phenomena from the territory of its deployment. For example, a sensor network may be deployed to monitor the humidity or the temperature of a certain region, or it may be deployed to detect the presence or absence of some objects, as well as the movement of objects within the area being monitored. Recent technological advances in the field of micro-electro-mechanical systems (MEMS) have made the development of such tiny, low-cost, low-powered and multi-functional sensor devices technically and economically feasible [3], [9]. These nodes are usually equipped with a sensing unit, a processing unit, a memory unit and a RF communication unit.

The data generated by each sensor by sensing its vicinity is required to be sent to a central point, known as *Base Station (BS)* or *sink*. The base station is not power constrained and its location is usually fixed. A general layout of a sensor network, including the sensor nodes and a base station, is shown in Fig. 1.1. The nodes in a sensor network are deployed inside or very close to the phenomenon being monitored, so that the sensing task can be carried out effectively. The placement of sensor nodes in a network can be pre-determined (e.g. the deployment of a sensor network in a factory or in the body of a human, an animal or a robot) or random (e.g. the deployment of nodes by dropping them from a helicopter/airplane or delivering them in an artillery shell or in a missile) [3], [9]. The data from the sensor nodes is collected at the *BS*. This data may be aggregated and forwarded to the user, possibly using the Internet, where it can be further analyzed and useful information can be extracted.

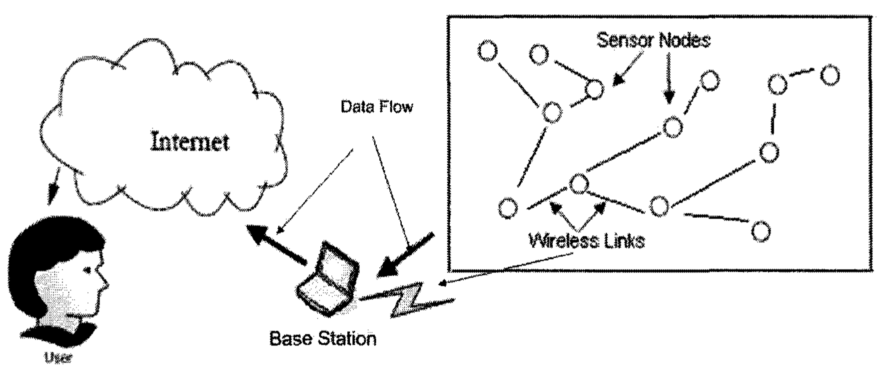


Figure 1.1: A general layout of sensor network

Although the capability of an individual sensor node is limited, a sensor network is usually able to perform bigger tasks through the collaborative efforts of a large number of sensor nodes (hundreds or even thousands) that are densely deployed within the sensing field [2], [3], [9]. There is a wide range of applications, for both military and civil purposes, where the use of sensor networks can be very useful.

Sensor networks pose many challenges in design, operation and maintenance in each layer of the networking protocol stack. Some important issues in the design of sensor networks include [3], [4]:

- *Network deployment in ad hoc manner:* The nodes in a sensor network, deployed in remote areas, need to self-configure and self-organize themselves to form networks.
- *Unattended operation with limited battery power:* Replacing or recharging batteries in sensor networks is usually not feasible, either physically or economically, so that, in many cases, the lifetime of a sensor network expires as soon as *critical* node(s) runs out of battery power [21], [40].
- *Changes in network condition:* Sensor networks need to be adaptive to node failure(s), node mobility and link failures.
- *Scalability:* As the size of networks may vary from one application to another, the protocols need to be scalable.
- *Connectivity:* The system needs to ensure that all the nodes are connected even in the event of failures.
- *Coverage:* As each sensor node can only cover a limited physical area around its vicinity, the entire area to be monitored needs to be covered by the nodes in the sensor network.

- *Data Aggregation:* To reduce the energy dissipated by the transmitting node, the volume of data should be reduced as much as possible.

- *Quality of Service:* There may be a trade-off between the quality of the result and the conservation of energy.

Many routing algorithms have been proposed in the past few years to address the challenges of routing in sensor networks. In general, these protocols can be broadly classified into two major categories, based on the network structure [4], as follows:

i) routing for *flat-architecture* [22], [28], [31], and

ii) routing for *hierarchical-architecture* [21], [33], [34].

In a network based on the flat-architecture, all nodes are treated equally so that each sensor node is responsible for sensing the environment and forwarding its own data as well as data from other nodes, which are using this node as an intermediate node in a multi-hop path towards the base station. In a hierarchical architecture, the network is organized as a number of clusters and each sensor node belongs to only one cluster. Certain nodes are treated as *cluster heads* and have some additional responsibilities (e.g. data gathering, data aggregation and routing) compared to the regular nodes.

In sensor networks, all data flow from the sensor nodes towards the base station, whose location is usually fixed. The transmission power dissipated by a source node to transmit each bit of data to a destination node increases significantly with the distance between the source and the destination [3], [9], [13], [16], [21]. As a result, the use of multi-hop paths has been proposed for conserving energy, in both flat and hierarchical architectures [22], [25], [26], [28], [48]. In the multi-hop routing scheme, nodes located further away from the base station use some intermediate nodes to forward the data to the base station. In such a data-gathering model, it is possible that some nodes are required to relay more data, which they have received from the neighboring nodes, compared to other nodes. Therefore, these nodes may dissipate energy at higher rates than the nodes which are not relaying (or relaying very little) data from other nodes. This uneven energy dissipation among the nodes may lead to the faster "death" of some nodes in the network due to the depletion of the batteries of these nodes, assuming that initial energy provisioning for all nodes are equal. Such unbalanced energy dissipation has an undesirable effect on the functionality of the sensor networks, as the inoperative node(s) will not be able to perform either sensing or routing. This can cause the entire network to prematurely lose its usefulness, even though many other nodes in the networks still retain power. Therefore, a careful load distribution scheme can be effective to prolong the useful lifetime of the network.

1.1.1 Relay Nodes

One method, that has been used to address the issue of uneven energy consumption, is to deploy some special nodes, called *relay nodes* (also called *Gateway nodes* and *Aggregating and Forwarding nodes (AFN))* [8], [14], [16], [25], [26], [29] within the network. The relay nodes have special functionalities and are used in sensor networks to achieve various objectives, e.g., balanced data gathering, reduction of transmission range, connectivity and fault tolerance [8], [14], [29], [40]. These relay nodes can also be provisioned with higher energy [48], compared to sensor nodes, and can be used as cluster-heads in hierarchical sensor networks [15], [16], [25], [26].

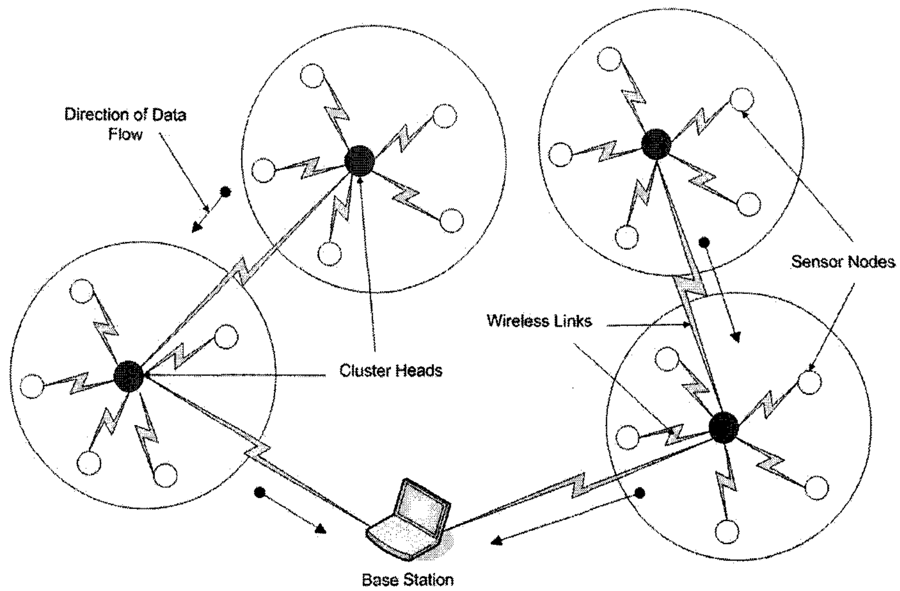


Figure 1.2: An example of hierarchical sensor network

In the *single-hop data transmission model* (also called the *direct transmission energy model* (DTEM)) [21], [20] the cluster heads send data directly to the base station. In the *multi-hop data transmission model* (MHDTM), [25], [26], [29], [48], relay nodes, acting as cluster heads, form a network among themselves to send data to the base station. In this case, the relay nodes not only transmit data gathered from the sensor nodes in their respective clusters but also forward data from other relay nodes towards the base station, as shown in Fig. 1.2.

1.2 Motivation

In a two-tiered sensor network, where the cluster heads use Multi-Hop Data Transmission Model (MHDTM), to transmit data to the base station, two important factors need to be considered for the relay nodes -

i) the placement strategy and

ii) the routing strategy.

The placement strategy attempts to find the minimal set of relay nodes, required within the network, such that each sensor node can communicate with at least one relay node and the relay node network is connected. The placement strategy is also responsible for determining the positions of each relay node, in this set. It has been shown in [46] that the problem of finding an optimal placement of relay nodes in sensor networks is NP-hard - even finding approximate solutions is NP-hard in some cases.

In MHDTM, the failure of a single relay node usually results in data loss from its own cluster and may prevent information flow of other relay nodes, which are using the failed node for forwarding data towards the base station. Therefore, it is important to have a placement strategy with some redundancy, so that, for a single relay node failure, data from all other relay nodes will still be able to reach the base station successfully.

Although the relay nodes can be provisioned with higher power, they are also battery operated and hence, power constrained. The goal of the routing strategy is to find a suitable data gathering schedule such that the lifetime of the network is maximized. Total depletion of the power of a relay node, specially in a hierarchical architecture, can impact the functionality of the network more severely [40] than the depletion of the battery of a simple sensor node. This is because, when the battery of a relay node is totally depleted (hence, *"die"*), the sensor nodes which are communicating with this relay node will no longer be able to send their data to the base station and an entire region within the network becomes inoperative. The death of a relay node may also put additional load on the surviving relay nodes, causing faster depletion of the batteries of other relay nodes. Therefore, maximizing the lifetime of a sensor network is directly related to maximizing the lifetime of the network of relay nodes. The lifetime of a network based on the MHDTM can vary considerably with the actual routing scheme used [25], [26], [29], [40].

1.3 Solution Outline and Contribution

In this thesis, we consider both the placement of relay nodes and some optimal routing strategies, in two-tiered sensor networks. We assume that the relay nodes are used as cluster heads and individual sensor nodes, in a cluster, communicate directly with the corresponding relay node. The relay nodes then use a multi-hop routing scheme to transmit data to the base station.

First, we propose a simple, efficient and scalable strategy for the placement of relay nodes in a specified sensing area so that the connectivity of the relay node network is ensured. Our approach divides the sensing region into imaginary cells, and creates an initial distribution of relay nodes at predetermined locations on the cell boundaries. Once the initial locations are determined, a simple heuristic is applied to remove any redundant relay node(s) form the initially set. Unlike existing placement schemes, our approach does not require complex computations [48]. Furthermore, we provide a theoretical upper bound on the worst case performance of our placement strategy. We also prove that our placement strategy guarantees that the resulting relay node network is at least 2-connected, and is capable of handling single faults in the relay node network.

After determining the positions of the relay nodes, we propose two Integer Linear Program (ILP) formulations that determine an optimal routing scheme to maximize the lifetime of the relay node network. Most existing formulations [14], [26], [29], [40] for maximizing the lifetime of the networks adopt the *flow-splitting* model where the flow of outgoing data is divided into a number of sub-flows and sent to different destination nodes. A more practical scheme is to allow each relay node to receive from any number of nodes but transmit to only one other relay node (or base station). This is the *non-flow-splitting* model [25] and this approach

a. simplifies the forwarding task and the use of directional antennas,

b. requires a minimum amount of packet-level power control,

c. relieves relay nodes from the burden of carrying multiple transmitters, and

d. frees the nodes from performing complicated routing functions.

The ILP formulations presented in this thesis find an optimal routing strategy for relay nodes, *without any flow splitting.* We have compared our approach with two widely used routing schemes - the direct transmission energy model (DTEM) and the minimum transmission energy model (MTEM) [20], [21]. The results indicate that our formulation outperforms both of these approaches. Finally, we have shown that, recomputing the routing strategy after predetermined intervals, results in additional lifetime improvements, compared to the situation where the routing strategy is fixed.

1.4 Thesis Organization

We provide a brief review of relevant background material in Chapter 2. Chapter 3 describes our placement strategy along with calculation of the performance bounds. Two ILP formulations for determining the routing strategy are presented in Chapter 4. The experimental results are discussed and analyzed in Chapter 5. Finally, we conclude and provide some future directions in Chapter 6.

Chapter 2

REVIEW OF LITERATURE

2.1 Sensor Nodes and Sensor Networks

A sensor node is a tiny computer, powered by lightweight batteries, and includes sensing device(s) to measure some physical phenomenon (e.g. temperature, humidity, temperature, illumination, pressure, movement-detection), an analog-to-digital converter (ADC) to covert the output of the sensing device to digital form and radio transiver, as shown in Fig. 2.1 (simplified from [3]). The nodes in a sensor network are deployed inside or very close to the phenomenon being observed (e.g., the temperature in the ocean bed) so that the sensing task can be carried out effectively. The placement of sensor nodes in a network can be pre-determined or random [3], as mentioned in Section 1.1.

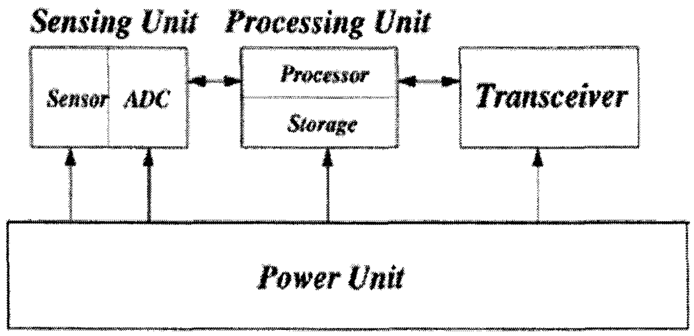


Figure 2.1: The components of a sensor node (simplified from [3], p. 399).

2.1.1 Sensor Energy Model

The power consumption in wireless communication is the most dominant factor in a sensor network. The first-order radio model [21] for energy dissipation in wireless communication is shown in Fig. 2.2 (redrawn from [21]). Energy is dissipated at a rate of E_{elec}/bit for both transmitting and receiving of data to run the transmitter and receiver circuitry. In addition, to transmit each bit to a destination at a unit distance, the amplifier of the transmitter in the source node dissipates ϵ_{amp} amount of energy. Typical values for these factors are $E_{elec} = 50nJ/bit$ and $\epsilon_{amp} = 100pJ/bit/m^2$ [21]. The energy loss due to channel transmission at a distance d is taken as d^m, where m is the path loss exponent, $2 \leq m \leq 4$, for free space and for short to medium-range radio communication [40]. Therefore, energy dissipated to receive (transmit over a distance d) k bits is given by, $E_{R_x}(k) = E_{elec} * k$ ($E_{T_x}(k,d) = E_{elec} * k + \epsilon_{amp} * k * d^m$).

2.1.2 Lifetime of Sensor Networks

The *lifetime* of a sensor network is defined as the time interval from the inception of the operation of the network, to the time when the power supplies of a number

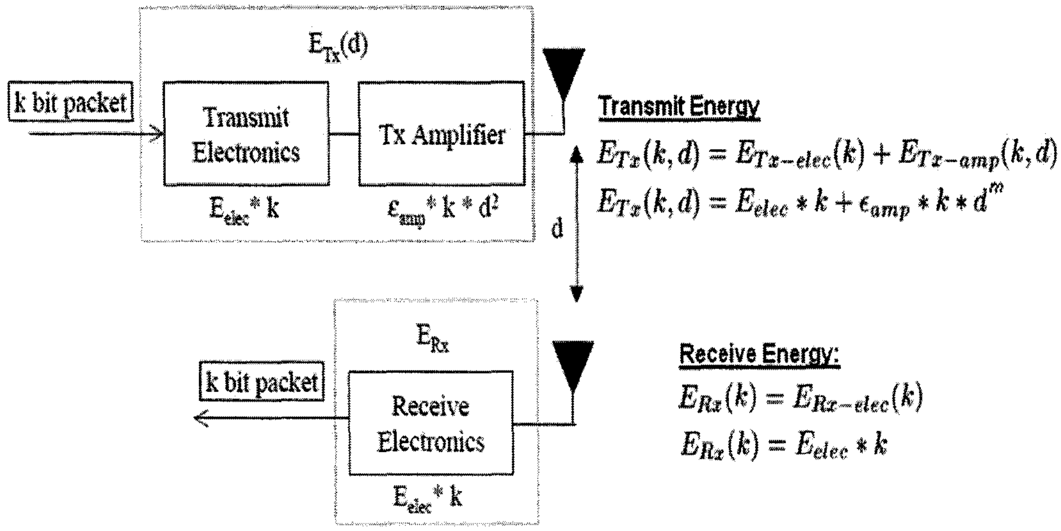


Figure 2.2: First order radio model ([21], p. 3006).

of *critical nodes* are depleted to such an extent that it results in a *routing hole* [42] within the network, a disconnected network, or a network with insufficient coverage. In sensor networks based on the flat architecture, the lifetime may be taken as the time when first node dies[1], or the last node dies or, more generally, a certain percent of nodes die.

In sensor networks based on the hierarchical architecture, the lifetime of the sensor nodes and that of the cluster heads need to be considered separately, as they have different impacts on the functional ability of the network. For example, if a sensor node dies, then the network suffers from the lack of sensing by this single node, which may only have a limited impact due to the inherent data redundancy in sensor networks. But if a cluster head dies, all the underlying sensor nodes of that cluster head become inaccessible from the other part of the network, a potentially

[1]when the power of a node is sufficiently depleted to affect its performance, the node is colloquially referred to as "dead" [21]

more important consideration.

In [40], three different ways to measure the lifetime of a hierarchical sensor network have been proposed as follows:

i) *N-of-N* lifetime, where the mission fails if any cluster-head node dies,

ii) *K-of-N* lifetime, where the mission survives if a minimum of K cluster-head nodes are alive and

iii) *m-in-K-of-N* lifetime, where the mission survives if all m pre-specified and overall a minimum of K cluster-head nodes are alive.

More information on the lifetime, including some upper bounds on the lifetime can be found in [42], [51]. An analysis of the energy consumption and the lifetime of heterogeneous sensor networks can be found in [13].

2.2 Relay Nodes in Sensor Networks

A number of approaches have been proposed to optimally balance the energy dissipation among all nodes in a sensor network [16], [17], [21], [29], [31], [33], [39]. One of these approaches is to use a special type of node in sensor networks, called *relay node*, whose job is only to relay the data generated by other sensor nodes, without sensing the environment. Relay nodes, are typically battery-operated devices with wireless communication capabilities. Relay nodes can prolong the lifetime of sensor networks, allow sensor nodes that are far away to communicate with each other and allow fault tolerance. Fig. 2.3 shows the use of relay nodes in a sensor network (redrawn from

[45]). Fig. 2.3(a), shows a traditional sensor network without relay nodes. The same network, with some relay nodes added to it is shown in Fig. 2.3(b). The topology shown in the Fig. 2.3(b), has reduced the transmission distances of nodes such as y, w, p or x, giving a network with an increased lifetime.

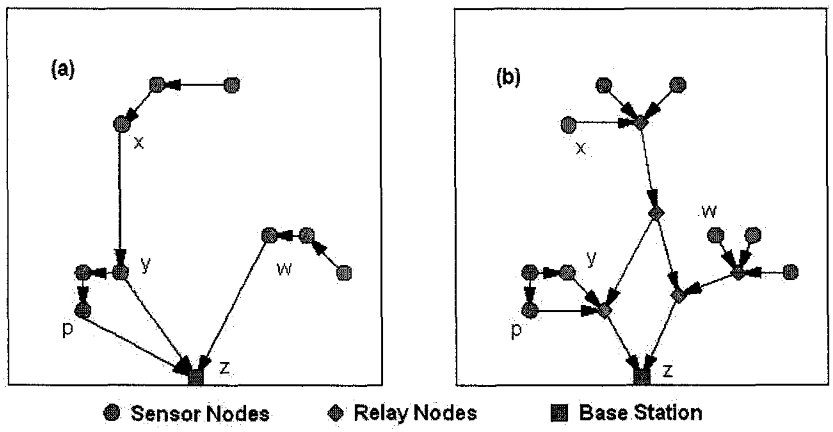


Figure 2.3: Use of relay nodes in flat sensor networks ([45], p. 1).

In the past few years, numerous papers [7], [8], [10], [11], [12], [14], [16], [26], [40], [41], [45] have studied the use of relay nodes with the following objectives:

1) Extending the lifetime of the network,

2) Energy-efficient data gathering,

3) Improving the connectivity,

4) Balanced data gathering,

5) Providing fault tolerance.

Relay nodes with different characteristics have been proposed to be used in flat architectures (Fig. 2.3) as well as in the hierarchical architectures (Fig. 2.4) [11],

[48].

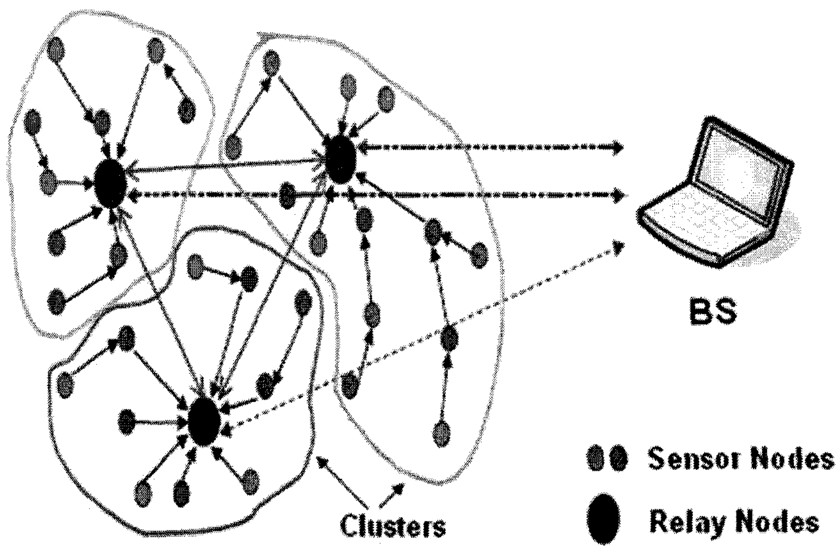


Figure 2.4: Use of relay nodes in hierarchical sensor networks ([16], p. 1848).

2.2.1 Relay Nodes in a Flat Architecture

The idea of deploying relay nodes in sensor networks, based on the flat architecture, was first introduced by Cheng et al. [8], in 2001, to investigate the effect on the total power consumption, if a small number of relay nodes is used in a network with pre-determined sensor locations.

In [11], Dasgupta et. al. have focused on maximizing the lifetime by studying the placement problem and the role-assignment problem in sensor networks of topology-aware nodes, where the sensor nodes may be mobile. Other work on mobile sensor networks appear in [27] and [43] which have high coverage, but does not address clearly how the lifetime of the network is affected by such placement. In [14], Falck et al. have attempted to achieve balanced data gathering against sufficient coverage, using relay nodes in a multi-hop sensor network and have solved the optimization problem using a Linear Programming (LP) formulation. This improves the work done in [30], [38] using non-linear solutions.

In [10], Coleri and Varaiya have studied ways to achieve a desired network lifetime, using minimum total energy in a sensor network that contains relay nodes. Two different formulations - one using a Linear Programming formulation and the other using a Non-Linear Programming formulation have been proposed.

2.2.2 Relay Nodes in Hierarchical Architectures

This problem was first considered in [16] and [40] to address the issue of load balancing in energy-constrained sensor networks, deployed uniformly in an inhospitable environmental condition. They have proposed an algorithm for clustering the sensor nodes around some higher-powered relay nodes (which they called *gateway nodes*), acting as cluster heads, to achieve the objective.

In [40], Pan et al. have attempted to maximize the topological network lifetime of sensor networks by arranging the base stations (BS) and by optimal inter-aggregation node (AN) relaying. They have proposed a two-tiered sensor network model where the sensor nodes lie in the lower tier and the Application Nodes (AN) as well as the Base Stations (BS) lie in the upper tier. In this model, the sensor nodes in the networks form clusters and send their readings directly to the respective AN. Their algorithms are based on Computational Geometry that finds the optimal locations of the BS's under the three definitions of lifetime discussed above. They have also established theoretical upper and lower bounds on the maximal topological lifetime of sensor networks.

In [26], Hou et al. have focused on prolonging the lifetime of sensor networks with energy provisioning to the existing nodes and deploying relay nodes within the two-tiered, cluster-based wireless sensor networks model that contain Aggregation and Forwarding Nodes (AFN) and relay nodes (RN's). They have focused on mainly two aspects,

a. provisioning additional energy to the existing nodes, and

b. the deployment of AFN's, and RN's to prolong the lifetime of the network by mitigating the geometric deficiencies of the network with the use of these nodes.

2.3 Node Placement and Coverage in Sensor Networks

Most of the research, discussed so far, has focused on the performance improvement of sensor networks with an assumption that the networks have been already deployed. The research on the node placement and the coverage problems in sensor networks, on the other hand, has focused on the efficient deployment of sensor nodes within the networking field. As mentioned earlier, each sensor node in a sensor network monitors a small area surrounding the node. The complete view of the area where the sensor network is deployed, for the attribute(s) being monitored, is constructed by putting together the data received from a large number of sensor nodes that are dispersed throughout the sensing field. Obviously, no data can be obtained from a region if it it not covered by at least one sensor node or if sensor node(s) covering the region get disconnected. This means that the placement of sensor nodes must take into consideration the *coverage* of sensor networks. Coverage is an important area in sensor networks and it has been studied in many papers, including [5], [49], [50] and [52].

A network model, which can be either *deterministic* or *probabilistic* [45], specifies the area covered by each sensor node in the network. In a deterministic model, the area covered by each sensor node is predetermined and the coverage is measured by the area that is covered by at least one sensor. If the model is probabilistic, then it specifies the probability that a phenomenon will be detected at a given location [35], [45].

In a two-tiered sensor network where relay nodes are used as cluster heads and equal-capability sensor nodes are randomly deployed, the placement of relay nodes should ensure that all the sensor nodes in the network are covered by the set of relay nodes, i.e., each sensor node should be able to communicate with at least one relay node. A sensor node can be considered as *covered* if at least one relay node lies in the area around the sensor node within its transmission range. As the radio transmission is inherently broadcast, in free space, the area covered by each relay node can be seen as a unit circle. Here the radius of the circle is the transmission range of the sensor node. Therefore, the problem of finding the minimum number of relay node to cover a sensor network may be reduced to the problem of finding the locations and the number of circles that can be used to cover the monitored area. For example, Fig. 2.5(a) (redrawn from [45]) illustrates the idea of covering a square area with 6 equal circles. The problem of covering a square with equal circles has been studied in [36] and [37].

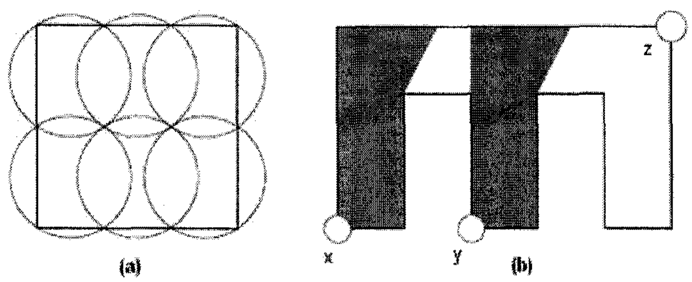


Figure 2.5: (a) Covering an area with 6 equal circle. (b) An sample solution of the art-gallery problem ([45], p. 17).

Another approach that has been explored for finding the locations of nodes, in a type of sensor network that has obstacles in the monitored area, is related to the well-known *art-gallery* problem [45]. Given the plan of the interior of an art-gallery, the art-gallery problem attempts to find the minimum number and the placement of guards for completely monitoring the gallery. An illustration of the art-gallery problem is shown in Fig. 2.5(b) (redrawn from [45]). The figure shows a solution for the art gallery problem, where the guards may be placed at locations x, y and z to completely cover the interior of the entire gallery. The placement problem of sensor nodes has also been addressed in [7], [41] and [48]. The complexity issues for the relay node placement problem have been studied in [45] and [46].

In [7], the problem of sensor node placement and the data transmission pattern (in terms of the network lifetime and the total power consumption) in sensor networks has been solved using a nonlinear program. Considering a region with a specified number of sensor/aggregation nodes and a certain coverage requirement, the paper shows how to optimize the network lifetime and the total cost.

In [41], Patel et al. have addressed the optimum placement problem of the sensor nodes, the relay nodes and the base station in a sensor network either to minimize the number of deployed sensor nodes, the total cost, the energy consumption, or to maximize the energy utilization or the lifetime of the network. An integer linear program formulation has been proposed for the placement problems for both reliable and unreliable/probabilistic detection models. The placement of these nodes is such that

i) each point of interest in the sensor field is covered by a subset of sensors of desired cardinality,

ii) the resulting sensor network is connected and

iii) the sensor network has sufficient bandwidth.

In [48], Tang, Hao and Sen have focused on the placement of relay nodes with guaranteed coverage and connectivity.

In [45], Suomela has studied the complexity of relay-node-placement problem in sensor networks and, for different optimization problems, proposed some algorithms to find k-optimal solutions of the balanced data gathering problem, based on the method proposed in [14]. The objective was to optimize relay node placement in two different senses,

i) *maximizing the utility*, given a fixed number of relays and

ii) *minimizing the number of relays*, given a target value of the utility function.

The k-approximate version of both problems turn out to be NP-hard.

In [46], Suomela has focused on finding the computational complexity of the relay-node-placement problem for balanced data gathering, where the utility function is a weighted sum of the minimum and average amounts of data collected from each sensor node. All of these problem classes are NP-hard, and, in some cases even finding approximate solutions is NP-hard [46].

2.4 Routing in Sensor Networks

Routing in sensor networks is a challenging task [4] due to the following problems:

- The number of nodes deployed in a sensor network may be very large.
- Sensor nodes are constrained by energy, processing, and storage capabilities.
- Once deployed, most of the sensor nodes are usually stationary, but some nodes may be allowed to move around, depending upon the requirements of the application.
- The requirements for the design of sensor networks may change with application.
- Data collection, in a sensor network, is usually location based, so that position awareness of sensor nodes is important.
- A large number of sensor nodes is usually densely deployed in a sensor network. As all sensor nodes usually monitor a common phenomena, it is highly probable that the data is redundant. Appropriate aggregation techniques are needed to take care of this redundancy so that the available bandwidth is utilized efficiently.

In addition to the above mentioned routing challenges, the resource constraints of the sensor nodes, especially the energy constraints, the unpredictable changes of the nodes and link status (e.g. due to node failure or mobility) and the corresponding topology changes make routing in sensor network a nontrivial task. In the past few years, many algorithms have been proposed to address these challenges for routing in sensor networks. The routing strategies proposed in most of the literature mainly concentrated on minimizing the energy consumption of the sensor nodes so that the lifetime of the network is maximized. Along with employing various standard tactics for routing in wireless networks, different papers have proposed techniques, such as clustering of sensor nodes, data-centric approach, load balancing, energy-efficient data gathering, data aggregation and in-network processing, role assignment nodes [4].

Routing protocols can be classified in a number of ways. One scheme based on the network structure has been described in Section 1.1. In another scheme, Al-Karaki and Kamal [4] have classified the routing protocols, based on the network structure ([6], [22], [28]) and the protocol operation ([19], [28]). These classifications are shown in Fig. 2.6 (modified from [4]).

One more classification is based on how a source finds a route to the destination which are characterized as *proactive, reactive,* and *hybrid.* Proactive protocols compute all routes beforehand, i.e. before routes are actually needed, reactive protocols compute routes on demand while hybrid protocols use a combination of both proactive and reactive schemes [2], [4].

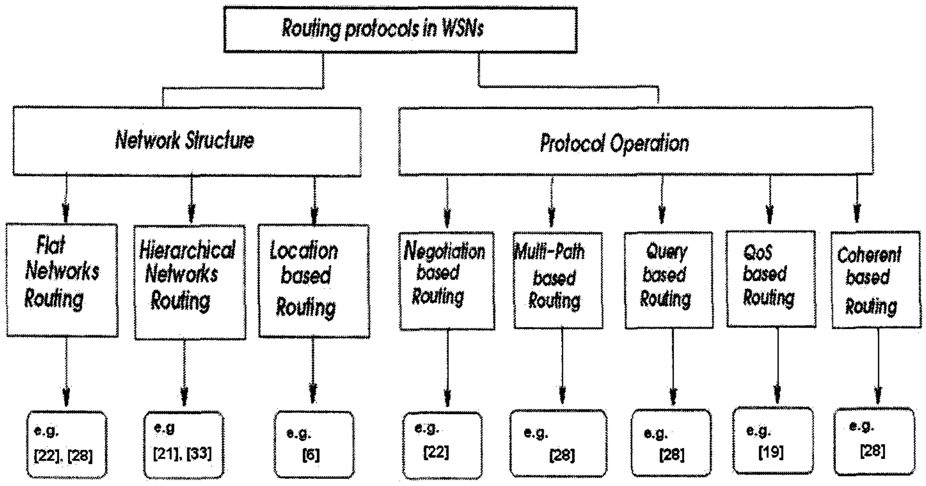


Figure 2.6: A taxonomy of routing protocols ([4], p. 7).

In a multi-hop *flat* network architecture, each sensor node is typically assigned the same functionality and plays the same role, i.e., each node does sensing and collaborate together to perform the networking task. These protocols use a data-centric approach for routing. In this type of routing, the base station (also known as the sink) sends queries to a certain region of the network. Upon reception of the queries, the sensor nodes located in the selected regions send the data being queried, towards the base station, each sensor node using a multi-hop path (Fig. 2.7). Some examples of flat routing protocols include Sensor Protocols for Information via Negotiation (SPIN) [22], Directed Diffusion [28] and many other protocols that use similar concepts [4].

It is well known that the hierarchical techniques offer advantages related to scalability and efficient communication [4]. These architectural advantages have been exploited to perform energy-efficient routing in sensor networks. Each sensor node in such a network belongs to one distinct cluster and sends data to only its own cluster

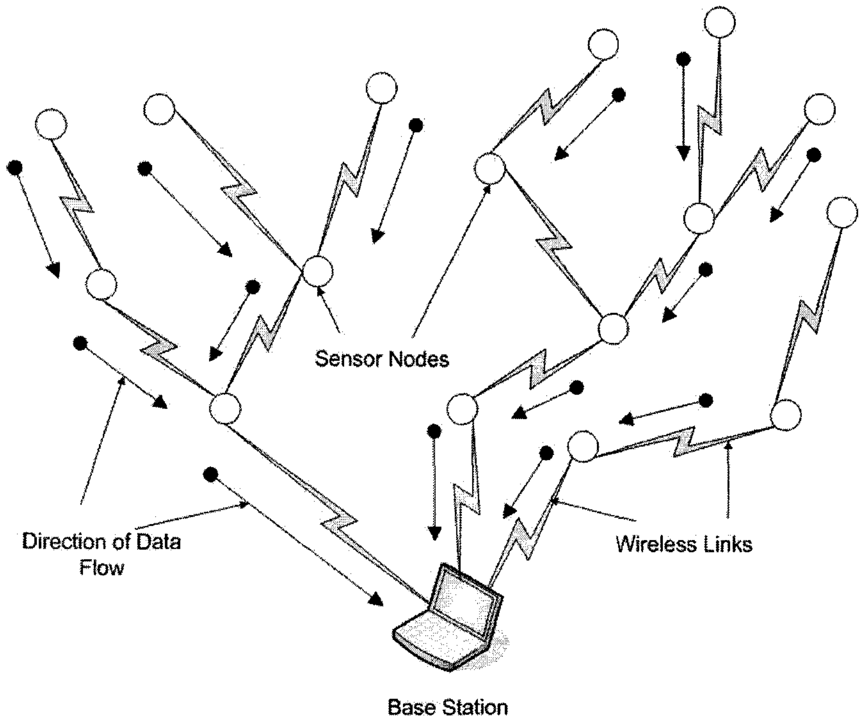


Figure 2.7: An example of routing in a flat architecture.

head. Cluster heads collect data from all the sensor nodes in its own cluster, process the data and send the result towards the base station. The cluster heads may also use multi-hop path to forward data towards the base station, where each cluster head also acts as a router for the data forwarded to it by the neighboring cluster head nodes (Fig. 2.8).

One of the advantages of the hierarchical architecture is that, higher-energy provisioned nodes can be used as cluster heads, as these nodes are expected to perform data processing, take part in routing and transmit data to the base station (possibly, using multi-hop paths), which may be lies at a distant location. Sensor nodes, on the other hand, can be low-energy nodes, as these nodes perform only the sensing in the proximity of the target and transmitting the sensed data to the immediate cluster head

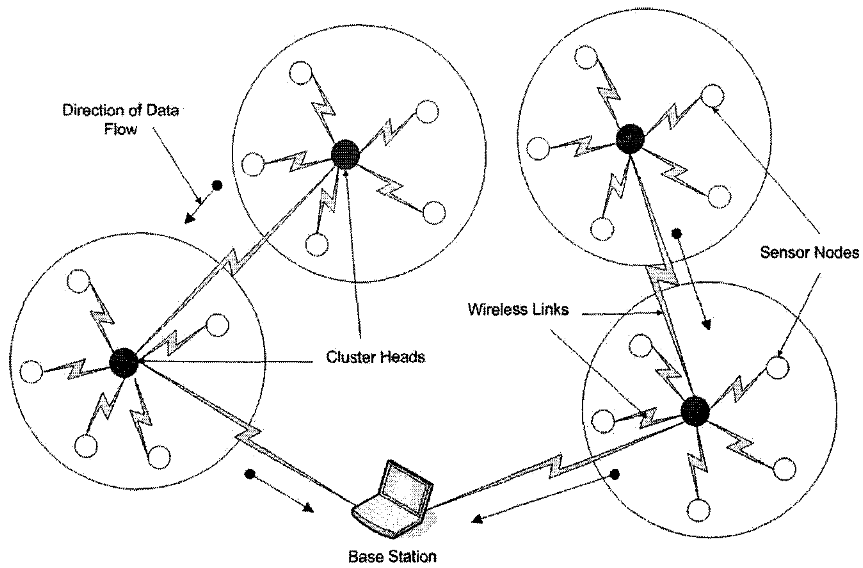


Figure 2.8: An example of hierarchical routing in sensor networks

only (which usually lies at a short distance) and may not participate in the routing. Even if nodes with the same capacity are used as cluster heads, the role of cluster heads can be rotated among the sensor nodes and the benefit of hierarchical architecture can be exploited [21]. Clustering in sensor networks contributes to the improvement of overall system performance including scalability, network lifetime, and efficient energy utilization [4]. Hierarchical routing can lower the energy consumption for intra-cluster communication and lower the energy consumption for inter-cluster communication by data aggregation and fusion [4], [16], [17], [21], [31], [33], [39].

Most of the proposed hierarchical routing protocols use two-layer routing. For communication from a sensor node to the base station, the first stage is to select the cluster-head. The next stage is to find a proper multi-hop route from the cluster head to the base station. Examples of hierarchical-routing protocols include the following:

- Low-Energy Adaptive Clustering Hierarchy (LEACH) [21], using randomization to select cluster heads,

- Threshold Sensitive Energy Efficient Sensor Network (TEEN) [33], the *reactive* approach in LEACH to further enhance the energy efficiency,
- Hybrid Energy-Efficient Distributed clustering (HEED) [39], using a distributed approach for the selection of cluster heads.

2.5 Fault Tolerance in Sensor Networks

The objective of fault tolerance in sensor networks is to ensure that the network remains functional even in the event of node and/or link failures. In general, sensor nodes are prone to failures due to reasons such as running out of battery power, physical damages and malicious attacks. Also, there can be infrequent link failures, which may occur due to the environmental interference or node mobility.

To withstand a node and/or link failure in a network, a traditional fault-tolerant approach is to establish node/link disjoint paths between all source, destination pairs. This approach ensures connectivity in the networks in the case of a failure, i.e. if some links and/or nodes fail, the remaining network still remains connected. A fault tolerant network should generally be at least 2-connected, but can be k-connected [48], where $k \geq 2$, depending upon the criticality of the mission of the network. Finding disjoint paths is an important research area in networking. For example, computing the minimum total-cost disjoint paths has been studied for general networks [47], as well as for wireless networks [44].

In a flat architecture, sensor nodes themselves are responsible for routing the data. Therefore, fault-tolerant schemes in this architecture need to take into consideration all the sensor nodes within the network in the same way. But in hierarchical architectures, sensor nodes (in *lower tier*) and cluster heads (in *upper tier*) must be treated differently. In the lower tier, each sensor node belongs to only one cluster and sends data only to its own cluster head in this architecture. Therefore, fault tolerance for *sensor nodes* attempts to ensure that, in case a cluster heads fails, the underlying sensor nodes are still able to communicate with some other cluster head, so that the data generated by these nodes is not lost. In the upper tier, since the cluster heads may also form networks among themselves and use multi-hop routing to send data to the base station, node/link disjoint paths are needed to be established between each pair of source-destination cluster heads so that the functionality of the network is not disrupted in case of single cluster head failures.

In [15], Gupta and Younis have addressed the issue of fault tolerance in two-tiered cluster-based sensor networks and proposed a mechanism for recovering sensor nodes that belongs to a cluster whose cluster head (called *gateway* nodes in [15]) has failed. Higher-powered gateway nodes act as cluster heads in the upper tier. Each sensor node lies in the lower tier can communicate with only one gateway node, which is the cluster head of the cluster containing the sensor node. Failure in gateway nodes are more severe in such a system since the underlying sensors covered by a failed gateway node will become inaccessible, although they are still fully functional. A mechanism to access the sensor nodes in the cluster corresponding to a failed gateway node, without a full-scale re-clustering and any redundant gateway nodes, have been proposed in [15].

In [18], Hao, Tang and Xue have focused on the problem of relay-node placement in two-tired, cluster-based, wireless sensor networks so that the network become fault-tolerant. In the upper tier, relay nodes are used as cluster heads and are responsible for collecting data from the sensor nodes of their respective cluster, aggregating received data, forming connected topology and transmitting the data towards the sink using multi-hop routing. The authors have formulated a fault-tolerant scheme for finding the minimum number of relay nodes, so that each sensor node is connected to at least two relay nodes and the relay node network itself is 2-connected. They also have proposed a polynomial-time approximation algorithm to solve the problem.

In [32], Liu, Wan, and Jia have considered a two-tiered sensor network model where relay nodes are deployed in the upper tier and are used to forward data packets from the sensor nodes towards the sink. They have attempted to solve the problem of finding the optimal number relay nodes as well as their placements for a fault tolerant network and proposed a number of approximation algorithms.

Fault tolerance in a two-tiered sensor networks is also studied in [48] where the entire sensing region is divided into cells of size $2r \times 2r$, where r is the communication range of each sensor node. They have focused on the placement of relay nodes with guaranteed coverage and connectivity and attempted to find the placements of relay nodes based on some initial set of probable locations.

Chapter 3

PLACEMENT OF RELAY NODES IN SENSOR NETWORK

3.1 The Placement Problem of Relay Nodes

The role of relay nodes as cluster heads in a two-tiered network, has been reviewed in Chapter 2. In such a network, it is important to place the relay nodes so that all the sensor nodes in the network can communicate with at least one relay node, i.e., each sensor nodes must be *covered* by at least one relay node. A sensor node must communicate with at least one relay node so that the data generated by the node may be collected by the network. Generally, a relay node can communicate with many sensor nodes. The *relay node placement problem* is that of finding the location of relay nodes in a sensor network, so that all the sensor nodes are covered using a minimum number of relay nodes. In a network where there is no obstacle, a sensor node can transmit in any direction within its transmission range. Therefore, the placement of relay nodes in such a network is the problem of covering the area corresponding to the network using a minimum number of discs,where the radius of each disc is the transmission range of a sensor node.

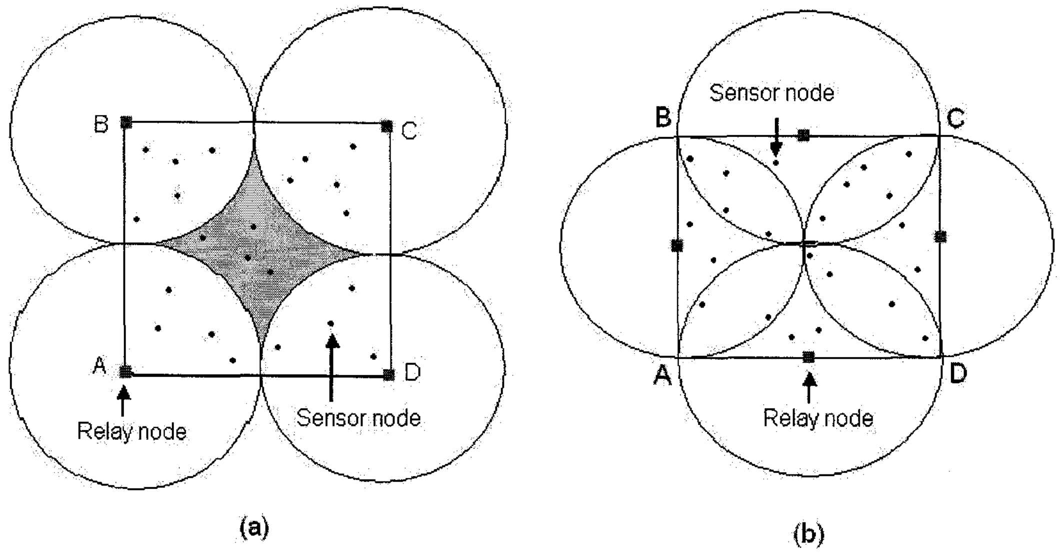


Figure 3.1: An example depicting the importance of the placement of relay node in sensor networks. (a) A placement of four relay nodes that does not cover all sensor nodes in the network. (b) A placement of four relay nodes that covers all the sensor nodes of the same network.

Fig. 3.1(a) shows, an arbitrary network bounded by the rectangle $ABCD$ with relay nodes placed at points A, B, C and D. The circles with centers A, B, C and D, having radius r, are also shown. Each circle represents the area covered by the corresponding relay node. All the sensor nodes in the network are not covered so that the sensor nodes lying within the shaded area will not be able to communicate with any of the relay node. Therefore, this placement of relay nodes is inadequate and at least one more relay node is required (e.g., at the center of the network area) to cover the entire network. But the same network can be covered by four relay nodes with an appropriate placement within the network, as shown in Fig. 3.1(b). Figures 3.1 show the importance of finding the appropriate placement of relay nodes. In this chapter, the topic of determining a "reasonably small" number of relay nodes to cover all the sensor nodes of the network is covered.

In a two-tiered sensor network where relay nodes are used as cluster heads and the relay nodes use the *multi-hop data transmission model* (MHDTM), [25], [26], [29], [48], to forward data towards the base station, the placement of the relay nodes must also make sure that the relay node network is connected. In this model, the relay nodes, acting as the cluster heads, not only transmit data gathered from the sensor nodes in their respective clusters but also forward data from other relay nodes towards the base station, as shown in Fig. 3.2. Hence, the relay nodes form a network among themselves and use multi-hop paths for routing data to the base station.

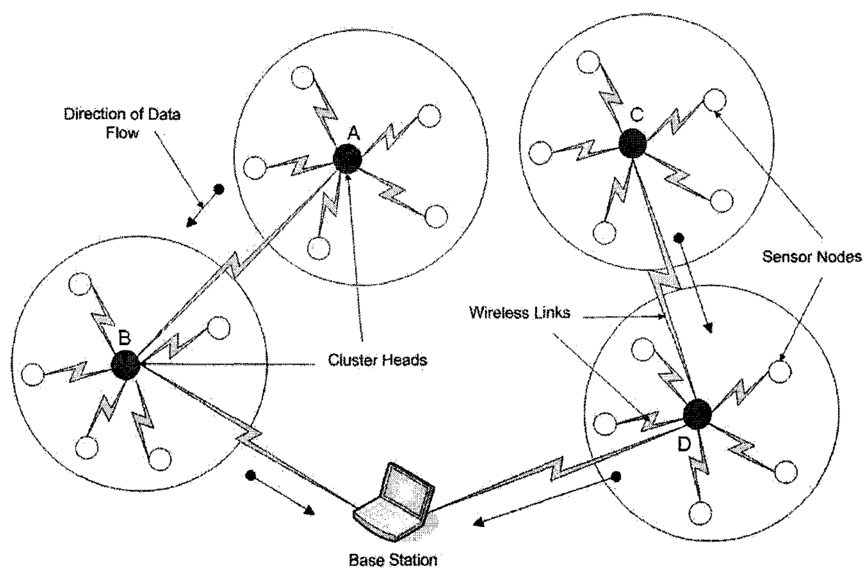


Figure 3.2: An example of MHDTM in a two-tired sensor networks

Moreover, in such a model, if the connectivity of the network is 1, a fault in a single relay node may severely impair the functionality of the network. This is because, in MHDTM, the failure of a single relay node not only results in data loss from its own cluster, but also may prevent information flow of other relay nodes, which are using the failed node for forwarding data towards the base station. It is important to have a placement strategy with some redundancy, so that, in the event of the failure of a single relay node, data from all other relay nodes will still be able to reach the base station successfully.

It has been shown in [46] that the problem of finding an optimal placement of relay nodes is NP-hard - even finding approximate solutions is NP-hard in some cases. In the following section, we present an efficient and scalable strategy for the placement of relay nodes in a specified sensing area, to achieve the desired coverage and, at the same time, be able to handle the failure of a single relay node. Our approach requires significantly less computation compared to existing schemes [48]. We also provide a theoretical upper bound on the worst case performance of our placement strategy, with respect to the optimal solution, and prove that our placement strategy guarantees that the resulting relay node network is at least 2-connected.

3.2 A Placement Strategy for Survivable Relay Node Network Design

We consider a two-tier network consisting of sensor nodes with communication range r and relay nodes with communication range R, where $R \geq 4r$. Following the approach used in [48], we start by dividing the entire sensing region into cells of size $2r \times 2r$, where r is the communication range of each sensor node. A sensor node s is *covered* by a relay node u, if the distance from s to u is less than or equal to r (i.e. s can transmit to u directly). Our objective is to find a placement of relay nodes such that each sensor node in the sensing area is covered by at least one relay node, and the number of relay nodes is minimized. We will use $\mathcal{S}$ to denote the set of all sensor nodes and use $\mathcal{S}^u$ to denote the set of all sensor nodes covered by relay node u. The steps of our placement strategy, which we call the *fixed placement (fp)* strategy, are given below.

Step 1: Divide the entire area into an imaginary grid with k_1 rows, numbered $1, 2, \ldots, k_1$, with each row having k_2 cells, numbered $1, 2, \ldots, k_2$, where each cell has size $2r \times 2r$.

Step 2: Put relay nodes on the center of the top boundary and the center of the left boundary of each cell in the imaginary grid.

Step 3: For the cells in row (column) number k_1 (k_2), put relay nodes on the center of the bottom (right) boundaries.

Step 4: Let $\mathcal{R}$ be the set of relay nodes found in steps 2 and 3. Using some heuristic for minimum set covering (one possible heuristic is described in section 3.5), find the set of relay nodes $\mathcal{R}_{min}$ with the smallest number of elements such that $\bigcup_{u \in \mathcal{R}_{min}} \mathcal{S}^u = \mathcal{S}$.

Our work was motivated by [48] and for comparison, we give some details of the approach given in [48].

Definition

The *P-positions* for a pair of sensor nodes at locations $\boldsymbol{x}$ and $\boldsymbol{y}$ are the point(s) of intersection (if any) of two circles of radius r with centers at $\boldsymbol{x}$ and $\boldsymbol{y}$ in the same cell.

In [48] the process starts by dividing the entire region into imaginary cells of size $2r \times 2r$ as described above. An optimal placement of relay nodes for each cell is computed from $\wp$, the set of P-positions for all pairs of sensor nodes within the cell, by checking *all* subsets of $\wp$ of size four or less.

The performance ratio (p_a) of a placement algorithm a is defined as the ratio of the size of the solution provided by the algorithm a, divided by the size of the optimal solution. By applying the shifting lemma [23], the authors in [48] have shown that for cells of side length $2r.l$, where l is an integer, if p_a is the performance ratio of the relay node placement algorithm within each cell, and p_{s_a} is the performance ratio of the algorithm for the entire area, obtained by combining the solutions for each cell, then $p_{s_a} \leq p_a(1+\frac{1}{l})^2$. In case $l = 1$, the performance bound of this strategy is given by

$$p_{s_a} \leq 4p_a \tag{3.1}$$

On their network model, the authors have proposed two schemes for the placement of relay nodes within the network. The first scheme focused on placing a minimum number of relay nodes within the network in such a way that each sensor node is connected with a minimum of one relay node and the relay nodes network itself is connected. They have formulated this optimization problem and named it as Connected Relay Node Single Cover (CRNSC) problem.

In the second scheme, the authors have addressed the issue of fault tolerance by enabling the network to survive the failure of single failure of a relay node. The scheme makes the network two-connected in both tiers such that each sensor node can communicate with at least two relay nodes and the network of the relay nodes are two connected. They have formulated this optimization problem and named it as 2-Connected Relay Node Double Cover (2CRNDC) problem. For the solution, they have proposed two approximation algorithms for each problem. Using the concept discussed in [24] and [37], they have proved that, in terms on the number of relay nodes used, the performance for CRNSC problem is bounded by 8 and 4.5 from the optimal solution (for proposed two approximate solutions respectively). And for the 2CRNDC problem, the bounds are 6 and 4.5 (for the proposed two approximate solutions respectively).

In [48] it is necessary to compute the set of P-positions $\wp$, for all pairs of sensor nodes within the cell and check *all* subsets of $\wp$ of size four or less. For a network with hundreds, or thousands of sensor nodes, this can require significant computational effort. It is important to note that our placement algorithm uses the same idea of dividing the sensing area into smaller cells, but requires much less computational effort.

3.3 Performance Bound for the Fixed Placement Strategy

In this section, we obtain a theoretical upper bound for the worst case performance of our strategy. We also show that, as the size of the sensing area increases (relative to the size of a single cell), our worst case performance bound approaches the same value as the more complex scheme proposed in [48].

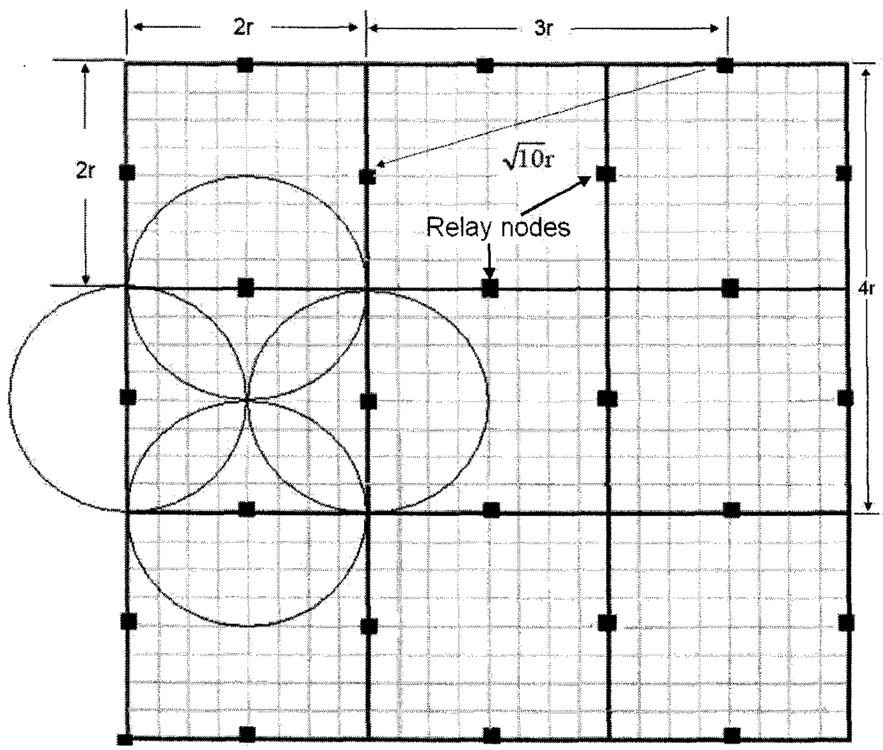


Figure 3.3: The placement of relay node in grids

We assume, in our analysis, that the number of sensor nodes is much higher than the number of relay nodes, and that the sensor nodes are densely distributed in the sensing area, so that there is at least one sensor node in each cell. A sensor area of size $2r.k \times 2r.k$ consists of k^2 cells of side length $2r$, and hence requires at most $2k^2 + 2k$ relay nodes, using our fixed placement strategy. Since, there is at least one sensor node in each cell, any placement algorithm that *optimally* places relay nodes in each cell (without considering the effect of neighboring cells) would require at least k^2 relay nodes. Therefore, the performance ratio of our fixed placement strategy, with respect to the optimal algorithm, for a single cell is given by $p_{fp} \leq \frac{2k^2+2k}{k^2} = 2 + \frac{2}{k}$. So, using Equation 3.1, the overall performance ratio of our scheme is $p_{s_{fp}} \leq 4(2+\frac{2}{k})$. For $k = 1$, the optimal solution for a cell is the same as the optimal solution for the entire area and $p_{s_{fp}} = 2 + \frac{2}{k} = 4$. The worst case scenario is for $k = 2$, in this case $p_{s_{fp}} \leq 12$. However, as k increases, $p_{s_{fp}}$ decreases and for large values of k, $p_{s_{fp}} \approx 8$. This is the same bound calculated for the algorithm in [48]. For the generalized case of a rectangular sensing area of size $2r.k_1 \times 2r.k_2$, the performance bound is given by $p_{fp} \leq \frac{2.k_1.k_2+k_1+k_2}{k_1.k_2}$.

The advantage of our approach is that it automatically guarantees that any individual sensor node is covered by at least one relay node, without requiring any complex computations. This is shown in Fig. 3.3, for a square sensing area of size $2r.k \times 2r.k$, for $k = 3$. We can see that, for any given sensor node s within a cell, there is at least one relay node within a distance r from s.

3.4 Connectivity of the Relay Node Network

In this section, we show that the relay node network generated by our placement scheme is at least *2-connected*. This means that even if a single relay node fails, the remaining nodes will still have a viable route to the base station. Only the local information from the region covered by the faulty relay node will be lost, and the rest of the network can continue to function. This is important for improving the survivability of the network.

Theorem 1: The *fp* strategy generates a 2-connected network of relay nodes.

Proof: First, we consider the case where none of the relay nodes, placed in steps 2 and 3 by the *fixed placement* strategy, are removed. In this case each relay node has at least one vertical neighbor and one horizontal neighbor at a distance $2r$ from itself. Since the communication range of a relay node is given by $R \geq 4r$, each node can communicate with at least two other nodes. Therefore, the network is at least 2-connected.

Next, we show that even if some relay nodes are removed by the algorithm in step 4, the relay node network still remains 2-connected. Since we have at least one sensor node inside each cell, and each sensor node is covered by at least one relay node, there must be a relay node on at least one of the boundary edges of each cell. Without loss of generality, we assume that, in a given cell i, there is a relay node located at the midpoint of the top boundary edge, and the relay nodes on all other boundaries of cell i have been removed. In this case, the distance to the farthest possible relay node on one of the boundary edges of a neighboring cell, in the horizontal direction, is $\sqrt{10}r$. Similarly, the distance to the farthest possible relay node on one of the boundary edges of a neighboring cell, in the vertical direction, is $4r$, as shown in Fig. 3.3. Since the communication range of a relay node is $R \geq 4r$, and each cell in the sensing area has at least one horizontal neighbor and one vertical neighbor, it follows that every relay node can communicate directly with at least two other relay nodes. Hence, the relay node network remains 2-connected, even if some relay nodes are removed in step 4 of our placement strategy.

3.5 A Heuristic to Minimize the Number of Relay Nodes

The performance bound presented in Section 3.3 for the *fixed placement* strategy gives the number of relay nodes required to cover a given area in the worst case scenario. The actual performance of the strategy is further improved by removing redundant relay nodes wherever possible (step 4, Section 3.2), using a heuristic. We have used a simple, greedy heuristic to perform this function, given in the algorithm *Minimum-Set-Cover*.

The heuristic first identifies the essential relay nodes (the first outer for loop). An *essential* relay node, $u \in \mathcal{R}$, is a relay node such that there exists a sensor node, $s \in \mathcal{S}$, in the network which can communicate only with u. Once an essential relay node, u, is identified, it is included in the set $\mathcal{R}_{min}$ as a required relay node. All sensor nodes, $s \in \mathcal{S}$, that can communicate with the selected relay node u are then assigned to the cluster of u and removed from the set of sensor node, $\mathcal{S}$ (the inner for loop of the first For loop). As each essential relay node is identified, all sensor nodes that can be included in the cluster corresponding to the essential relay node is removed from the set of sensor nodes $\mathcal{S}$. The algorithm then finds a relay nodes $u \in \mathcal{R}$ which covers the maximum number of nodes from the set of remaining sensor node, $\mathcal{S}$, (inside the while loop). This u is then added to the set $\mathcal{R}_{min}$ and sensor nodes in $\mathcal{S}$ that may communicate with u are assigned to the cluster for u (inner for loop within the while loop). The sensor nodes that may communicate with u are removed from $\mathcal{S}$ and the process continues until all sensor nodes are assigned into a cluster.

Algorithm 1 ***Minimum-Set-Cover***

Input: A set of relay nodes, $\mathcal{R}$ and a set of sensor nodes, $\mathcal{S}$
Output: $\mathcal{R}_{min}$, which is a minimal subset of $\mathcal{R}$ such that each sensor node $s \in \mathcal{S}$ is covered by at least one $u \in \mathcal{R}_{min}$ and the cluster for each $u \in \mathcal{R}_{min}$, $\mathcal{S}^u$, such that $\bigcup_{u \in \mathcal{R}_{min}} \mathcal{S}^u = \mathcal{S}$.

```
begin
    R_min ⟵ NULL
    for Each s ∈ S do
        Find a u ∈ R such that s is covered by only u.
        R_min ⟵ R_min ∪ {u}
        for Each s ∈ S do
            If s is covered by u, Then
            S^u ⟵ S^u ∪ {s}
            S ⟵ S − {s}
        end
    end
    while S ≠ NULL do
        Find a u ∈ R such that u covers maximum number of s ∈ S.
        R_min ⟵ R_min ∪ {u}
        for Each s ∈ S do
            If s is covered by u, Then
            S^u ⟵ S^u ∪ {s}
            S ⟵ S − {s}
        end
    end
end
return R_min, S^u, ∀u ∈ R_min.
```

3.5.1 An Example of *fixed placement* Strategy

We illustrate the idea of the fixed placement strategy in this section with the help of Fig. 3.4 and 3.5. As shown in Fig. 3.4, let the network area be ABCD and let the sensor nodes be randomly distributed within the network area. At first, we divide the network area into imaginary cells with each side equal to twice the transmission range of the sensor nodes (step 1, Section 3.2). The figure shows that the area has to be divided into nine cells. Once the cells are formed, we find the initial positions of the relay nodes, following step 2 and 3 of the *fp* strategy (Section 3.2). This initial positions of relay nodes are shown in Fig 3.4. From the figure, we can see that 24 relay nodes are used to cover the entire network.

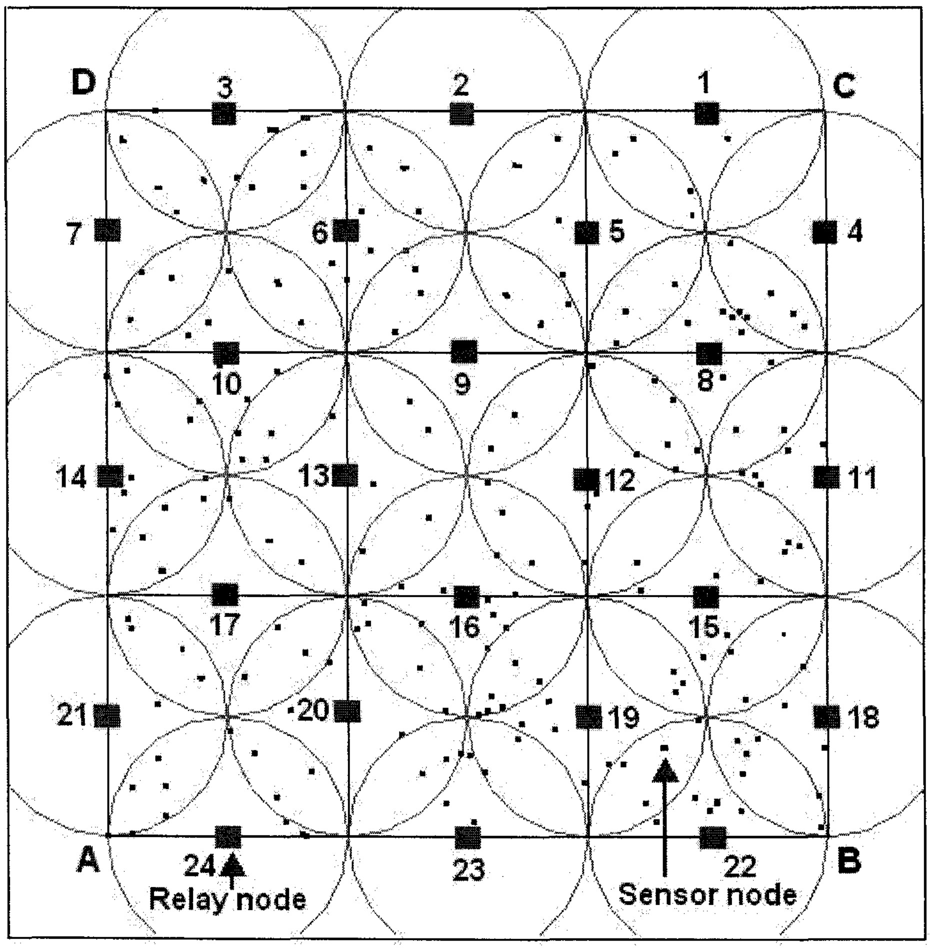


Figure 3.4: An example of the placement of relay nodes using *fp* strategy.

The number of relay nodes may be reduced using the heuristics given in Section 3.5, which is step 4 of the *fp* strategy (Section 3.2). The result of the application of the heuristic on the initial placement is shown in Fig. 3.5. As shown in the figure, the heuristics has removed five relay nodes, relay nodes 2, 5, 7, 9 and 17, from the initial set of twenty four relay nodes. It may also be noted that although relay node 5 can cover all, except one sensor node that are covered by relay node 1. Therefore, the heuristic keeps the relay node 1, otherwise, the data generated by that single sensor node cannot be accessed.

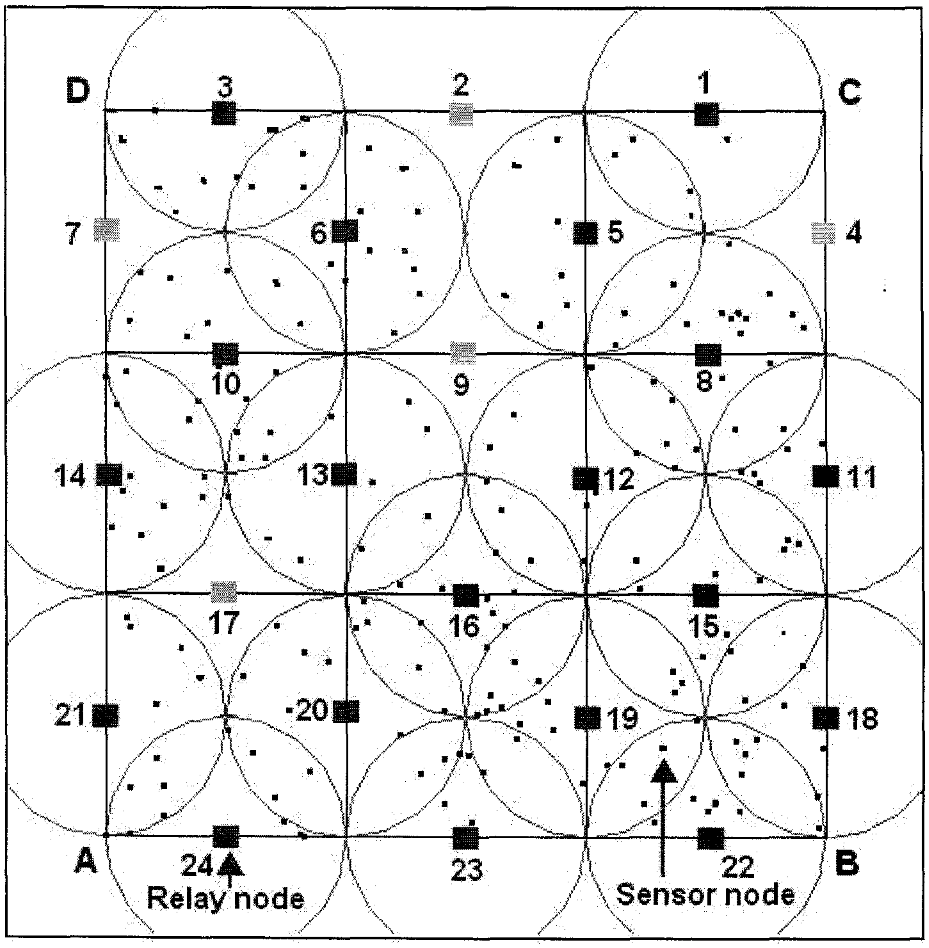


Figure 3.5: The placement of relay nodes in the network in Fig. 3.4 after applying the heuristics.

Chapter 4

ENERGY EFFICIENT ROUTING STRATEGIES

4.1 Introduction

In this chapter we will outline our approaches for determining a routing scheme for the relay node network. We assume that the number and the positions of the relay nodes have been previously determined by a suitable placement scheme, such as that given in Chapter 3. The main objective our approach is to find a routing scheme which maximizes the *lifetime* of the relay node network. A standard way to measure the lifetime is in terms of the number of *rounds* (defined in Section 4.3), until one relay node ceases to function. We have also used this measure in our work.

The dominant factor in power consumption in sensor networks is the power needed for wireless communication. To review the power model in sensor networks, the transmission power dissipated by a source node to transmit each bit of data to a destination node is given by $\alpha + \beta d^m$, where α and β are distance-independent constants, d is the distance between sender and receiver and m is the path loss exponent such that $2 \leq m \leq 4$. This cost model makes direct communication between two distant nodes much more energy consuming than communicating via a multi-hop path with smaller hop distances. The energy dissipated by a relay node, in a multi-hop routing scheme depends on a number of factors such as:

i) the number of bits of data gathered from its own cluster,

ii) the number of bits data, from other clusters, that it must forward,

iii) the distance to the next hop.

In such a data-gathering model, it is possible that certain nodes are required to relay more data, received from their neighboring nodes, as compared to some other nodes. Nodes transmitting more data will dissipate much more energy compared to the remaining nodes. Determining an optimal routing scheme that balances the load on different nodes and maximizes the network lifetime is a non-trivial task. In this chapter, we present two integer linear program (ILP) formulations for optimal data gathering and forwarding in a relay node network.

4.2 Routing for Maximizing Network Lifetime

We consider a two-tiered, cluster based sensor network where higher-powered relay nodes are used as cluster heads, each sensor node belongs to only one cluster and sends data to its respective cluster head. Each relay node is responsible for collecting and forwarding data from its own cluster as well as any data it receives from the neighboring relay node(s). The relay nodes can send data, received from their own cluster, either directly to the base station (if the base station lies within the transmission range of all relay nodes) or they can form a network and use multi-hop paths to forward data to the base station. An example of multi-hop routing by relay nodes in a two-tired sensor networks is shown in Fig. 4.2, repeated from Section 3.1. As shown in the figure, relay node A(C) collects and forwards the data from its own cluster to the relay node B(D). Node B(D) collects data from its own cluster and forwards this data to the base station, along with the data it receives from A(C).

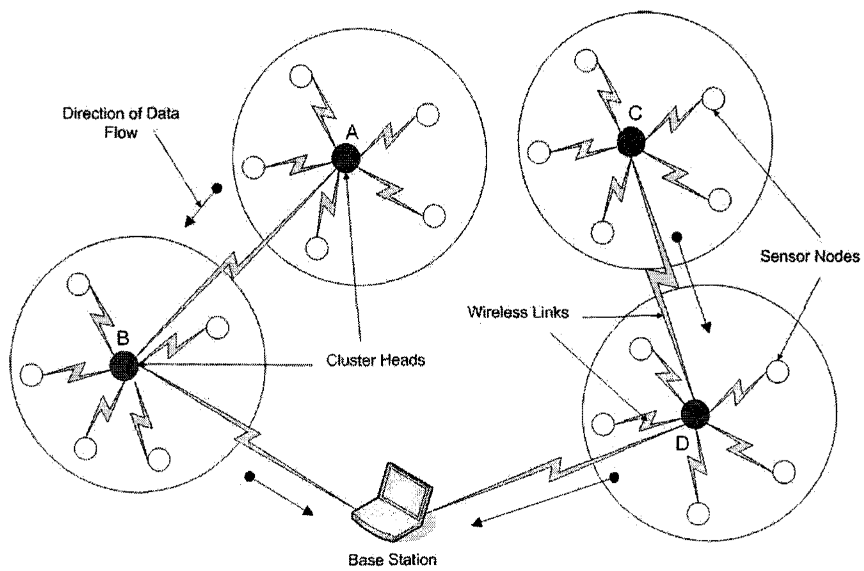


Figure 4.1: An example of multi-hop routing by relay nodes in a two-tired sensor networks

Most papers dealing with routing in a network of relay nodes adopt the "flow-splitting" model. This means that the flow of outgoing data from a single node can be divided into a number of sub-flows and sent to different destination nodes simultaneously. This approach simplifies the LP formulations, but has a number of drawbacks. The flow-splitting model requires the relay nodes to maintain routing tables and perform complicated routing functions. It also requires that each relay node be equipped with multiple transmitters. In the "non-flow-splitting" model, a relay node may receive from any number of other relay nodes, but it transmits to only one relay node or to the base station.

In our ILP formulations we adopt the non-flow-splitting model. Two widely used routing strategies, under the non-flow-splitting model are:

i) the direct transmission energy model (DTEM) and

ii) the minimum transmission energy model (MTEM)

In DTEM, each relay node transmits its data directly to the base station, in a single hop. In MTEM, each node n_i transmits to its nearest neighbor n_j, where n_j is closer to the base than n_i. If there is more than one such node, only one is selected. Assuming that initial energy provisioning for all nodes are equal and the amount of data generated in each cluster is relatively uniform, Fig. 4.2 illustrates the relative energy dissipation of different nodes, under the above two models. It is clear that in DTEM, nodes located further away from the base station dissipate more power. Therefore, their power are depleted earlier than the nodes located closer to the base station (Fig. 4.2(a)). When MTEM [21] is used, nodes located closer to the base station need to relay data at much higher rates than the nodes located further away from the base station. Therefore, nodes located near the base station deplete their energy at a faster rate and *die* sooner (Fig. 4.2(b)).

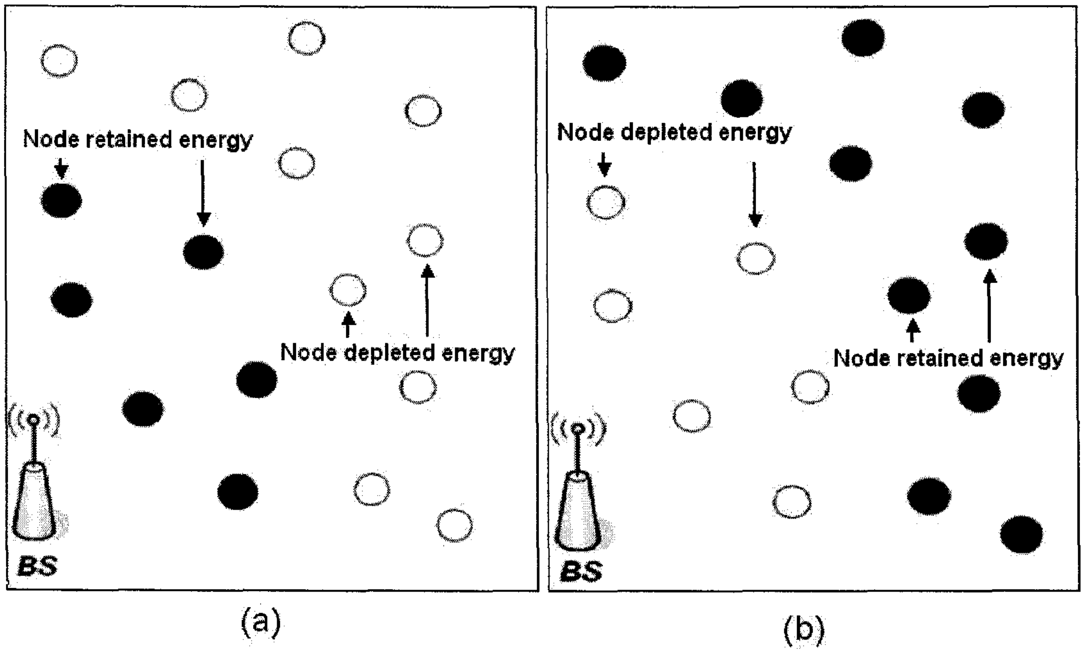


Figure 4.2: The pattern of energy dissipation by individual nodes with respect to DTEM and MTEM. (a) Effect when DTEM is used. (b) Effect when MTEM is used.

4.3 Network Model

We consider a two-tiered wireless sensor network model with $n-1$ relay nodes, labeled as node numbers $1, 2, 3, 4, ..., n-1$ and one base station, labeled as node number n. Each sensor node belongs to only one cluster and, in each cluster, one relay node acts as a cluster head of that cluster. In other words, let $\mathcal{S}$ be the set of all sensor nodes, and $\mathcal{S}^i, 1 \leq i \leq n-1$, be the set of sensor nodes belongs to the i^{th} cluster. Then, $\mathcal{S} = \mathcal{S}^1 \cup \mathcal{S}^2 \cup \ldots \cup \mathcal{S}^{n-1}$ and $\mathcal{S}^i \cap \mathcal{S}^j = \emptyset$, for $i \neq j$. Fig. 4.3 shows an example of the network model with 12 relay nodes, labeled from 1 to 12, and one base station, labeled as 13.

We have assumed that the routing schedule is computed beforehand by some centralized entity and the average amount of data generated by each cluster is known.

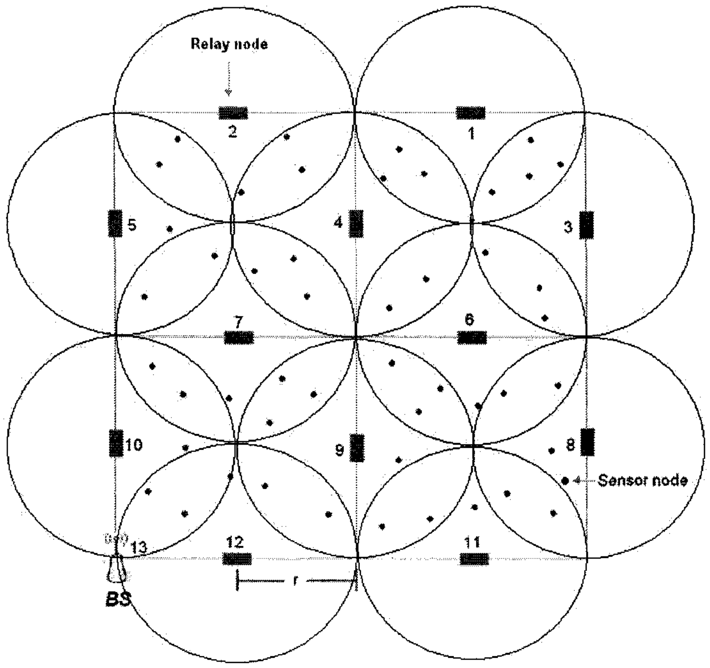


Figure 4.3: An example of the network model

We also assume that the wireless links are symmetric, i.e. if node i can transmit to node j, then node j can also transmit to node i. The connectivity of all nodes are ensured by a suitable placement strategy, applied during the deployment phase of the network. Finally, we use the first order radio model, as explained in Section 2.1.1, for representing energy dissipation of the nodes.

In our model, data gathering is *proactive*, i.e., data are collected and forwarded to the base station periodically, following a predefined schedule. Each period of data gathering is referred to as a *round* [29]. In each round of data gathering, each relay node gathers the data it receives from its own cluster and transmits that data towards the base station using multi-hop paths. It also relays any data it receives from neighboring relay nodes.

In the following sections, we will present two Integer Linear Programs, ILP-I (Section 4.5) and ILP-II (Section 4.6), for optimal data gathering and forwarding in a relay node network. To keep the formulations simple, we neglect the amount of energy, dissipated by a source relay node to receive data from its own cluster, as it has minimal impact on the total energy dissipation by the node (for receiving data from other relay node(s) and to transmit it to the destination node). ILP-I is a straightforward implementation to maximize the *lifetime* of the relay node network. ILP-II is very similar, but achieves the same objective with significantly fewer integer variables and constraints. We measure the *lifetime* of the network by the number of rounds until one relay node ceases functioning. In this situation, it is much more important to minimize the energy dissipation of the most heavily loaded relay node, than to decrease the average energy dissipation. This is exactly what we have done in our formulations.

4.4 Notation Used

In this section, we define the notation used in the two ILP formulations. Given a collection of relay nodes and a base station, along with their locations, the objective of the formulations is to find a schedule for data gathering such that the lifetime of the network is maximized.

In our ILP formulations, we are given the following data as input:

- α_1 (α_2): Energy coefficient for transmission (reception).
- β: Energy coefficient for amplifier.

- m: Path loss exponent.
- C: A large constant (for ILP-II only).
- b_i: Number of bits generated by the sensor nodes belonging to cluster i.
- $n-1$: Total number of relay nodes, with each relay node having a unique index lying between 1 and $n-1$.
- n: Index of the base station.
- d_{max}: Transmission range of each relay node.
- $d_{i,j}$: Euclidean distance from relay node i to relay node j.

We define the following *continuous* variables for the ILP formulations:

- T_i: Number of bits transmitted by node i.
- G_i: Amount of energy needed by the amplifier in relay node i to send its data to the next node in its path to the base station.
- R_i: Number of bits received by node i from other relay nodes.
- $f_{i,j}$: Amount of flow from node i to node j (used only in ILP-II).

We define the following *binary* variables for ILP-I only:

- $X^k_{i,j}$: Binary variable defined as follows:

$$X^k_{i,j} = \begin{cases} 1 & \text{if data originating in cluster } k \text{ uses the link } i \rightarrow j, \\ 0 & \text{otherwise.} \end{cases}$$

We define the following *binary* variables for ILP-II only:

- $Y_{i,j}$: Binary variable defined as follows:

$$Y_{i,j} = \begin{cases} 1 & \text{if relay node } i \text{ uses the link } i \rightarrow j \text{ to relay node } j, \\ 0 & \text{otherwise.} \end{cases}$$

4.5 The Initial Formulation (ILP-I)

Using the notation from Section 4.4, we formulate ILP-I as follows:

$$\textbf{Minimize } F_{max} \tag{4.1}$$

Subject to:

1. Flow constraint.

$$\sum_j X_{i,j}^k - \sum_j X_{j,i}^k = \begin{cases} 1 & \text{if } i = k, \forall k, i : k \neq n, \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

2. Calculate the total number of bits transmitted by node i.

$$T_i = \sum_k \sum_j b_k X_{i,j}^k, \forall i: \ i \neq n \tag{4.3}$$

3. Calculate the amplifier energy dissipated by node i to transmit to the next node.

$$G_i = \beta \sum_k \sum_j b_k X_{i,j}^k d_{i,j}^m, \forall i, i \neq n \tag{4.4}$$

4. Calculate the total number of bits received by node i from other relay node(s).

$$R_i = \sum_k \sum_j b_k X_{j,i}^k, \forall i, \; k \neq n \tag{4.5}$$

5. Constraint ensuring that all data from a given cluster k are forwarded along the same link, from node i.

$$\sum_j X_{i,j}^k \leq 1, \; \forall k, i: \; k, i \neq n \tag{4.6}$$

6. Constraint to prevent flow-splitting.

$$X_{i,j}^k \leq X_{i,j}^i, \; \forall k, i, j: \; k, i \neq n \tag{4.7}$$

7. Transmission range constraint.

$$X_{i,j}^k d_{i,j} \leq d_{max}, \forall k, i, j: \; k, i \neq n \tag{4.8}$$

8. Constraint limiting the total energy dissipated by node i.

$$\alpha_1 R_i + \alpha_2 T_i + G_i \leq F_{max}, \; \forall i: \; i \neq n \tag{4.9}$$

4.5.1 Justification of the ILP-I Equations

Equation 4.1 is the objective function that minimizes the maximum energy dissipation at individual relay nodes in one round of data gathering. Since we assume that the initial energy at each node is fixed, the lower the value of F_{max}, the higher the number of rounds of data gathering that the network can sustain. The most heavily loaded node(s) will be the one(s) that use the most energy per round, and these will be the first ones to run out of battery power. Since the *lifetime* is measured by the number of rounds before the first node depletes its battery power, minimizing F_{max} effectively maximizes the lifetime of the network.

1. Equation 4.2 is the standard flow constraints [1]. It is used to find a route, over the network, for the data originating in cluster k to the destination node n, which is the base station. In the remainder of this thesis, we will refer to the data originating in cluster k, as *commodity* k. For each commodity k, Equation 4.2 must be satisfied at each node i in the network. We have to consider three cases.

   Case 1 ($i = k$): $\sum_j X_{k,j}^k - \sum_j X_{j,k}^k = 1$

   The above equation states that there is one outgoing link (k, j) from relay node k to node j, such that $X_{k,j}^k = 1$. This is the first link in the route (from k to n), and none of the incoming edges for node k are on the route for commodity k ($\sum_j X_{j,k}^k = 0$).

   Case 2 ($i \neq k, i \neq n$): $\sum_j X_{i,j}^k - \sum_j X_{j,i}^k = 0$

   This equation holds for all nodes in the network, other than the base station n and the source node for commodity k. In this case, if i is an intermediate node in the path from k to n, there is exactly one incoming link to node i and one outgoing link from node i which are on the route associated with the k^{th} commodity. In this case, $\sum_j X_{i,j}^k = \sum_j X_{j,i}^k = 1$. If a node i is not on the selected route for commodity

k, then $\sum_j X^k_{i,j} = \sum_j X^k_{j,i} = 0$.

Case 3 ($i = n$): $\sum_j X^k_{n,j} - \sum_j X^k_{j,n} = -1$

This case is not stated explicitly, but is implied by the remaining constraints. The destination node (or sink) for all commodities k is the base station n. Therefore, the above equation states that there is one incoming link (j, n) such that $X^k_{j,n} = 1$ and this is the last link in the route (from k to n). Since n is the sink for all commodities, there are no outgoing links from n, so $\sum_j X^k_{n,j} = 0$.

2. Equation 4.3 specifies the total number of bits T_i transmitted by node i. This is obtained by summing the number of bits b_k, for each commodity whose route contains node i. This includes the data (b_i) generated in its own cluster. If a commodity is not routed over node i, then $\sum_j X^k_{i,j} = 0$ (case 2 above). Therefore, summing this value over all commodities k, will generate the total number of bits to be forwarded by node i from all the clusters.

3. Equation 4.4 is used to calculate G_i, the total amplifier energy needed at node i, by directly applying the first order radio model.

4. Equation 4.5 is used to calculate the total number bits R_i received at node i from other relay node(s), and is similar to Equation 4.3.

5. Equation 4.6 specifies that the total data from each cluster is transmitted along a single route. For each node i on the route for commodity k, $\sum_j X^k_{i,j} \leq 1$. This means there is exactly one outgoing link (i, j) from node i, that carries the data corresponding to commodity k.

6. Equation 4.7 specifies that each relay node can transmit to only one other node in the network. Since Equation 4.6 must be satisfied for all commodities, $\sum_j X^i_{i,j} = 1$. Equation 4.7 then states that all commodities routed through node i must use the same outgoing link used by commodity i. This effectively enforces the "non-flow-splitting" constraint, since it ensures that each node i cannot transmit to more than one node.

7. Equation 4.8 enforces the transmission range constraint. This equation ensures that a node i does not use link (i, j) if the distance d_{ij}, from node i to node j, is greater than the maximum transmission range d_{max}. If $d_{ij} \geq d_{max}$, equation 4.8 can only be satisfied by setting $X^k_{i,j} = 0$. If $d_{ij} \leq d_{max}$, equation 4.8 can be satisfied by either $X^k_{i,j} = 0$ or $X^k_{i,j} = 1$.

8. Equation 4.9 gives the total energy dissipated by each relay node, which the model attempts to minimize. The energy dissipated by a node i has three components:

   i) the receive energy $\alpha_1 R_i$,

   ii) the transmit electronics energy $\alpha_2 T_i$, and

   iii) the transmit amplifier energy G_i.

   The total energy dissipated by a node cannot exceed F_{max}, which the formulation attempts to minimize.

4.6 The Second Formulation (ILP-II)

The ILP presented in the previous section requires a large number of integer variables and constraints (analysis given in Section 4.7). ILP-II is a similar formulation, but the number of integer variables and constraints needed in this formulation is considerably lower. Using the notation from Section 4.4, we formulate ILP-II as follows:

$$\textbf{Minimize } F_{max} \tag{4.10}$$

Subject to:

1. Non flow-splitting constraint.

$$\sum_{j} Y_{i,j} = 1, \forall i: \ i \neq n \tag{4.11}$$

2. Calculate the total number of bits transmitted by node i.

$$T_i = \sum_{j} f_{i,j}, \forall i: \ i \neq n \tag{4.12}$$

3. Calculate the amplifier energy dissipated by node i to transmit to the next node.

$$G_i = \beta \sum_{j} f_{i,j} d_{i,j}^{m}, \forall i, i \neq n \tag{4.13}$$

4. Calculate the number of bits received by node i from other relay node(s).

$$R_i = \sum_j f_{j,i}, \forall i, i \neq n \tag{4.14}$$

5. Base station does not transmit.

$$f_{n,j} = 0, \forall j \tag{4.15}$$

6. Only one outgoing link can have non-zero data flow.

$$f_{i,j} \leq CY_{i,j}, \ \forall i, j, i \neq n \tag{4.16}$$

7. Flow constraint.

$$\sum_j f_{i,j} - \sum_j f_{j,i} = b_i \tag{4.17}$$

8. Transmission range constraint.

$$Y_{i,j} d_{i,j} \leq d_{max}, \forall i, j : \ i \neq n \tag{4.18}$$

9. Energy dissipated by node i.

$$\alpha_1 R_i + \alpha_2 T_i + G_i \leq F_{max}, \ \forall i : \ i \neq n \tag{4.19}$$

4.6.1 Justification of the ILP-II Equations

The objective function to be minimized is the same as in ILP-I (Section 4.5.1).

1. Equation 4.11 prevents flow-splitting by specifying that a node i can transmit to only one other node j.
2. Equation 4.12 calculates the total number of bits T_i transmitted by node i, by summing the data transmitted over all outgoing links from node i.
3. Equation 4.13 calculates the amplifier energy G_i, by summing the energy required for each link. In the actual solution, only one outgoing link will have non-zero data flow.
4. Equation 4.14 specifies the total number of bits received at node i from other relay node(s), by summing the data flow on all incoming links.
5. Equation 4.15 specifies that the base station n, does not transmit to any other node.
6. Equation 4.16 specifies that data can be sent from node i to node j, only if link (i, j) is selected as the single outgoing link by Equation 4.11, i.e. $Y_{i,j} = 1$. If $Y_{i,j} = 0$, then Equation 4.16 forces $f_{i,j} = 0$. The constant C is needed since the value of $f_{i,j}$ may be greater than 1. The value of C should be large enough to allow the maximum possible data flow on link (i, j). We have set $C = \sum_i b_i$.
7. Equation 4.17 corresponds to the standard flow constraints [1], and states that the total data flowing from node i ($\sum_j f_{i,j}$)is equal to the total incoming data from other relay nodes ($\sum_j f_{j,i}$) plus the data generated in cluster i (b_i).

8. Equation 4.18 enforces the transmission range constraint, stating that a node i cannot transmit to node j if they are separated by a distance greater than d_{max}. This is very similar to Equation 4.8.

9. Equation 4.19 gives the total energy dissipated by individual relay nodes, which the model attempts to minimize and is similar to Equation 4.9.

4.7 Complexity of ILP Formulations

The number of integer variables is the crucial factor determining the time required to solve an ILP. We will measure the complexity of our ILP formulations in terms of three parameters:

a. the number of integer variables,

b. the number of continuous variables, and

c. the number of constraints.

Among these three, the number of integer variables is the crucial factor determining the time required required to solve a mixed integer linear program. Table 4.1 shows the number of integer variables, the number of continuous variables and the number of constraints needed in the formulation for ILP-I and ILP-II. We can see that ILP-II requires fewer integer variables and constraints compared to ILP-I. This is accomplished at the cost of introducing some additional continuous variables.

	Number of integer variables	Number of continuous variables	Number of constraints
ILP-I	n^3	$3n+1$	$2n^3+2n^2+3n$
ILP-II	n^2	n^2+3n+1	$2n^2+7n$

Table 4.1: Number of variables and constraints needed in ILP formulations

4.8 Rescheduling the Data Gathering Scheme

The objective of our ILP formulations was to minimize the energy dissipation of the most heavily loaded node(s). However, with a fixed routing schedule, the same node has the highest load in each round. The lifetime can be further improved by recomputing the routing schedule at certain intervals. The idea is to redistribute the load on different nodes, taking into account the available residual energy of each node. To implement this, we first compute the number of rounds that can be sustained by the current schedule. We then allow the current schedule to continue for a specified number of rounds before re-computing the routing schedule. This re-computation takes into account the available residual energy of individual relay nodes, at the time of re-computation. We do this by introducing a new input, w_i that indicates the ratio of available energy to the initial energy for each relay node $i, 1 \leq i \leq n-1$. We assume that each relay node has the same initial energy, and set $w_i = 1$, $\forall i$, for calculating the initial schedule. We then update these values prior to each rescheduling to reflect the current residual energy for each relay node. Equation 4.9 for ILP formulation-I (Equation 4.19 for ILP formulation-II) in our model can then be replaced by equation (4.20). We note that in Equation 4.20, the values of w_i are treated as constants, so that it remains a *linear* constraint. This also results in a generalized formulation which can handle nodes with different levels of initial energy.

$$\alpha_1 R_i + \alpha_2 T_i + \beta G_i \leq w_i F_{min},\ \forall i:\ i \neq n \tag{4.20}$$

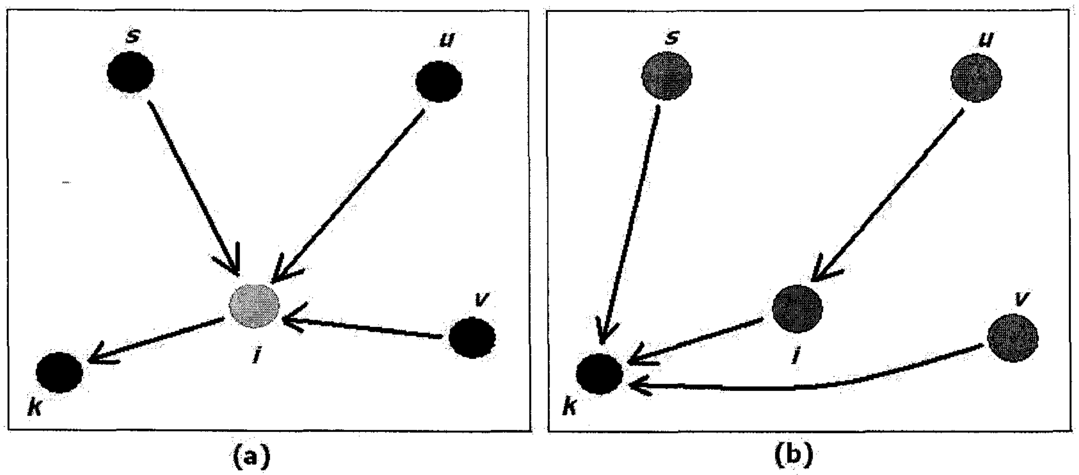


Figure 4.4: An example of recomputing the data gathering schedule in a muti-hop data transmission model

The idea of rescheduling is depicted in Fig. 4.4 with an arbitrary example. The figure shows a portion of sensor network containing 5 relay nodes, s, u, v, i and k, that is using MHDTM for forwarding data to the base station. Let, the ILP initially compute a schedule where nodes s, u and v are using node i as a hop to forward data to node k (which can be a base station), as shown in Fig. 4.4(a). This is simply because the distances, $d_{s,k}, d_{u,k}$ and $d_{v,k}$ is larger than $d_{i,k}$. Therefore, such a schedule should minimize the maximum energy dissipated by each node. But it is possible that node i dissipates more energy in each round than nodes s, u, v as it is transmitting not only the data of its own cluster, but also the data it receives from these nodes. Therefore, the network lifetime will be over as soon as the power of node i is completely depleted. Now, instead of running the network with the initial schedule for the entire time, we can let the network to operate with this schedule for certain number of rounds and then compute the residual energy of each node. If the available energy of node i is less than that of the other nodes, equation 4.20 will try to reduce the load on i by requiring the energy dissipation of node i to be lower, compared to the other nodes. This will likely result in node i being relieved of some of its burden for data forwarding. For example, as shown in Fig. 4.4(b), nodes s and v could send their data directly to k, instead of routing through i, in order to reduce the energy dissipation of node i. In this case, s and v will dissipate more energy per round (as they are transmitting to a larger distance). But since these nodes have a higher residual energy, the total lifetime of the network will be improved, compared to the lifetime that can be achieved by using the original schedule.

Such re-computation of the data gathering schedule can be performed multiple times at predetermined intervals. Lifetime improvement after each rescheduling contributes to the total lifetime of the network. The rescheduling can be performed until the nodes drain out of power or no significant improvement on the lifetime is observed.

Chapter 5

ANALYSIS OF RESULTS

5.1 Performance Evaluation for the Placement Strategy

In this section, we present the simulation results for our placement strategy. Our objective was to minimize the number of relay nodes required to form a 2-connected network, where each sensor node was covered by least one relay node. We have used an experimental setup similar to [48], where the sensor nodes were randomly distributed over a $480 \times 480\ m^2$ area. We have assumed that the transmission ranges of all sensor nodes in the network were equal and varied from $r = 24m$ to $r = 40m$. We set the transmission ranges of all relay nodes to $R = 200m$. For each value of the range of the sensor nodes, we have repeated the experiments with 600, 800, 1000, 1200 and 1400 sensor nodes, randomly distributed within the network region.

Table 5.1 gives the results of the experiments, for $r = 24m$, $r = 30m$ and $r = 40m$. For each range, the table shows the initial number of relay nodes, computed following step 2 and 3 of the *fp* strategy (Section 3.2), after dividing the networking

Sensor range	Strategy step	600 nodes	800 nodes	1000 nodes	1200 nodes	1400 nodes
r=40	RN-fixed	84	84	84	84	84
	RN-minimized	66	68	71	72	73
r=30	RN-fixed	144	144	144	144	144
	RN-minimized	102	105	110	117	117
r=24	RN-fixed	220	220	220	220	220
	RN-minimized	137	149	157	165	172

Table 5.1: Number of relay node required for different number of sensor nodes with respect to different transmission range.

area into imaginary cells of size $2r \times 2r$ (step 1 of the *fp* strategy). We have indicated these values as "Number of Relay Nodes with fixed placement (RN-fixed)" in Table 5.1. In the *fp* strategy, the initial positions of the relay nodes depends only on the network size and the transmission ranges of the sensor nodes, and not on the number of sensor nodes. Therefore, in Table 5.1, this value does not vary with the number of sensor nodes.

However, the actual number of relay nodes, required to cover the network, varies with the number and the distribution of the sensor nodes. We have obtained these values using the heuristic, given in Section 3.5 (step 4 of the *fp* strategy (Section 3.2)). We have indicated, in Table 5.1, the required numbers of relay nodes, as computed by the heuristic, as "Minimized number of Relay Nodes (RN-minimized)". Our heuristic reduces, considerably, the number of relay nodes required to cover the entire network, compared to the initial assessment.

Fig. 5.1 shows, for different values of the communication ranges (r) of the sensor nodes, how the number of relay nodes changes with the number of sensor nodes. As expected, we see that, as the communication range of a sensor node is decreased, more relay nodes are required to adequately cover the same sensing area.

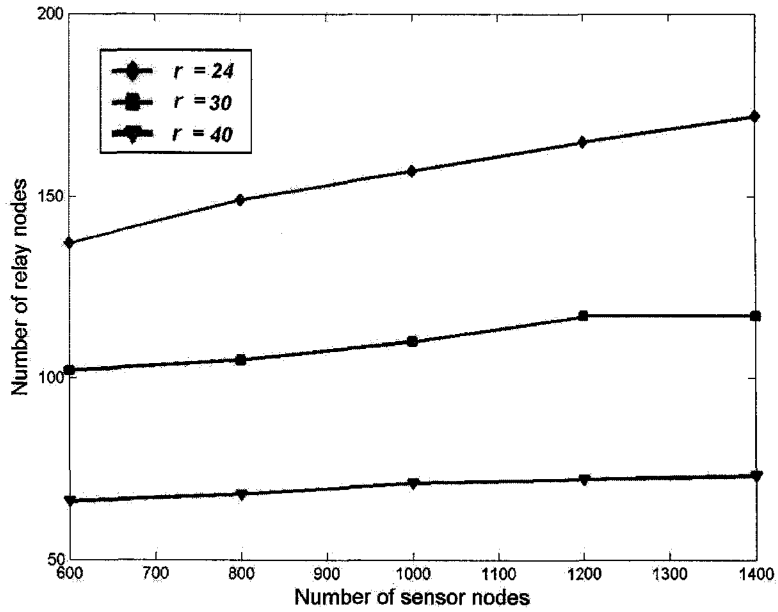


Figure 5.1: Variation of the number of relay nodes with the number of sensor nodes.

5.2 ILP Performance Evaluation

We have simulated our routing scheme using a network with twelve relay nodes, and with sensor nodes randomly distributed in a 160m × 160m field. We have assumed that the base station is located at coordinate $(0, 0)$. We have shown our experimental setup in Fig. 5.2.

We have measured the achieved lifetime of the network by the number of rounds until the first relay node runs out of battery power. The arrangement of the relay nodes is similar to that shown in Fig. 3.3. For experimental purposes, we have assumed that each relay node receives data at a rate of 1000 *bits/round*, from sensor nodes in its cluster. Such uniformity for the amount of data is not a requirement for our model as long as the average amount of data generated by each cluster is known

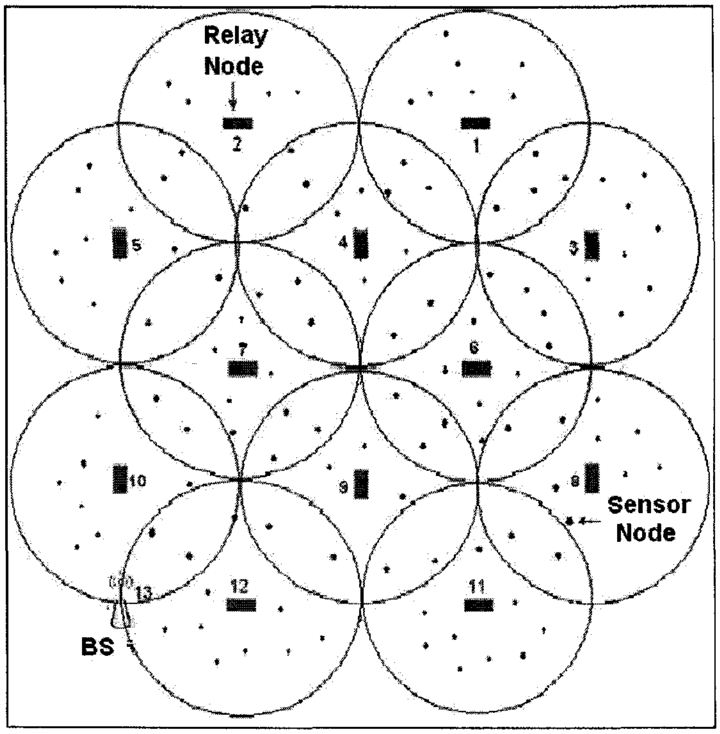


Figure 5.2: The experimental network setup for the performance evaluation of the ILP.

beforehand.

We have assumed that:

i. the communication energy dissipation is based on the first order radio model, described in Section 2.1.1.

ii. the values for the constants are the same as in [21], so that:

a. $\alpha_1 = \alpha_2 = 50nJ/bit$,

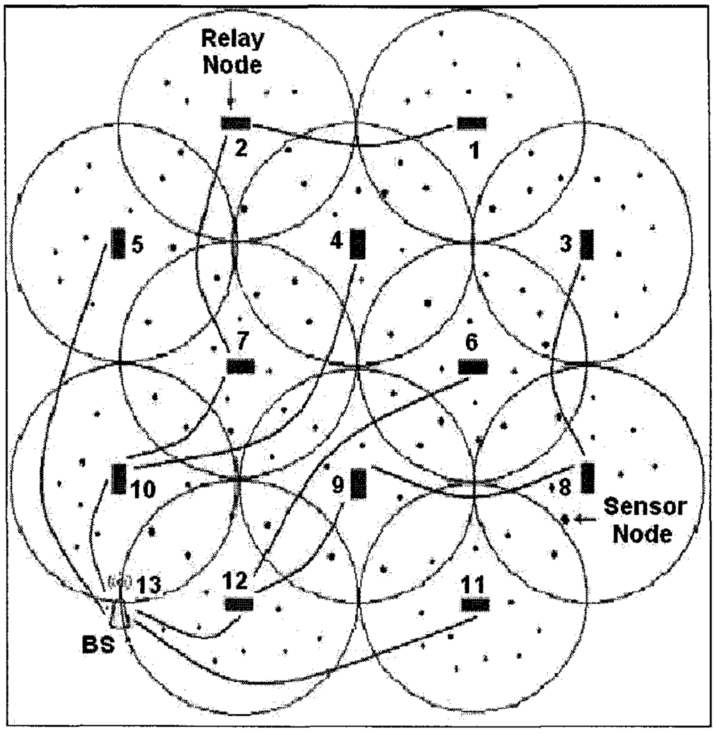
b. $\beta = 100pJ/bit/m^2$ and

c. the path-loss exponent, $m = 2$.

iii. the range of each sensor (relay) node is $40m$ ($200m$), as in [48].

iv. the initial energy of each relay node was $5J$, as in [48].

We have shown, in Fig. 5.3, the data gathering schedule computed by the ILP for a network with 12 relay nodes.



(a)

Figure 5.3: The data gathering schedule computed by the ILP.

For the small network described above, the direct energy model is applicable. The results show that our method can achieve an improvement of more than 2.71 times the network lifetime, compared to *Direct Transmission Energy Model (DTEM)*[21]. The experiment shows that relay nodes 5 and 11 transmit directly to the base station and dissipate the largest amount of energy in each round. Therefore, these are the nodes that decide the lifetime of the network.

We have further improved this initial solution by recomputing the routing schedule at certain intervals. To implement this, we first computed the number of rounds that can be sustained by the current schedule. We then allowed the current schedule to continue until 20% of the maximum lifetime, that can be achieved by the node(s) dissipating the most power in the current schedule, has expired. At that point, we have re-computed the routing schedule. We have shown an example of the effect of rescheduling in Fig. 5.4. Fig. 5.4(a) gives the initial schedule. In this schedule, relay nodes 5 and 11 dissipate the most power by directly transmitting to the base station. Fig. 5.4(b) shows the new data gathering scheme after the first rescheduling. As shown in the figure, after considering the residual energy, the new schedule is such that node 5 transmits to node 10 and node 11 transmits to node 12. In this way, we have reduced the transmission distance, and hence the rate of energy dissipation, for both nodes 5 and 11. The load on nodes 10 and 12 are increased, but since they had higher residual energy, we have improved the overall lifetime of the network. Subsequent rescheduling keeps on reassigning the data gathering scheme in a similar way based on the residual energy of each relay node.

In Table 5.2, we have compared the achieved lifetime of the network, using our model, the *Maximum Lifetime for Relay Nodes Model (MLRNM)*, with the Direct Transmission Energy Model (DTEM) [21], at different rescheduling points. The first row shows the achieved lifetimes, without any rescheduling. The remaining rows indicate the values after five, ten, fifteen and twenty rescheduling, respectively. The first column in the Table 5.2 indicates the rescheduling points (labeled as "Resch. pts."). The second column and the third columns show the lifetime that can be

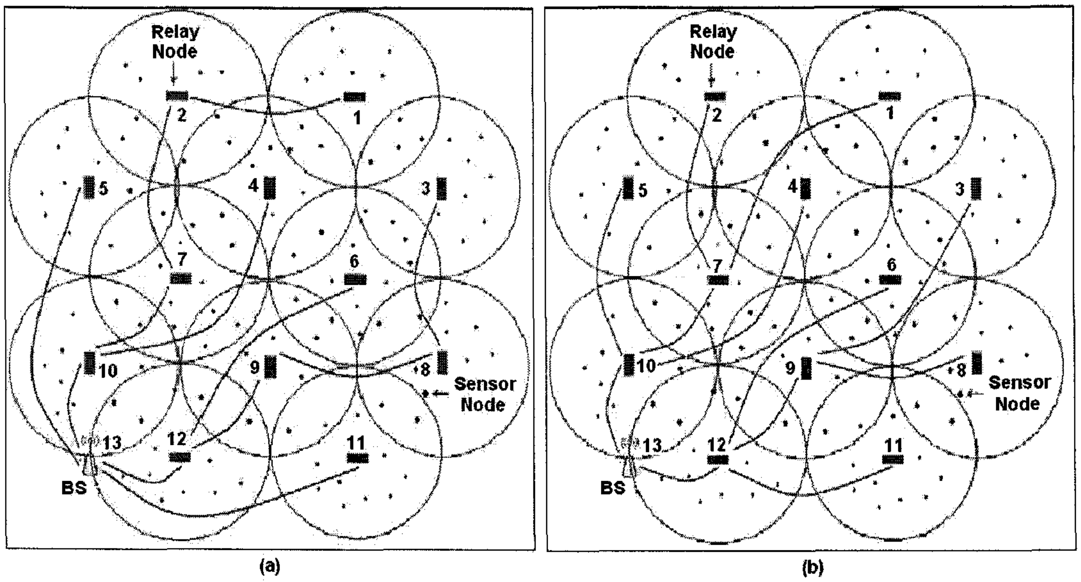


Figure 5.4: An example of the effect of rescheduling data gathering scheme using ILP.

Resch. pts.	*DTEM* lifetime	*MLRNM* lifetime	Improvement on $MLRNM_0$	Improvement on *DTEM*
0	1234	3355		271%
5	1234	3651	9%	296%
10	1234	3836	14%	310%
15	1234	3931	17%	318%
20	1234	3961	18%	321%

Table 5.2: Number of rounds achieved by each relay nodes after rescheduling.

achieved by the DTEM, and the MLRNM, respectively. We have denoted the lifetime achieved by the first schedule using MLRNM as $MLRNM_0$. The fourth column of the Table 5.2 indicates the improvements of the lifetime on $MLRNM_0$, after using the number of reschedules indicated in the corresponding rows. The last column shows the lifetime improvements on DTEM, using MLRNM and rescheduling. Table 5.2 shows that, after twenty rescheduling, our model can achieve a performance improvement of 18% over the initial schedule, $MLRNM_0$, and an improvement up to 321% over the direct transmission energy model, DTEM.

We have shown that the lifetime improvement of the network after each rescheduling in Fig. 5.5. The x-axis of the figure represents the rescheduling points and the y-axis represents the number of rounds before the first relay node runs out of battery power. Fig. 5.5 indicates that, initially, the rescheduling results in a quick increase of the network lifetime. Then, the rate of improvement with successive rescheduling becomes slower, and after about 20 rescheduling, the improvement is negligible.

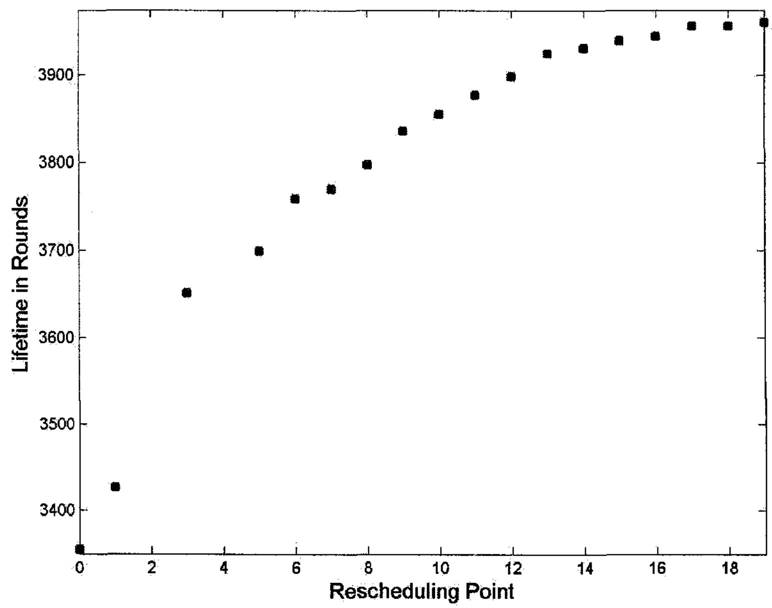


Figure 5.5: Lifetime improvement by rescheduling

In a second set of experiments, we have considered a $240 \times 160m$ sensing area, containing 18 relay nodes. For larger networks, such as this, the direct energy model (DTEM) may not always be applicable, due to the limited communication range of the sensor nodes. Therefore, for this case, we have compared our method to the Minimum Transmit Energy model (MTEM) [21], as well as to DTEM. Initial results indicate a lifetime improvement of more than 20% over MTEM, and 3.5 times over DTEM (which was tested after appropriately extending the range of relay nodes) without any rescheduling.

In our model, we have assumed that there is a central agent to compute the relay schedule. We have also assumed that the average amount of data generated by each cluster is known beforehand. Once we have computed the schedule, we can either load the schedule in the relay nodes during the deployment of the network or the base station can broadcast the schedule to each relay node. Broadcasting the schedule from the base station may be particularly useful in networks where the average amount of data generated by each cluster is not known *a priori*, or changes with time. In such networks, we can compute the relay schedule reactively and on the fly. For this, we may require the relay nodes to report their residual energy periodically to the base station. The base station can then determine the relay schedule, based on this information, and broadcast the schedule to the relay nodes in the network.

Chapter 6

CONCLUSION AND FUTURE WORKS

In this thesis we have presented:

i. a placement strategy for relay nodes, and

ii. an optimal routing scheme, for relay nodes, in two-tiered sensor networks.

Our placement strategy is scalable and efficient. This approach can provide fast solutions, requiring very little computation. The performance ratio of our placement scheme is comparable to existing schemes, which require significantly more computation. We have also proved that our placement scheme guarantees a topology where the relay node network is at least 2-connected, so that our scheme can handle single faults.

We have proposed two ILP formulations that can maximize the lifetime of the relay node network by making the routing decisions in an energy efficient way.

Finally, we have introduced a rescheduling technique that can further extend the maximized lifetime of a sensor network. We have shown that our model can extend the network lifetime, compared to both the direct transmission energy model, as well as the minimum-transmission-energy model.

6.1 Future Work

In this thesis, we have addressed the placement and routing problems separately. However, the two problems are interrelated and greater improvements may be achieved, if they are considered together. We are currently investigating the joint problem of maximizing the lifetime and minimizing the number of relay nodes by determining optimal location of relay nodes within the sensing field.

In Chapter 3, we have used a simple, greedy heuristic to assign each sensor node to a cluster. An efficient clustering scheme can also play an important role in extending the liftime of the sensor network. Such a scheme would take into account the load on each relay node and its distance from its neighbors, before assigning sensor nodes to its cluster.

Our placement strategy considers the fault tolerance of the relay node network and ensures that it is 2-connected. However, if a relay node fails, the sensor nodes it its cluster become disconnected from the network. The placement strategy could be extended so that each sensor node is *covered* by at least two relay nodes. This will ensure that, even if a relay node fails, the data from the sensors in the affected cluster will not be lost.

Bibliography

[1] R. K. Ahuja, T. L. Magnanti, and J. B. OrLin. *Network flows.* Prentice-Hall, Inc., 1993.

[2] K. Akkaya and M. Younis. A survey on routing protocols for wireless sensor networks. *IEEE Transactions On Mobile Computing*, 3(3):325–349, 2005.

[3] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38:393–422, 2002.

[4] J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor networks: A survey., 2005. to appear in *IEEE Wireless Communications*, http://vulcan.ee.iastate.edu/~ kamal/Docs/kk04.pdf.

[5] K. Chakrabarty, S. S. Iyengar, H. Qi, and E. Cho. Grid coverage for surveillance and target location in distributed sensor networks. *IEEE Transactions on Computers*, 51(12):1448–1453, 2002.

[6] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. SPAN: an energy efficient coordination algorithm for topology maintenance in ad hoc wireless networks. In *Proc. ACM Mobicom*, volume 283, pages 85–96, 2001.

[7] P. Cheng, C-N. Chuah, and X. Liu. Energy-aware node placement in wireless sensor network. In *Proc. of IEEE Global Telecommunications Conference*, volume 5, pages 3210–3214, 2004.

[8] X. Cheng, D-Z. Du, L. Wang, and B. B. Xu. Relay Sensor Placement in Wireless Sensor Networks. *IEEE Transactions on Computers*, 2001. http://citeseer.ist.psu.edu/cheng01relay.html.

[9] C.-Y. Chong and S. P. Kumar. Sensor Networks: Evolution, Opportunities, and Challenges. *Proc. of the IEEE*, 91(8):1247–1256, 2003.

[10] S. Coleri and P. Varaiya. Optimal Placement of Relay Nodes in Sensor Networks. submitted to *ICPP*, 2005.

[11] K. Dasgupta, M. Kukreja, and K. Kalpaki. Topology-aware placement and role assignment for energy-efficient information gathering in sensor networks. In *Proc. of Eighth IEEE International Symposium on Computer and Communication*, pages 341–348, 2003.

[12] S. S. Dhillon and K. Chakrabarty. Sensor placement for effective coverage and surveillance in distributed sensor networks. In *Proc.of IEEE Wireless Communications and Networking Conference*, volume 3, pages 1609–1614, 2003.

[13] E. J. Duarte-Melo and M. Liu. Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks. In *Proc. IEEE Global Telecommunications Conference*, volume 1, pages 21–25, Taipei, Taiwan, 2002.

[14] E. Falck, P. Floren, P. Kaski, J. Kohonen, and P. Orponen. Balanced data gathering in energy-constrained sensor networks., 2004. In S. Nikoletseas and J. D. P. Rolim, editors, Algorithmic Aspects of Wireless Sensor Networks: First International Workshop (ALGOSENSORS 2004, Turku, Finland, July 2004), volume 3121 of Lecture Notes in Computer Science, pages 59–70, Berlin Heidelberg, 2004. Springer-Verlag.

[15] G. Gupta and M. Younis. Fault-tolerant clustering of wireless sensor networks. In *Proceedings of IEEE WCNC* , pages 1579–1584, 2003.

[16] G. Gupta and M. Younis. Load-balanced clustering of wireless sensor networks. In *IEEE International Conference on Communications*, volume 3, pages 1848–1852, 2003.

[17] G. Gupta and M. Younis. Performance evaluation of load-balanced clustering of wireless sensor networks. In *10th International Conference on Telecommunications*, volume 2, pages 1577–1583, 2003.

[18] B. Hao, J. Tang, and G. Xue. Fault-tolerant relay node placement in wireless sensor networks: formulation and approximation. In *Workshop on High Performance Switching and Routing (HPSR)*, pages 246–250, 2004.

[19] T. He, J.A. Stankovic, C. Lu, and T. Abdelzaher. SPEED: A Stateless Protocol for Real-Time Communication in sensor networks. In *Proceedings of 23rd International Conference on Distributed Computing Systems*, pages 46–55, 2003.

[20] W. Heinzelman. *Application-specific Protocol Architectures for Wireless Networks.* PhD thesis, Massachusetts Institute of Technology, 2000.

[21] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy efficient communication protocol for wireless micro-sensor networks. In *Proc. of the 33rd HICSS*, pages 3005–3014, Maui, Hawaii, 2000.

[22] W. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive Protocols for Information Dissemination in Wireless Sensor Networks. In *Proc. 5th ACM/IEEE Mobicom Conference*, pages 174–185, 1999.

[23] D. S. Hochbaum and W. Maass. Approximation schemes for covering and packing problems in image processing and VLSI. In *journal of ACM*, volume 32, pages 130–136, 1985.

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

[25] Y. T. Hou, Y. Shi, J. Pan, and S. F. Midkiff. Lifetime-optimal data routing in wireless sensor networks without flow splitting. In *Workshop on Broadband Advanced Sensor Networks*, San Jose, CA, 2004. http://www.broadnets.org/2004/basenets.html.

[26] Y. T. Hou, Y. Shi, H. D. Sherali, and S. F. Midkiff. On Energy Provisioning and Relay Node Placement for Wireless Sensor Networks. *IEEE Transactions on Wireless Communications*, 4(5):2579–2590, 2005.

[27] A. Howard, M. J. Mataric, and G. S. Sukhatme. Mobile Sensor Network Deployment using Potential Fields: A Distributed, Scalable Solution to the Area Coverage Problem. In *Distributed Autonomous Robotics Systems*, pages 299–308, 2002. H. Asama and T. Arai and T. Fukuda and T. Hasegawa (eds), Springer, 2002.

[28] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: a scalable and robust communication paradigm for sensor networks. In *ACM Mobicom*, pages 56–67, 2000.

[29] K. Kalpakis, K. Dasgupta, and Namjoshi. P. Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks. Technical Report TR CS-02-12, University of Maryland Baltimore County, August 2002. http://www.csee.umbc.edu/ kalpakis/homepage/papers/tr-02-12.pdf.

[30] B. Krishnamachari and F. Ordóñez. Analysis of Energy-Efficient, Fair Routing in Wireless Sensor Networks through Non-linear Optimization. In *Proceedings of IEEE 58th Vehicular Technology Conference*, volume 5, pages 2844–2848, 2003.

[31] J. Kulik, W. R. Heinzelman, and H. Balakrishnan. Negotiation-based protpcols for disseminating information in Wireless Sensor Networks. In *Wireless Networks*, volume 8, pages 169–185, , 2002.

[32] H. Liu, P. Wan, and W. Jia. Fault-Tolerant Relay Node Placement in Wireless Sensor Networks. In *COCOON*, 2005. submitted for journal publication, also available at http://www.public.asu.edu/~jtang3/Doc/FTRP_HPSR04.pdf.

[33] A. Manjeshwar and D. P. Agarwal. TEEN: a routing protocol for enhanced efficiency in wireless sensor networks. In *Proceedings 15th International Parallel and Distributed Processing Symposium*, pages 2009–2015, 2001.

[34] A. Manjeshwar and D. P. Agarwal. APTEEN: A hybrid protocol for ecient routing and comprehensive information retrieval in wireless sensor networks. In *Proceedings International Parallel and Distributed Processing Symposium*, pages 195–202, 2002.

[35] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. Srivastava. Coverage problems in wireless ad-hoc sensor networks. In *Proc. IEEE Infocom*, pages 1380–1387, 2001.

[36] J. B. M. Melissen and P. C. Schuur. Improved coverings of a square with six and eight equal circles. *Electronic Journal of Combinatorics*, 3(1), 1996. R32, http://www.combinatorics.org/Volume_3/Abstracts/v3i1r32.html.

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

[38] F. Ordóñez and B. Krishnamachari. Optimal Information Extraction in Energy-Limited Wireless Sensor Networks. 2004. To appear in IEEE JSAC, special issue on Fundamental Performance Limits of Wireless Sensor Networks.

[39] Y. Ossama and F. Sonia. HEED: A Hybrid, Energy-Efficient, Distributed Clus-

tering Approach for Ad Hoc Sensor Networks. *IEEE Transactions On Mobile Computing*, 3(1):366–379, 2004.

[40] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen. Topology Control for Wireless Sensor Networks. In *Proc. of Ninth Annual International Conference on Mobile Computing and Networking*, pages 286–299, 2003.

[41] M. Patel, R. Chandrasekaran, and S. Venkatesan. Energy Efficient Sensor, Relay and Base Station Placements for Coverage, Connectivity and Routing. In *24th IEEE International Performance Computing and Communications Conference (IPCCC)*, pages 581 - 586, 2005.

[42] V. Rai and R. Mahapatra. Lifetime Modeling of a Sensor Network. In *IEEE Intl. Conf. on Design, Automation and Test in Europe (DATE).*, volume 1, pages 202–203, 2005.

[43] G. T. Sibley, M. H. Rahimi, and G. S. Sukhatme. A Tiny Mobile Robot Platform for Large-Scale Ad-hoc Sensor Networks. In *Proceedings of International Conference on Robotics and Automation*, volume 2, pages 1143–1148, 2002.

[44] A. Srinivas and E. Modiano. Minimum energy disjoint path routing in wireless ad-hoc networks. In *Proceedings of the 9th annual international conference on Mobile computing and networking.*, pages 122–133, San Diego, CA, USA, 2003.

[45] J. Suomela. Relay Placement in Sensor Networks. Master's thesis, University Of Helsinki, Department of Computer Science, 2005.

[46] J. Suomela. Computational Complexity of Relay Placement in Sensor Networks., 2006. Accepted for the *SOFSEM* 2006 conference.

[47] J. W. Suurballe. Disjoint paths in a network. In *Networks*, volume 4, pages 125–145, 1974.

[48] J. Tang, B. Hao, and A. Sen. Relay node placement in large scale wireless sensor networks. *Computer Communications*, 29(4):490–501, 2006.

[49] D. Tian and N. D. Georganas. A coverage-preserving node scheduling scheme for large wireless sensor networks. In *Proceedings of ACM WSNA*, pages 32–41, 2002.

[50] W. Wang, G. Xing, Y. Zhang, C. Lu, R. Pless, and C. Gill. Integrated coverage and connectivity configuration in wireless sensor networks. In *Proceedings of ACM SenSys*, pages 28–39, 2003.

[51] H. Zhang and J. C. Hou. On the upper bound of α-lifetime for large sensor networks. In *ACM Transaction on Sensor Networks.* To appear.

[52] H. Zhang and J. C. Hou. Maintaining Sensing Coverage and Connectivity in Large Sensor Networks. In *Ad Hoc and Sensor Wireless Networks*, volume 1, pages 89–124, 2005.