2009

# Integrated placement and routing of relay nodes for fault-tolerant hierarchical sensor networks

Yufei Xu

*University of Windsor*

# INTEGRATED PLACEMENT AND ROUTING OF RELAY NODES FOR FAULT-TOLERANT HIERARCHICAL SESNOR NETWORKS

**BY**
**Yufei Xu**

A Thesis
Submitted to the Faculty of Graduate Studies
through the School of Computer Science
in Partial Fulfillment of the Requirements for
the Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada

2009

*Your file Votre référence*
*ISBN:* 978-0-494-57638-0
*Our file Notre référence*
*ISBN:* 978-0-494-57638-0

# Declaration of Co-Authorship

I hereby declare that this thesis incorporates material that is result of joint research, as follows:

*This thesis also incorporates the outcome of a joint research undertaken in collaboration with Ataul Bari under the supervision of Dr. Arunita Jaekel. The collaboration is covered in Chapter 3, Chapter 4 and Chapter 5 of the thesis. In all cases, the key ideas, primary contributions, experimental design, data analysis and interpretation, were performed by the author, and the contribution of co-authors primarily through the provision of some key ideas and constructive criticism.*

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

# Abstract

In two-tiered sensor networks, using higher-powered relay nodes as cluster heads has been shown to lead to further improvements in network performance. Placement of such relay nodes focuses on achieving specified coverage and connectivity requirements with as few relay nodes as possible. Existing placement strategies typically are unaware of energy dissipation due to routing and are not capable of optimizing the routing scheme and placement concurrently.

We, in this thesis, propose an integrated integer linear program (ILP) formulation that determines the minimum number of relay nodes, along with their locations and a suitable communication strategy such that the network has a guaranteed lifetime as well as ensuring the pre-specified level of coverage ($k_s$) and connectivity ($k_r$). We also present an intersection based approach for creating the initial set of potential relay node positions, which are used by our ILP, and evaluate its performance under different conditions. Experimental results on networks with hundreds of sensor nodes show that our approach leads to significant improvement over existing energy-unaware placement schemes.

*To all the people who work me through the way of making this achievement,*

*especially to my parents, my father and mother in law and my wife.*

# Acknowledgements

The research work presented in this thesis received help from many people. First and most important, I would like to take this opportunity to express my gratitude to my supervisor, Dr. Arunita Jaekel. Without her advice, guidance and encouragement, this research work would not have been completed.

I would also like to appreciate my committee members, Dr. Janguo Lu, Dr. Xiang Chen and Dr. Xiaobu Yuan for their precious time spent in reading my thesis and for their valuable comments and suggestions on my work.

Appreciation also goes to my peer, Autal Bari, for his help and advices which inspired me.

Finally, I am grateful to all my family members for providing me continuous support and encouragements during the way of pursuing my master degree.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1 Introduction

## 1.1 Sensor Networks

A sensor network, as its name suggests, interconnects a number of tiny, low-cost, low-power, and multifunctional sensing devices (called sensors) and is usually deployed to measure/detect intended physical phenomena within a geographical area. Sensor nodes in networks combine technological advances in sensing, computation, communication and operate, among themselves, in a cooperative manner to achieve the objective of deployment. Sensor networks, in recent years, have gained popularity in both military and civilian applications due to significant cost efficiency, ease of deployment and reliable performance even in hostile environment. Application scenarios of a sensor network have been extended to various aspects such as real-time tracking, habitat monitoring, parameter measurements, and military surveillance, etc.

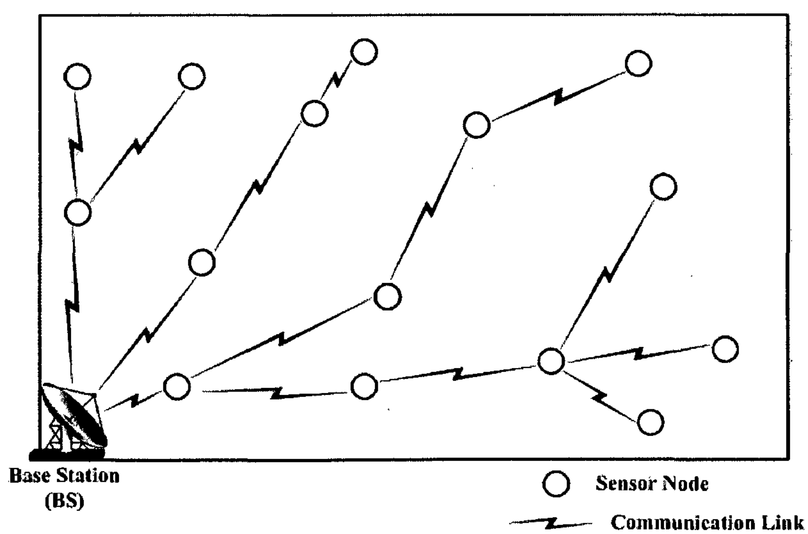


**Figure 1.1 General layout of a sensor network**

In addition to regular sensor nodes, a sensor network typically contains a base station (BS) which serves as a central repository to collect sensed data from all sensor nodes. Unlike regular sensor nodes, a base station, in a sensor network, is usually located at a fixed position and supplied with unlimited power (e.g. plugged to a wall outlet).

As shown in Figure 1.1, a sensor network is usually deployed within a geographical area (called *sensing field* shown as a rectangle border) containing the physical phenomena of interest. Sensor nodes (shown as white dots) are distributed inside the sensing field in order to carry out the sensing task effectively and accurately. Once in operation, data obtained from sensor nodes by sensing their respective vicinities, is continuously reported to the base station (shown as satellite dish) following an appropriate routing path (shown as communication links). A base station, on the other hand, is responsible for processing, analyzing and extracting meaningful information from those collected data to provide an entire view of the sensing field being monitored.

Factors, such as tiny in dimension, unattended operation and cost concerns, pose restrictions in the designated capabilities of sensor nodes. The major limitations that constrain the functionality of sensor nodes include [3], [8] and [26]:

- *Limited transmission range*: The built-in communication unit of a sensor node has limited transmission range. Therefore, if the base station is located too far away from a sensor node, that sensor node might not able to directly transmit its sensed data to the base station.

- *Prone to failures*: Nodes in sensor networks are often prone to failures, particularly when deployed in hostile environment, where chances of damage/destructions are significantly high. The physical failure of a sensing node can lead to lose of data from area where the failed node is deployed.

- *Limited energy supply*: A sensor node is usually powered by a small battery which is supplied with a limited amount of energy. In sensor networks, recharging or exchanging the batteries of sensor nodes is generally considered too costly to carry out. Therefore, once the battery is completely dissipated, a sensing device will be out of operation and lose its functionality [3], [26].

Presented with such challenges, the major concerns in the design of sensor networks are scalability, fault tolerance, and energy conservation [8]. Scalability requires sensor networks to be adaptive to frequent changes in operating conditions which include, for example, addition/removal of sensor nodes in a network or the scale variation of the sensing field. Factors, such as energy depletion, harsh environmental conditions and malicious attacks, might lead to node failures in sensor networks. Fault tolerance techniques allow a network to survive form failures and continue operation in the presence of faults. Battery power, in sensor networks, is considered one of the most precious resources as recharging or replacement of battery is infeasible for both economical and physical concerns. Given initial energy supply, a sensor node, if operating at a high data transmission rate, can only remain functional for a fairly short period of time. Therefore, an energy-aware network design is directly related to the lifetime of the network.

### 1.1.1 Relay Nodes in Hierarchical Sensor Networks

To address the above mentioned issues, hierarchical sensor networks (also known as *two-tiered sensor networks*) have been proposed in recent years. In hierarchical architecture, as shown in Figure 1.2, sensor nodes (shown as white dots) are grouped into clusters (enclosed in a dashed circle) and form the lower-tier. Each cluster is assigned a cluster head (shown as red dot) and all cluster heads plus base station (BS) compose the upper tier. Each sensor node belongs to only one cluster and sends sensed data directly to its cluster head instead of the base station. Cluster heads in upper tier are dedicated for reporting data collected from their clusters to the base station. Separating sensing and routing tasks into different tiers helps to improve network performance with respect to lifetime, fault tolerance and scalability.

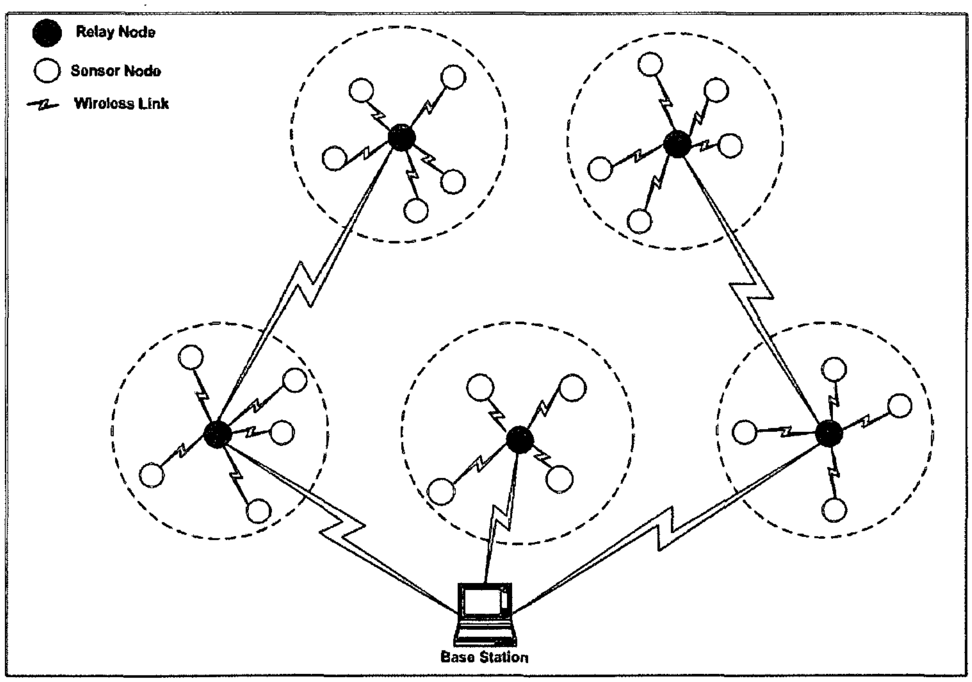


**Figure 1.2 General layout of a hierarchical sensor network**

Cluster heads, while forwarding data to the base station, normally employ the *multi-hop data transmission model* (MHDTM) [8], [12], [14], [27] in order to achieve energy conservation in routing. In MHDTM, cluster heads located too far to reach the base station with a single hop use other cluster heads as intermediate nodes to relay data to the base station. Referring to this scenario, it is possible that some cluster heads are required to transmit more data compared with other cluster heads. Thus, these cluster heads may dissipate energy at higher rates than those not relaying (or relaying very little) data from other cluster heads. Such uneven energy dissipation among cluster heads may lead to the faster "death" of some cluster heads due to the complete depletion of batteries. This unbalanced energy dissipation also has an undesirable impact on network lifetime as it may cause a network to prematurely lose its usefulness while many other cluster heads still retain power.

One method, proposed in [3], [12], to address the uneven energy dissipation among cluster heads, is to deploy a special kind of nodes called *relay nodes* (also called *Gateway nodes* or *Aggregation and Forwarding nodes* (AFN)) as cluster heads. Relay nodes are equipped with high-power batteries and built with additional capabilities in order to achieve various objectives [6], [7], [8], [10], [12], such as balanced data gathering, reduction of transmission range, connectivity and fault tolerance [3], [4], [5].

## 1.2 Motivation

In hierarchical sensor networks where relay nodes are used as cluster heads to form the upper tier network and communicate in a multi-hop fashion, two important design issues need to be considered:

- The placement strategy of relay nodes
- The routing strategy among relay nodes

The placement strategy is responsible for determining the minimum number of relay nodes along with their locations such that each sensor node can communicate with at least one relay and the relay node network is connected. It has been proved in [15] that finding the optimal placement of relay nodes in sensor networks is NP-hard. Under fault-free conditions, a network will function as long as each sensor node can communicate with at least one relay node and the relay node network is connected, so that each relay node is able to find a path to the base station. However, in such a network, the failure of even a single relay node results in data loss not only from all sensor nodes belonging to its own cluster, but also from other relay nodes, which are using the failed node to forward data towards the base station. In order to protect the network against faults, it is necessary to introduce redundancy in the network, in the form of additional relay nodes, so that each sensor node can communicate with multiple ($k_s$) relay nodes and each relay node can forward its data to multiple ($k_r$) relay nodes (or directly to the base station). The desired level of redundancy (i.e. the values of $k_s$ and $k_r$) will depend on the intended application and the goal is to achieve this with as few relay nodes as possible.

The lifetime of a sensor network is typically determined by the battery power of the "critical node(s)" in the network [2], [3]. Therefore, it is extremely important to devise strategies that extend the lifetime of the sensor network as a whole. The relay nodes, although provisioned with higher power, are also battery operated. As the transmit energy dissipation increases rapidly with the distance between the source and the destination nodes [2], the actual routing strategy has a significant impact on the network lifetime and must be determined with care.

## 1.3 Objective of Study and Contribution

Existing placement strategies decouple the placement and routing schemes. First, the positions of relay nodes are determined and then an appropriate routing schedule is developed based on this information. Therefore, the placement algorithms do not take into account the energy dissipation of the relay nodes, which requires knowledge of the routing scheme. Unlike previous approaches, we focused on jointly optimizing both placement and routing of relay nodes in hierarchical sensor networks. The proposed approach not only designs a network that meets the coverage and connectivity requirements, but also finds a routing schedule that ensures the energy dissipation of each relay node does not exceed a specified amount. The main contributions of this thesis are as follows:

1. We present an ILP formulation that jointly optimizes the placement and routing of relay nodes in a hierarchical sensor network such that the network meets specified coverage, connectivity and energy requirements.

2. We propose an intersection based approach, for determining the potential positions of relay nodes.

3. We provide experimental results to demonstrate that our joint optimization approach can lead to significant improvements in network design.

In our model, we have used a centralized approach for computing the optimal relay node positions and routing schedule. This is applicable for networks where the relay nodes can be positioned accurately and nodes are mostly stationary after deployment. A centralized approach has been adopted in a number of recent papers [12], [13] and can be used in different application areas, such as habitat monitoring, environment monitoring, building monitoring, or surveillance [22], [23].

## 1.4 Thesis Organization

The remainder of this thesis is organized as follows. In Chapter 2, a brief review of the background knowledge will be provided and chapter 3 presents our LIP formulation for optimal relay node placement and routing in hierarchical sensor networks. We will discuss and analyze various experimental results in Chapter 4. Finally, in Chapter 5, we conclude with a critical summary and provide some future work directions.

# Chapter 2 Background Information

## 2.1 Sensor Nodes and Sensor Networks

Sensor networks combine research advancements from various areas such as sensing, communication and computing (including both hardware and software). Similar to the development of many other technologies, research and development of sensor networks were initially driven by the requirement of military applications. The Distributed Sensor Network (DSN) program, initiated by the Defense Advanced Research Project Agency (DARPA) in the late 70's, symbolizes the modern research on sensor networks.

The recent Technological advances in the field of micro-electro-mechanical system (MEMS) have made the development of tiny, low-powered and multifunctional sensing devices technically and economically feasible [1], [28]. Given such type of sensing devices, modern sensor networks can be constructed by establishing the communication links among the deployed sensor nodes. Although the capability of an individual sensor node is limited, sensor networks are able to perform complex sensing tasks through the collaborative effort of a large number of deployed sensor nodes. Departing from its initial motivation, sensor networks nowadays are employed in a wide range of both civilian and military applications. For example, sensor networks, in civilian domain, can be used to measure the temperature/humidity of a certain region or to monitor the traffic along a highway segment. Scenarios of using sensor networks in military domain include target detection, battle field surveillance and equipment/ammunition monitoring.

### 2.1.1 Sensor Nodes and Deployment

Sensor nodes are underlying building bricks of sensor networks. A typical sensor node, as shown in Figure 2.1 (Simplified from [1]), is usually equipped with a sensing unit for measuring the intentional target (e.g. temperature, humidity, pressure and object-presence/absence etc.). After sensing its vicinity, the raw data generated by a sensing unit is generally in an analogous format which is not computer-readable; therefore, an analog-to-digital convertor (ADC) is normally required to transform the analog data into digital format which, in turn, is further processed by a processing unit. The resultant data from a processing unit is cached into the local memory and when it comes the turn for a sensor node to transmit, the cached data is sent out by the radio communication unit following a pre-established routing path to base station.

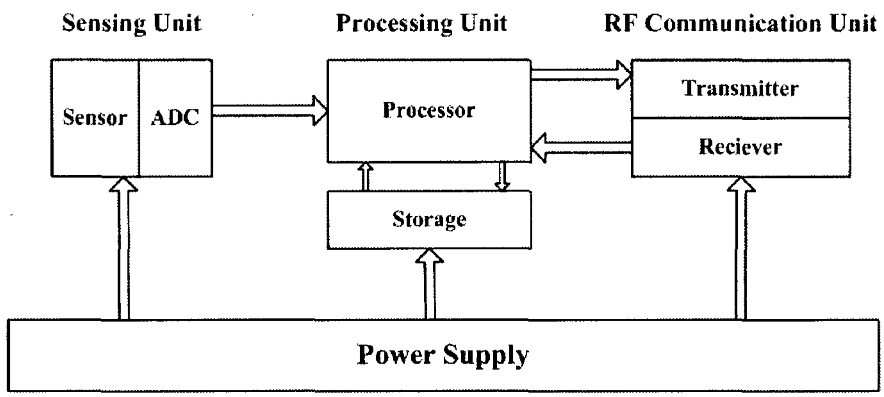


**Figure 2.1 Components of a sensor node**

Sensor nodes in the network are normally deployed inside or very close to the phenomenon, so that the sensing task can be carried out effectively. Positioning sensor nodes within a sensing field can be executed either in a pre-determined fashion or a random scenario. The pre-determined placement of sensor nodes applies to the situation where it is possible to know the actual location of sensor nodes prior to the deployment of the network (e.g. deployment of sensor network in factories or in the bodies of human/

animals). However, in certain cases, especially when working in hostile environment such as battle field or poisoned region, randomly deploying sensor nodes is more practical (e.g. deployment of sensor nodes by dropping them from helicopter/airplane or delivering them in artillery shell or missiles) [1], [28]. The capability of random deployment requires self-organized routing schemes and distributed-network algorithms to be incorporated in sensor networks, which are relatively complex. However, it is the power of random deployment, which makes sensor networks suitable for applying in hostile territories as well as in disaster-relief operations.

### 2.1.2 Architecture Model of Sensor Networks

Sensor networks, according to their internal architecture, can be broadly classified into two categories known as *flat sensor networks* and *hierarchical sensor networks* respectively. In flat sensor networks (e.g. networks as shown in Fig. 1.1), all sensor nodes are assigned the same roles. They are responsible for not only sensing the environment, but also forwarding the sensed the data to the base station.

Unlike flat sensor networks, hierarchical sensor networks (also known as *two-tiered sensor networks*) separate sensing and routing tasks into two different tiers. As shown in Figure 1.2, sensor nodes which are dedicated to the sensing task lie in the lower tier and are grouped into various clusters identified by an assigned cluster head. Each sensor node usually belongs to only one cluster and communicates directly to its cluster head, instead of the base station. All cluster heads, lying in the upper tier, collect sensed data from their respective clusters and form a network among themselves in order to send the collected data to the base station. Compared to flat architecture, hierarchical model achieves advantages in various design objectives: energy conservation, data aggregation, load balancing and connectivity. For example, in hierarchical architecture, sensor nodes in the lower tier are relived from the burden of routing and forwarding, which reduces the energy consumption of these nodes. Because of these mentioned advantages, hierarchical architecture has gained increased popularity in the research and development of sensor networks.

### 2.1.3 Energy Consumption Model of Sensor Nodes

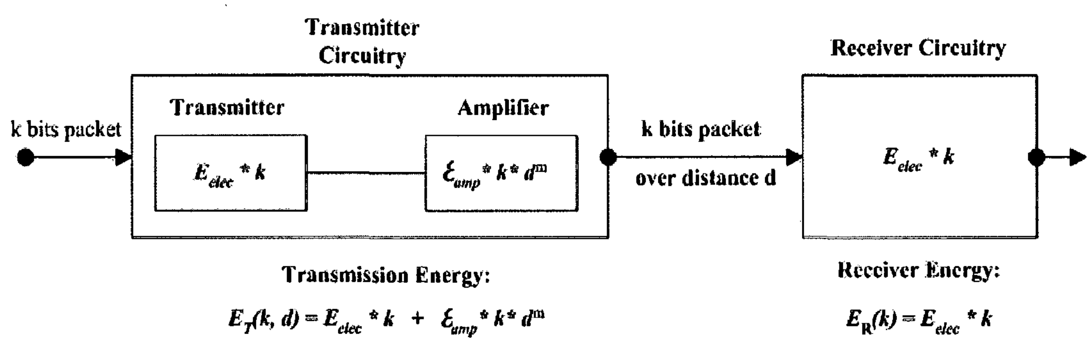


**Figure 2.2 First order radio model**

Energy is considered as one of the most precious resources since it is generally infeasible to recharge/replace batteries within sensor nodes. To manage the energy consumption, it first requires an approach so that the energy dissipated at each sensor node becomes measurable. In the literature, the most commonly employed approach is known as *first-order radio model* (depicted in Figure 2.2 (simplified from [2])), which was proposed by Heinzelman et al. in [2].

According to this model, energy consumed at a sensor node communicating a $k$-bit packet is decomposed into two parts for receiver and transmitter circuitry respectively. The receiver circuitry spends $E_{elec}$ amount of energy in receiving per unit bit of data.

Therefore, for receiving a $k$-bit packet, the total amount of energy dissipated at receiver circuitry is measured as ($E_{elec}$ * $k$) joule. On the other hand, energy dissipated at the transmitter circuitry can be expressed in two terms. The first term considers the amount of energy dissipated by the transmitter circuitry and is calculated by using the same expression, ($E_{elec}$ * $k$) joule, as the receiver circuitry for $k$-bit data. The second term, however, calculates the amount of energy consumed by the amplifier circuitry in order to compensate the signal depression along the transmission channel. The amplifier, in order to transmit 1 bit of data over unit distance, consumes $\varepsilon_{amp}$ joule of energy. The energy loss over distance $d$ is taken care by the term $d^m$, where $m$ is the path loss exponent, $2 \leq m \leq 4$, for free space and for short to medium-range radio communication [3]. Therefore, to transmit $k$ bit data over distance $d$, the total amount of data dissipated at the amplifier is calculated as ($\varepsilon_{amp}$ * $k$ * $d^m$) joule. As shown in Fig. 2.2, by using first-radio model, the total energy dissipation at a sensor node for communicating a $k$ bit packet over distance $d$ can be expressed as the following equation:

$$E_{total} = E_T(k, d) + E_R(k) = 2 * E_{elec} * k + \varepsilon_{amp} * k * d^2$$

$$where\ E_{elec} = 50nJ/bit\ and\ \varepsilon_{amp} = 100pJ/bit/m^2$$ [2]

### 2.1.4 Communication Model of Sensor Networks

All sensed data, in sensor networks, flow from sensor nodes to the base station through inter-communication among deployed sensor/relay nodes. Communication model of sensor networks defines how data packets are transmitted from a source sensor node to the base station. The communication models employed in sensor networks can be broadly classified into the following two groups:

1. *Single-hop transmission model* (DTEM)

2. *Multi-hop data transmission model* (MHDTM)

In the single-hop data transmission model (also called the direct transmission energy model (DTEM)) [2], [26], sensed data is directly transmitted to the base station provided that the base station lies within the transmission range of all sensor/relay nodes. However, in large scale networks, ensuring a base station to be reachable by every node is usually infeasible due to the limited transmission range of sensor/relay nodes. In this case, the multi-hop data transmission model (MHDTM [8], [12], [14], [27]) can be applied. According to multi-hop data transmission model, nodes that cannot reach the base station with a single hop use other nodes as intermediate nodes to relay their data to the base station. Multi-hop data transmission model helps to reduce the transmission distance of the sender and therefore saves the energy dissipation at the sender, which results in an extended lifetime of the network.

Referring to how a source node finds a communication path to the destination (the base station), communication model in sensor networks can also be characterized as *proactive*, *reactive* and *hybrid.* Proactive communication requires all communication paths to be calculated before the actual transmitting action happens. Reactive communication, on the other hand, computes transmission path on demand. In this context, proactive communication can be viewed as a static paradigm which prepares all paths beforehand, while reactive communication operates in a dynamic fashion which generates routing paths upon request of each transmission round. Hybrid communication, as its name suggests, uses a combination of both proactive and reactive communication paradigms [29], [30].

### 2.1.5 Lifetime of Sensor Networks

Lifetime of a sensor network measures the time period during which a sensor network guarantees to fully possess its designated usefulness. In [31], the lifetime time of a sensor network is formally defined as time interval from the inception of the network's operation to the time when the power supplies of a number of *critical nodes* are depleted to such an extent that it results in a *routing hole* [31] within the network, a disconnected network or a network with insufficient coverage. In sensor networks based on flat architecture, network lifetime, varying from application to application, can be taken as the time when the first node, last node or more generally a certain percentage of nodes completely runs out of energy.

However, in hierarchical sensor networks, energy depletion of a sensor node and a cluster head has different impacts on the lifetime of the network and needs to be considered differently. Hierarchical sensor networks usually contain a large number of sensor nodes which are densely deployed in the sensing field in order to carry out the designated sensing tasks accurately. In such context, the lack of sensing by a "dead" sensor node will be compensated by one or more adjacent alive sensor nodes. On the other hand, if a cluster head runs of energy and becomes dead, all sensor nodes communicating to this cluster head are inaccessible from other part of the network. If a cluster head also appears in multi-hop routing paths of other cluster heads, complete energy depletion of this node has even more severe impact on the lifetime of the network. In this case, the set of inaccessible sensor nodes includes not only its own cluster but also other clusters whose cluster heads use this node as intermediate node in their multi-hop communication paths. Therefore, energy depletion of cluster heads plays a more important role in the lifetime of hierarchical sensor networks. In [3], Pan et al. have measured the lifetime of hierarchical sensor networks in three different ways which are summarized as follows:

1. *N-of-N lifetime,* the network lifetime expires as soon as the first cluster head dies.
2. *K-of-N lifetime,* the network survives as long as *K* cluster heads are still alive.
3. *M-in-K-of-N lifetime,* the network survives if a minimum of *m* pre-specified cluster heads and overall a minimum of *K* cluster heads are still alive.

## 2.2 Relay Nodes in Sensor Networks

In the past few years, a number of researches have focused on deploying relay nodes in sensor networks. The main objectives of applying relay nodes to sensor networks can be summarized as follows:

- Extending the network lifetime
- Reduction of transmission range
- Energy-efficient data gathering
- Balanced data gathering
- Improved connectivity and fault tolerance

As regular sensor nodes, relay nodes in sensor networks are also battery-operated devices with wireless communication capabilities and hence are energy restricted. However, relay nodes in sensor networks only take care of relaying sensed data generated by other nodes, without sensing the environment. For example, a typical relay node may receive incoming data packets from multiple sensor nodes, generate outgoing packets and transmit them to the next relay node or the base station. Relay nodes with different characteristics can be used in both flat and hierarchical sensor networks.

### 2.2.1 Relay Nodes in Flat Sensor Networks

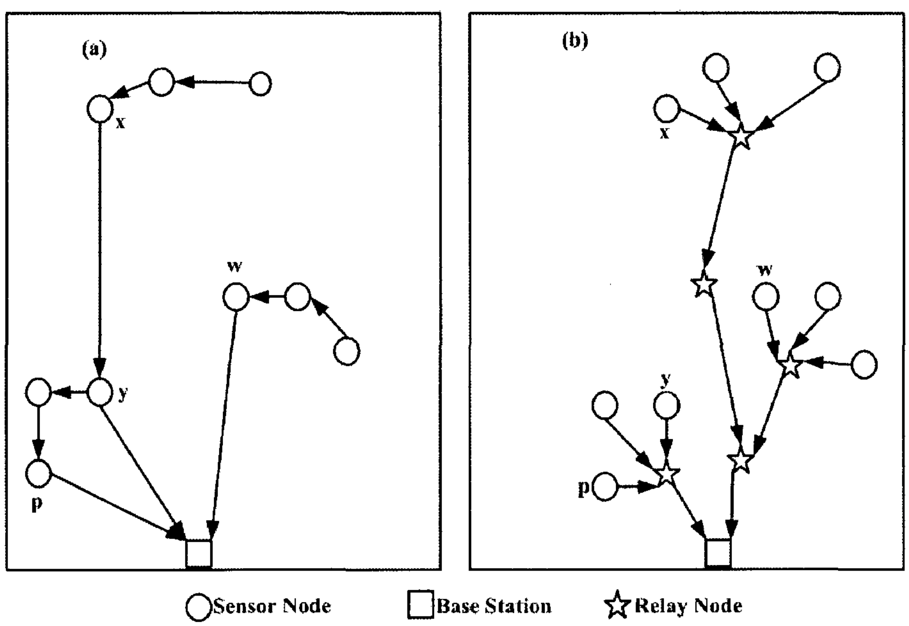


**Figure 2.3 Using relay nodes in flat sensor networks**

Figure 2.3 (Re-depicted from [35]) shows a basic example of employing relay nodes in a flat sensor network. Fig. 2.3 (a) presents a general flat sensor network without using any 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 range of nodes such as *x, y, p, w*, giving a network with an increased lifetime.

In literature, deploying relay node in flat sensor networks was first proposed by Cheng et al. [4] in 2001 when they studied the problem of "maintaining connectivity with minimum-per-node transmission power in wireless sensor networks" [4]. They have formulated this problem based on a network optimization problem called *Steiner Minimum Tree with Minimum Number of Steiner Points* [32] and proposed two optimization algorithms to solve the connectivity problem in flat sensor networks. Through performance study by simulation, they have claimed that introducing a small number of relay nodes helps to reduce the total number of power consumption while still maintains the global network connectivity. Dasgupta et al. in [33] considered flat sensor networks consisting of sensor nodes and relay nodes, where all nodes are of equal capabilities but can be assigned the role of either a relay node or a sensor node. They focused on the placement of nodes within the network and assigning their roles in a way that the lifetime of sensor networks is maximized while the coverage of the entire region is ensured. The algorithm proposed by them is named as *Sensor Placement and Role Assignment for Energy-efficient Information Gathering* (SPRING). Given the placement of the base station and the deployed nodes as well as their initial role assignment, SPRING is able to find the location along with their assigned roles so that the network lifetime is maximized while ensuring the coverage of entire sensing field.

Falck et al. in [5] introduced relay nodes in flat sensor networks with multi-hop communication in order to achieve balanced data gathering against sufficient coverage of the monitored area. They have studied the effect of deploying a small number of relay nodes within the network and proposed an approximation algorithm for their placement. The presented simulation results demonstrated that employing a small number of relay nodes in flat sensor networks can lead to a significant improvement in the balanced data-gathering, while retaining sufficient coverage of the sensing field. The proposed linear programming (LP) solution also improves the work done in [34], [36] in which non-linear solutions are used.

### 2.2.2 Relay Nodes in Hierarchical Sensor Networks

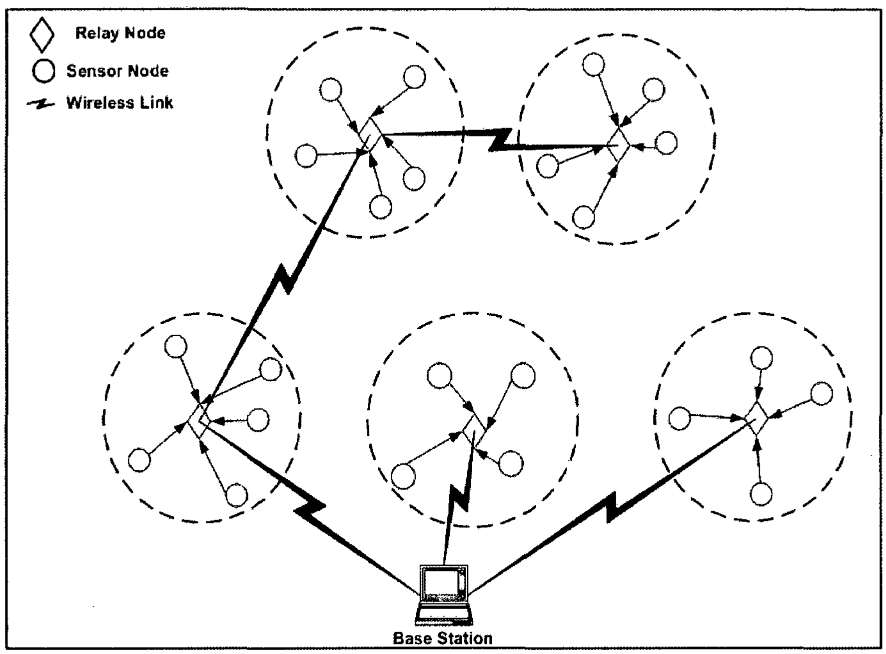


**Figure 2.4 Using relay nodes in hierarchical sensor networks**

Relay nodes, in hierarchical sensor networks, usually serve as cluster heads (Figure 2.4) to achieve energy-efficient data gathering, extended network lifetime and balanced data loading. Fig. 2.4 presents a typical usage of relay nodes in hierarchical sensor networks.

As sensor networks are usually deployed to measure a target parameter, it is highly possible that sensed data within a cluster (shown as a dash circle in Fig. 2.4) involves certain degree of data redundancy. Employing relay nodes as cluster head introduces a data-gathering pattern such that data redundancy within a cluster can be pre-removed before sending to the base station, which contributes to energy conservation as it saves the energy spent in transmitting the redundant data volume and therefore results in an extended network lifetime. Using relay nodes as cluster heads could also lead to an increased network bandwidth usage due to data volume reduction.

In literature, the employment of relay nodes in hierarchical sensor networks was first proposed in 2003, in two different publications [6] and [3]. Gupta et al. in [6] focused on the load balancing problem of sensor networks where the energy constrained sensor nodes are not uniformly distributed. They solved this problem by introducing the notion of deploying relatively-less energy-constrained relay nodes (e.g. *gateway nodes* named by them in [6]). In their proposed model, the deployed relay nodes group sensor nodes into distinct clusters and each relay node acts as cluster head of its corresponding cluster. Each sensor node, on the other hand, belongs to only one cluster and communicates directly to the cluster head. Relay nodes serving as cluster heads collect data from cluster members, perform data aggregation and relay the resultant data packets directly/through other relay nodes to the base station. With respect to such a model, they proposed an optimization heuristic algorithm that clusters the sensor nodes based on the deployed relay nodes and balances the data flow among the introduced relay nodes.

Considering the similar network model as in [6], Pan et al. in [3], introducing relay nodes (named as *Application Nodes (AN)* by them in [3]), have attempted to maximize the topological lifetime of the network by strategically placing the base station (BS) and optimizing inter-application node (AN) relaying. Under the assumption that the locations of BSs are relatively flexible, authors in [3] have proposed computational-geometry-based algorithms which find the optimal locations of BSs so that the topological lifetime of the network is maximized. Hou et al. in [12] spent their research effort in prolonging the lifetime of hierarchical, cluster-based sensor networks in which the upper tier contains *Aggregation and Forwarding Nodes* (AFNs) as well as relay nodes. They formulated this problem as *Energy Provisioning Relay Node Placement* (EP-RNP) and proposed a polynomial-time heuristic algorithm known as "SPINDS" which attempts to provision additional energy to the existing nodes and deploy AFNs and RNs to mitigate the geometric deficiency of the network so that the network's lifetime if extended.

Most researches discussed so far concentrated on applying relay nodes to achieve performance improvement in hierarchical sensor networks with the assumption that relay nodes have been deployed within the sensing field. Researches on the placement and coverage of relay nodes, on the other hand, have focused on how to effectively deploy relay nodes within hierarchical sensor networks. As mentioned earlier, relay nodes in hierarchical sensor networks usually lie in the upper tier and serve as cluster heads while equal capability sensor nodes are randomly deployed, the placement of relay nodes has to ensure that every sensor node is covered by at least one relay node. A sensor node, in real application, is considered as covered by a relay node if there is a relay node positioned with the transmission range of that sensor node. Relay nodes acting as cluster heads, on the other hand, are responsible for delivering/relaying data packets to the base station through either a single or multi-hop routing path. Therefore, placement of relay nodes also needs to ensure that the upper tier network (including relay nodes and the base station) is connected so that for each placed relay node, there is a routing path along which a relay node is able to convey data packets to the base station.

Placing relay nodes in hierarchical sensor networks has been addressed in [15] and [35] while the complexity of relay nodes placement problem has been investigated in [8] and [20]. Suomela in [15] and [35] has examined the complexity of relay node placement problem with respect to various optimization problems in hierarchical sensor networks, and shown that all these problems are NP-hard, in some cases even the approximations are NP-hard.

Tang et al. in [8] have concentrated on the problem of placing minimum number of relay nodes such that each sensor node is able to communicate with at least one relay node and relay nodes themselves are connected. They formulated this problem as *Connected Relay Node Single Cover* (CRNSC) problem. Introducing the concept of *P-Position* (Potential Position), they have proposed approximation algorithms of polynomial time complexity to solve the CRNSC problem. To incorporate the fault-tolerant capability, Tang et al. extended the CRNSC problem to a 2-Connected Relay Node Double Cover (2CRNDC) problem which is referred as finding minimum number of relay nodes in a way that each sensor node can communicated with at least two relay nodes and the network of placed relay nodes are 2-connected. The proposed polynomial time approximation algorithm to 2CRNDC problem is derived from solution of CRNSC by adding some redundant relay nodes to the solution set of CRNSC problem.

Bari et al. in [20] focused on a more general scenario with the objective to find minimum number of relay nodes along with their locations, such that each sensor node can communicate with at least *k* number of relay nodes and the relay nodes network is *r* connected. They proposed an *Integer Liner Programming* (ILP) formulation which takes a set of candidate relay node locations as well as a set of deployed sensor nodes locations and produces the set of selected relay node locations. Facing the infinite number of possible relay node locations in the sensing field, they presented a grid-based approach for preparing the potential set of relay nodes positions. In the grid-based approach, the entire sensing field is divided into a set of cells by the latitude and longitude lines. The centers of each cell are picked to initialize the set of input relay nodes. Given this initial set of relay nodes and the set of sensor locations, their ILP formulation selects, from the input set of relay nodes, the minimum number of relay nodes such that each sensor node can communicate with at least *k* number of relay nodes and the selected relay nodes are *r* connected.

## 2.3 Routing in Hierarchical Sensor Networks

Compared to flat sensor networks, hierarchical sensor networks have gained popularity in recent years, due to their ability to facilitate energy conservation, load-balanced data gathering, fault-tolerance as well as increased network coverage and connectivity. Routing in hierarchical sensor networks is considered as a challenging task [30] because of the inherent characteristics which distinguish hierarchical sensor networks for other kinds of wireless networks, e.g. mobile ad hoc networks or cellular networks. The major characteristics that pose difficulties to routing in hierarchical sensor networks are summarized as following [30]:

- The number of sensor nodes deployed in sensor network may be very large.

Considering high overhead for maintaining IDs of such a large number of nodes, traditional IP-based protocols may not be directly applicable in sensor networks.

- Sensor nodes are constrained by resources, e.g. energy, processing, and storage capacities, therefore, resource management is very important in sensor networks.
- Once deployed, most of sensor nodes are usually stationary, but some nodes may be allowed to move around, depending on the requirements of the application.
- The requirements for the design of sensor networks may change with application.
- Data collection in sensor networks is usually location-based, so that position awareness of sensor nodes is important.
- As sensor networks consist of large number of sensor nodes deployed to measure a common target parameter, it is highly possible to have data redundancy, which should be taken into consideration by the routing mechanism to ensure energy efficiency and bandwidth utilization.

On the other hand, as the dominant amount of energy is consumed in data transmission, routing strategy plays a significant role in conserving energy and needs to be taken into careful consideration. Presented with the above challenges, lots of research efforts, in the past few years, have been spent in designing routing mechanisms that are suitable for hierarchical sensor networks. In the literature, various hierarchical-sensor-network-oriented routing tactics have been proposed in [2], [37] and [38] to exploit the architectural advantages and perform energy-efficient routing in hierarchical sensor networks.

Heinzelman et al. in [2] proposed a self-organizing, adaptive routing protocol called *Low-Energy Adaptive Clustering Hierarchy* (LEACH) with the objective to minimize the total energy consumption of cluster-based hierarchical sensor networks. In LEACH, the distributed sensor nodes are grouped into a set of local clusters and one node is assigned the role as cluster head which is responsible for collecting, aggregating sensed data within its own cluster as well as transmitting the resultant data packets to the base station. Therefore, cluster heads in LEACH consume energy at a much higher rate than regular sensor nodes. However, the proposed LEACH protocol attempts to randomly rotate the role of cluster head within a cluster to ensure the energy dissipation is evenly distributed among all nodes within the sensor networks.

Manjeshwar et al. in [37] presented a *Threshold Sensitive Energy Efficient Sensor Network* (TEEN) protocol targeting to maximize the lifetime of the cluster-based, hierarchical sensor networks. Unlike LEACH in which data collection and transmission are performed at predetermined time intervals, the underlying idea of TEEN relies on the fact that cluster heads are required to transmit only when there are significant changes in the monitored environment. TEEN employs two threshold values, *hard threshold* and *soft threshold* respectively, to determine whether or not to execute data collection and transmission at cluster heads. Hard threshold is the absolute value of the sensed attribute that triggers a cluster head to transmit while a soft threshold is a small change in the attribute value that triggers a cluster head to transmit.

Ossama et al. in [38], attempting to maximize the life of cluster-based, hierarchical sensor networks, have improved LEACH [2] and proposed a routing protocol named as *Hybrid Energy-Efficient Distributed Clustering* (HEED). In HEED, cluster heads are periodically and probabilistically selected based on the residual energy of each node. They also included a secondary clustering parameter (e.g. node proximity to its neighbors, the node degree) in order to reduce the intra-cluster communication cost therefore increased the energy efficiency and further prolongs the lifetime of networks. When selecting the cluster head, a sensor node may refer to this secondary parameter so that the communication cost is minimized.

## 2.4 Fault-tolerance in Hierarchical Sensor Networks

Nodes in sensor networks are prone to failure because of running out of batteries, physical damages, and malicious attacks. In certain circumstance, there exists infrequent link failure in wireless communication due to the environmental interference. Fault-tolerance, in sensor networks, refers to the ability of surviving from such kinds of node or link failure. In other words, a fault-tolerant sensor network, even in the presence of node or link failure, should still sustain its designated functionality without interruption.

A traditional approach to enable fault-tolerance in sensor networks is to construct node or link disjoint paths between source and destination. Keeping multiple routing paths among all pairs of source and destination ensures the connectivity of the network. For instance, if some links or nodes fail, the alternative path can take part in the data routing and the network still remains connected. In general, a sensor network should be at lease 2-connected (i.e. for each pair of source and destination, at least two node disjoint paths are prepared). However, base on the criticality of the fault in an application, a sensor network may require to be $k$ connected where $k \geq 2$.

As mentioned in the previous section, different architectural models (flat and hierarchical sensor networks respectively) have been proposed for sensor networks. Therefore, fault-tolerance needs to be treated differently according to the characteristics of each model. In flat architecture, fault-tolerance focuses on establishing nodes/links disjoint paths between all sensor nodes and the base station so that even in case of nodes/links failure, the alternative paths are available and can be used to deliver data packets.

Under hierarchical model, it is highly possible that the lacking of sensing due to a single sensor node failure will be compensated by the other sensor nodes within the same cluster. However, failure of a cluster head has much more severe effect than the failure of a sensor node in that it makes not only all underlying sensor nodes covered by the failed cluster head become inaccessible, but also leads to data loss from other cluster heads which use the failed cluster head as intermediate node to relay data to the base station, if a multi-hop communication model is employed. Therefore, fault-tolerance in hierarchical sensor networks needs to pay more attention to deal with the failure of cluster heads.

In recent years, the issue of fault-tolerance in hierarchical sensor networks has been studied in various papers including [7], [16] and [17]. Gupta et al. in [7] proposed a solution to deal with the failure of relay nodes (cluster heads) in hierarchical sensor networks. Their solution focused on recovering the cluster members (sensor nodes) from a failed relay node. In their approach, the system periodically queries the status of relay nodes so that the system is able to detect the failure of any relay node. In the presence of relay node failure, their scheme will re-assign all affected cluster members to a backup relay node which is created during the clustering phase. Therefore, it eliminates the necessity of a full-scale re-clustering of the entire network.

Hao et al. in [16] have focused on enhancing the fault-tolerance capability of hierarchical sensor networks through the placement of relay nodes. They formulated their fault-tolerant scheme as placing the minimum number of relay nodes such that each sensor node is connected to at least 2 relay nodes and the upper tier relay nodes network is 2 connected. Therefore, even in case of a relay node failure, the affected sensor nodes can connect to at least one backup relay node and the remaining relay nodes network is at least one connected. They solved this relay nodes placement problem by proposing a polynomial-time approximation algorithm.

Liu et al. in [17] also focused on placing optimal number of relay nodes so that the hierarchical sensor network becomes fault-tolerant. Unlike the solution proposed in [16], they solved this problem through a two-phase approach. The goal of the first step is to ensure that the network becomes connected and the second stop focused on double connecting the relay nodes network by adding redundant relay nodes to the solution generated in the first step. They have presented an approximation algorithm to solve the problem addressed in the first step and two approximation algorithms which solves the problem addressed in the second step.

# Chapter 3 Network Design with Performance Guarantees

## 3.1 Problem Statement

Due to the ability to facilitate optimization of network lifetime, load-balanced data gathering, fault-tolerance as well as increased network coverage and connectivity, hierarchical sensor networks have gained popularity in recent years. Using higher-powered relay nodes as cluster heads can lead to further improvements in network performance and has been addressed in various researches. There are two important problems need to be taken into careful consideration when attempting to employ higher-powered relay nodes as cluster heads in hierarchical sensor networks:

i. Relay node placement strategies

ii. Routing strategies among the deployed relay nodes

Although significant amount of research effort has been spent in relay node placement and routing problems, current researches in this field separate the placement and routing of relay nodes into two steps. Positioning relay nodes is executed in the first step and then a particular routing scheme is employed based on locations of relay nodes. While focusing on the relay nodes placement, the typical consideration is coverage and connectivity, and does not take into account of the energy dissipation of relay nodes which requires knowledge of the routing schemes.

Unlike previous approaches, we, in this thesis, focus on the joint optimization of both placement and routing of relay nodes and define our problem as:

*Find minimum number of relay nodes, along with their locations, and a suitable* *communication strategy such that:*

i. *All sensor nodes are covered by at least $k_s$ relay nodes.*

ii. *The upper tier relay node network is at least $k_r$ connected.*

iii. *The network has a guaranteed lifetime.*

To solve the above defined problem, we have proposed an approach by using an ILP formulation, which is presented in section 3.4. Our proposed ILP formulation approach not only designs a network that meets the requirements of coverage and connectivity, but also finds a routing schedule which provides guarantees of network's lifetime. In addition to the ILP formulation, we also present an intersection based approach for determining the potential positions of relay nodes which are used as input for our ILP formulation.

## 3.2 Network Model

For our model, we consider a hierarchical (or two-tiered) wireless sensor network, where the lower tier consists of $n$ sensor nodes, randomly distributed within the sensing area. Our objective is to determine the minimum number and positions of relay nodes (cluster heads) to form the upper tier network, with a specified degree of redundancy. We also determine a suitable routing strategy such that the energy dissipation of the relay nodes is reduced as much as possible. A sensor node $i$ is said to be *covered* by a relay node $r_j$ at location $j$, if $i$ can transmit its data directly to $r_j$. Our proposed formulation designs the upper tier relay node network, such that each sensor node is covered by at least $k_s$ relay node(s), where $k_s$ = 1, 2, 3, …, and each relay node can forward its data to $k_r$, $k_r$ = 1, 2, 3, …, other relay node(s) (or directly to the base station). This means that each sensor node can still transmit its data to at least one rely node, even if up to $k_s - 1$ relay nodes fail. Similarly, it guarantees that each relay node has a viable path to the base station, even if up to $k_r - 1$ relay nodes fail. For proper functioning of network it is required that, at a minimum, $k_s = 1$, i.e. each sensor node is capable of communicating with at least one relay node and $k_r = 1$, i.e. the upper tier relay node network is connected.

We assume that the positions of the sensor nodes are known beforehand, or can be determined (e.g. using GPS), and that the relay nodes can be placed at the locations determined by our placement strategy. The ILP formulation proposed here assumes that a set $\boldsymbol{R}$ of potential locations for the relay nodes is given as input. In section 3.6, we describe an intersection based approach for generating the $\boldsymbol{R}$ potential locations. However, our formulation does not depend on how $\boldsymbol{R}$ is generated, and other approaches such as a grid based approach [20] or that given in [8] can easily be used.

The dominant factor in power consumption in sensor networks is the power needed for communication. In the first order radio model [2], receiver (transmitter) circuitry consumes $\alpha_1$ *nJ/bit* ($\alpha_2$ *nJ/bit*) of energy. The total energy to receive $b$ bits is given by, $E_{Rx}(b) = \alpha_1 b$ while the total energy needed to transmit $b$ bits over a distance $d$ is given by $E_{Tx}(b) = \alpha_2 b + \beta b d^q$, where $q$ is the path lose exponent, $2 \leq q \leq 4$ [3] and $\beta$ is the amplifier energy to transmit unit bit of data over unit distance. In our experiments, we have used $\alpha_1 = \alpha_2 = 50nJ/bit$, $\beta = 100pJ/bit/m^2$ and the path-loss exponent, $q = 2$.

We assume that data gathering is proactive, i.e., data are collected and forwarded to the base station periodically, following a schedule. We have called one period of proactive data gathering (starting from sensing until all data reach the base station) as one "round" [14]. We define the lifetime of a hierarchical sensor network as the number of rounds that the network can sustain, from the deployment of the network, to the time the data from any cluster head fails to reach the base station. The sleep/wake scheduling and the underlying synchronization protocols are handled separately by a state-of-the-art method in the MAC layer, such as those proposed in [21], [25].

## 3.3 Notation Used

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

- $n$: The total number of sensor nodes, with each sensor node having a unique index $i$, $1 \leq i \leq n$.
- $m$: The total number of possible positions of relay nodes, each position having a unique index $j$, $n+1 \leq j \leq n+m$.
- $r_j$: The relay node at location $j$, $n+1 \leq j \leq n+m$.
- $n+m+1$: The index of the base station.
- $r_{max}$: The transmission range of each sensor node.
- $d_{max}$: The transmission range of each relay node.
- $d_{i,j}$: The Euclidean distance from node $i$ to node $j$.
- $k_s$: The minimum number of relay nodes covering each sensor node.
- $k_r$: Desired connectivity of the relay nodes network.
- $\alpha_2$ $(\alpha_1)$: Energy coefficient for transmission (reception).

- $\beta$: Energy coefficient for amplifier.

- $D$: A large constant, $D \geq \sum_{i=1}^{n} bi$, $1 \leq i \leq n$

- $b_i$: Number of bits generated by sensor node $i$.

- $e_{max}$: Maximum allowable energy dissipation (per round) of a relay node.

We also define the following variables:

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

$$Z_{i,j} = \begin{cases} 1 \text{ if the sensor node } i \text{ can transmit to the relay node } j \\ 0 \text{ otherwise} \end{cases}$$

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

$$X_{i,j} = \begin{cases} 1 \text{ if the sensor node } i \text{ selects relay node } j \text{ as its cluster head} \\ 0 \text{ otherwise} \end{cases}$$

- $Y_j$: Binary variable defined as follows:

$$Y_j = \begin{cases} 1 \text{ if the relay node at location } j \text{ is included in the upper tier} \\ 0 \text{ otherwise} \end{cases}$$

- $C_j$: Continuous variable indicating the number of other relay node(s) that may be used by relay node $r_j$ to forward data towards the base station.

- $T_j$: Continuous variable indicating the number of bits transmitted by relay node $j$.

- $G_j$: Continuous variable indicating the amount of energy dissipated by the amplifier in relay node $j$ to send its data to the next node in its path to the base station.

- $R_j$: Continuous variable indicating the number of bits received by relay node $j$ from other relay nodes.

- $E_j$: Continuous variable indicating the total energy spent per round by the relay node $j$.

- $w_j$: Continuous variable indicating the total number of bits generated by all sensor nodes in cluster $j$.

- $f_{j,k}$: Continuous variable indicating the amount of flow from a relay node $j$ to node $k$ (may be another relay node or the BS)

## 3.4 ILP Formulation for Integrated Placement and Routing

In this section, we propose a formulation that guarantees the coverage of each sensor node by at least $k_s$, $k_s$ = 1, 2, ..., relay node(s) and relay nodes network that is $k_r$-*connected* ($k_r$ = 1, 2, …). The objective function is to minimize the number of relay nodes while maintaining a desired lifetime of the network. By setting the appropriate values for

$k_s$ and $k_r$, this formulation can also ensure fault tolerance. We note that this formulation may select relay nodes which are not acting as cluster heads for any sensor nodes. Such nodes are used to maintain the required degree of connectivity and/or to achieve the desired network lifetime, and are included in the topology only if necessary.

$$\textbf{Minimize} \sum_{j=n+1}^{n+m} Y_j \qquad (1)$$

Subject to:

a) A sensor node $i$ can transmit to a relay node $j$, only if the distance between $i$ and $j$ is less than the transmission range $r_{max}$ of the sensor node $i$.

$$Z_{i,j} \bullet d_{i,j} \le r_{\max} \qquad \begin{array}{l} \forall i,\ 1 \le i \le n, \\ \forall j,\ n+1 \le j \le n+m \end{array} \qquad (2)$$

b) A relay node $j$ can transmit to a relay node $k$, only if the distance between $j$ and $k$ is less than the transmission range $d_{max}$ of the relay node $j$:

$$f_{j,k} = 0 \qquad \forall j,k:\ d_{j,k} > d_{\max} \qquad (3)$$

c) The relay node at location $j$ is included in the upper tier network, if it is selected as a potential cluster head by at least one sensor nodes $i$.

$$Y_j \ge Z_{i,j} \qquad \begin{array}{l} \forall i,\ 1 \le i \le n, \\ \forall j,\ n+1 \le j \le n+m \end{array} \qquad (4)$$

d) A sensor node must be covered by at least $k_s$ relay nodes.

$$\sum_{j=n+1}^{n+m} Z_{i,j} \ge k_s \qquad \forall i,\ 1 \le i \le n \qquad (5)$$

e) A sensor node $i$ transmits to a relay node $j$, only if the relay node $j$ is selected by sensor node $i$ as its cluster head.

$$X_{i,j} \le Z_{i,j} \qquad \begin{array}{l} \forall i,\ 1 \le i \le n, \\ \forall j,\ n+1 \le j \le n+m \end{array} \tag{6}$$

f) A sensor node transmits sensed data to exactly one relay node.

$$\sum_{j=n+1}^{n+m} X_{i,j} = 1 \qquad \forall i,\ i \le 1 \le n \tag{7}$$

g) Calculate the number of the relay node that the relay node $j$ can use to route data towards the base station.

$$C_j = \sum_{w:(d_{j,w} \le d_{\max})\ AND\ (d_{w,\ n+m+1} < d_{j,n+m+1})} Y_w \tag{8}$$

Constraint (8) has to be applied for all $j$, $n \le j \le n+m$.

h) If the base station lies outside of the transmission range of relay node $r_j$, there must be $k_r$ other relay nodes where $r_j$ can forward its data.

$$C_j \ge k_r \bullet Y_j \qquad \forall j:\ d_{j,n+m+1} \ge d_{\max} \tag{9}$$

Constraints (8) and (9) together determine the connectivity of the relay node network.

i) Calculate the total number of bits generated in the cluster $j$.

$$w_j = \sum_i b_i \bullet X_{i,j} \qquad \begin{array}{l} \forall i,\ 1 \le i \le n, \\ \forall j,\ n+1 \le j \le n+m \end{array} \tag{10}$$

j) Flow constraint.

$$\sum_k f_{j,k} - \sum_k f_{k,j} = w_j \tag{11}$$

k) Calculate the total number of bits transmitted by the relay node $j$.

$$T_j = \sum_j f_{j,k} \qquad \forall j, k \neq n+m+1 \tag{12}$$

l) Calculate the amplifier energy dissipated by relay node $j$ to transmit to the next node.

$$G_j = \beta \sum_k f_{j,k} \bullet (d_{j,k})^q \qquad \forall j, k \neq n+m+1 \tag{13}$$

m) Calculate the number of bits received by node $j$ from other relay node(s).

$$R_j = \sum_k f_{k,j} \qquad \forall j,\ \ n < j \leq n+m+1 \tag{14}$$

n) Base station does not transmit.

$$f_{n+m+1,k} = 0 \qquad \forall k,\ \ 1 \leq k \leq n+m+1 \tag{15}$$

o) A link, from relay node $j$ to a relay node at location $k$, can have non-zero data flow only if the relay node $k$ is selected to be in the upper tier.

$$f_{j,k} \leq D \bullet Y_k \qquad \forall k, j,\ \ j \neq n+m+1 \tag{16}$$

p) Calculate the energy dissipated at relay node $j$.

$$\alpha_1 (R_j + w_j) + \alpha_2 T_j + G_j = E_j \qquad \forall j:\ \ j \neq n+m+1 \tag{17}$$

q) Constraint for maximum energy dissipation.

$$E_j \leq e_{\max} \qquad \forall j: \ j \neq n+m+1 \qquad (18)$$

## 3.5 Justification of the ILP Equations

Equation (1) is the objective function of the ILP formulation that minimizes the total number of selected relay nodes which form the upper tier relay nodes network. The minimization of the number of relay nodes is obtained after ensuring the required coverage of sensor nodes and the connectivity requirement of elected relay nodes, as well as ensuring the desired network lifetime.

a. Constraint (2) enforces the restriction that a sensor node can only transmit to a relay node, if the relay node is within the transmission range of that sensor node.

b. Constraint (3) specifies that if the distance between two different relay nodes exceeds the transmission of the relay node, the amount of flow between them is 0. In other words, constraint (3) enforces the restriction that a relay node can only transmit to another relay node (or to the base station) if the destination node is within the transmission range of the relay node transmitting data.

c. Constraint (4) ensure that if the relay node $r_j$ at location $j$ is chosen as a potential cluster head by one or more sensor nodes, then $r_j$ must be included in the set of relay nodes selected to form the upper tier network. If a relay node $r_j$ is not chosen as a potential cluster head for any sensor node, normally it should not be selected (unless it is needed to maintain required connectivity). This is not specially enforced by any constraint, but is taken care by the objective function, which will set $Y_j = 0$, if this does not violate any the other constraints.

d. Constraint (5) requires that each sensor node to be covered by at least $k_s$ relay nodes other than one. The actual value of $k_s$, can be chosen based on the intended application. For most applications $k_s = 2$ or 3 should be suffice. Under fault-free conditions, each sensor node will select one relay node (from the $k_s$ relay nodes it is associated with) to send its data. If that node fails, it can switch to another backup cluster head from the remaining $(k_s - 1)$ nodes.

e. Following constraint (5), constraint (6) and (7) jointly enforce that a sensor node transmits its sensed data to only one particular cluster head, even though a sensor node should maintain a certain level of redundant cluster heads (e.g. specified by the value of $k_s$ in constraint (5)) for fault tolerance purpose.

f. Constraint (8) and (9) ensure the connectivity of the upper tier relay nodes network, according to the pre-specified connectivity requirement ($k_r$). More specifically, constraint (8) specifies the approach to calculate the total number of other relay nodes that can be used by a relay node $j$ to forward data towards the base station ($C_j$). Given a relay node $j$, the value of $C_j$ is obtained by summing over all other selected relay nodes which are within the transmission range of relay node $j$ and closer to the base station than the given relay node $j$. Constraints (9) further states that for a selected relay node $j$ which cannot reach the base station with a single hop ($d_{j,\, n+m+1} \geq d_{max}$), there should be at least $k_r$ other relay nodes available for relay node $j$ to forward its data towards the base station. Therefore, constraint (8) and (9) jointly guarantee that there is at least one via path for each relay node to the base station even up to $k_r - 1$ relay nodes failed. Similar to $k_s$, the actual value of $k_r$ is set up depending on the intended application.

g. Constraint (10) calculates the total number of bits ($w_j$) generated in cluster $j$, by summing the data transmitted to it from all the sensor nodes belonging to the cluster $j$.

h. Constraint (11) corresponds to the standard constraints [41], and states that the total amount of outgoing data from relay node $j$ ($\sum_k f_{j,k}$) is equal to the total incoming data from other relay nodes ($\sum_k f_{k,j}$) plus the data generated within cluster $j$ ($w_j$).

i. Constraint (12) calculates the total number of bits ($T_j$) transmitted by the relay node $j$, by summing the data transmitted over all outgoing links from node $j$.

j. Constraint (13) calculates the amplifier energy ($G_j$) dissipated at relay node $j$ by summing the amplifier energy required along each link.

k. Constraint (14) specifies the total number of bits received at relay node $j$ form other relay node(s), by summing the data flowing along all incoming links.

l. Constraint (15) specifies that the base station indexed as $n+m+1$ does not transmit to any other node because base station serves as data repository which only receives data.

m. Constraint (16) specifies that data can be sent from relay node $j$ to relay node $k$ through link ($j$, $k$), only if relay node $k$ is also selected to be in the upper tier relay nodes network. For example, if $Y_k = 0$, constraint (16) will force $f_{j,k} = 0$. The constant $D$ is needed since the value of $f_{j,k}$ may be greater than 1. The value of $D$ should be large enough to allow the maximum possible data flow on link ($j$, $k$).

We have set $D \geq \sum_{i=1}^{n} bi$, $1 \leq i \leq n$.

n. Constraint (17) computes the total energy $E_j$ dissipated by a relay node $r_j$, in one round of data gathering. The energy dissipated by the relay node $j$ has three components:

   i. the receiver energy $\alpha_1(R_j + w_j)$,

   ii. the transmitter electronics energy $\alpha_2 T_j$, and

   iii. the transmitter amplifier energy $G_j$

o. Constraint (18) ensures that the total energy dissipated by a relay node cannot exceed $e_{max}$, which specifies the maximal per-round-energy-dissipation of a selected relay node and is supplied as input data to the formulation.

*Theorem 1: Constraint (8) and (9) guarantee that the relay nodes network can survive $k_r$ -1 faults.*

*Proof*: For each relay node $r_j$ in the upper tier network, constraint (8) computes the number of relay nodes that are:

i. within the transmission range of $r_j$, and

ii. closer to the base station than $r_j$.

These are the nodes that may be used by $r_j$ to forward its data to the base station, if the base station is not within its transmission range. Constraint (9) ensures that there are at least $k_r$ such nodes, for any selected relay node which cannot transmit to the base station directly. This means that even if up to $k_r - 1$ relay nodes failure, there will still be at least one surviving node within the transmission range of $r_j$, which is closer to the base station than $r_j$. Since this is true for all relay nodes, constraint (9) ensures that there will be a viable path from each relay node to the base station, even in the presence of $k_r - 1$ relay node failures. This guarantees that the relay nodes network has the desired connectivity.

## 3.6 Finding Potential Locations of Relay Nodes

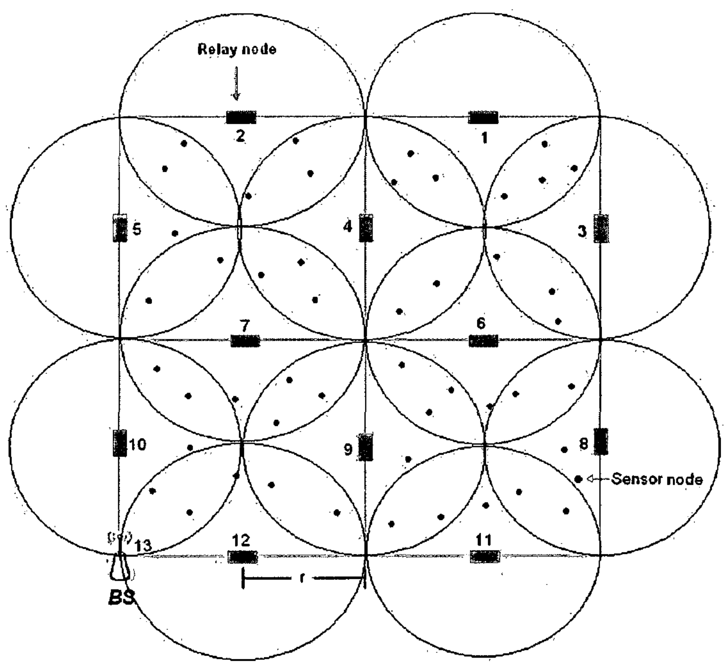


**Figure 3.1 Grid based placement of relay nodes**

In the previous section, we have presented an ILP formulation that optimally selected the positions of the relay nodes (from a set of *potential* positions) and determines a routing schedule that meets certain criteria such as coverage, connectivity and energy requirements. Experimental results (in Chapter 4) demonstrate that addition of a few properly placed relay nodes can significantly extend the network lifetime. In this context it is extremely important that a set of "potential" relay node positions $\boldsymbol{R}$, given to the ILP as input, should be chosen appropriately. If the elements of $\boldsymbol{R}$ are not selected properly, it is possible that the required connectivity and coverage cannot be achieved, even if all elements of $\boldsymbol{R}$ are included in the solution.

The number of potential position in a real plan can be infinite. Therefore, we need some heuristic to limit this number to a level where the ILP becomes computationally tractable. One such heuristic is the *grid based* approach [20], where the entire networking area is viewed as an imaginary grid and the center positions (shown as small red rectangles in Fig. 3.1 (redraw from [20])) of each cell boundary are selected as potential relay node positions. The spacing between grid lines must be small enough (e.g. at most $2r$ where $r$ is the transmission range of a sensor node) that all sensor nodes have at least one potential relay node position within it transmission range. A grid based approach can provide good solutions when the network area is small and sensor nodes are densely deployed within the network. For large area, the grid based approach results in too many potential positions, since grid line spacing cannot be increased beyond a certain point (e.g. $2r$), and the ILP becomes intractable. The grid based approach is also not suitable when the sensor nodes are sparsely distributed in the sensing area.

To address the limitations of grid based approach, we propose an *intersection based approach* (e.g. depicted in Fig.3.2) in this thesis. The steps for this approach are given below:

1) Taking each sensor node $i$ (shown as yellow dot in Fig. 3.2) as center, draw an imaginary circle (shown as a dash circle in Fig. 3.2) around each sensor node, where the radius of the circle is the maximum transmission range of sensor

nodes.

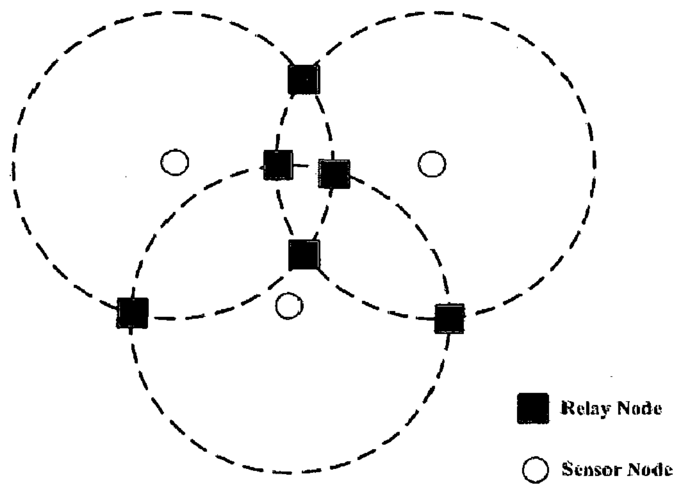


**Figure 3.2 Intersection based placement of relay nodes**

2) Pick all intersection points between pairs of circles generated in step 1 as potential relay node positions (shown as red squares in Fig. 3.2). The idea is that each intersection point is guaranteed to cover at least (possibly more) two sensor nodes and is therefore a good candidate for a potential relay node position.

3) If a sensor node has less than $k_s$ intersection points on its circumference (e.g. an isolated node having no other nodes within its transmission range), add extra potential relay node positions at random locations on its circumference.

4) If a potential relay node position, $j$, is not within the transmission range of at least $k_r$ other potential relay positions, randomly insert some additional relay positions on the circle circumference centered at $j$ with radius of rely node transmission range.

If the problem size is small enough that all intersection points may be included in the set of potential relay positions. However, for a dense distribution of sensor nodes, this number may be too high and make the ILP intractable. Therefore, if necessary, a simple heuristic (e.g. reduction heuristic presented in Appendix I) is used to remove some of the potential intersection points such that, even after removal, the remaining positions can still satisfy the coverage and connectivity requirements. This final set ***R*** of potential relay node positions is then provided as input to the ILP formulation.

# Chapter 4 Analysis of Simulation Results

## 4.1 Experimental Environment

In this chapter, we present the simulation results for our placement strategy and routing scheme. Our objective is to minimize the number of relay nodes required to form the upper tier relay node network, with respect to specified connectivity ($k_r$), coverage ($k_s$) and maximum per round energy dissipation ($e_{max}$). We compare our results to the existing placement strategies that attempt to minimize the number of relay nodes, without considering the routing scheme and corresponding energy dissipation of selected relay nodes.

We have used an experimental setup similar to [8], where the sensor nodes are randomly distributed over a $200 \times 280\ m^2$ area. The communication range of each sensor node is assumed to be $r_{max} = 40\ m$ and the communication range of each relay node is $d_{max} = 200\ m$. All relay nodes are assumed to have the same initial energy supply of the amount of $5J$. For measuring the energy dissipated by relay nodes, we adopt the *First-order Radio Model* described in section 2.1.3 and as in [20], we set up the same values for $\alpha_1$, $\alpha_2$, $\beta$ and $q$ as $\alpha_1 = \alpha_2 = 50nJ/bit$, $\beta = 100pJ/bit/m^2$ and $q = 2$ [20]. We further assume the average amount of data generated by each sensor node $i$ is $b_i = 10$ bits/round. Simulation results are obtained by CPLEX 9.1 solver.

## 4.2 Simulation Results

Table I compares the results of our intersection based approach with the grid based approach [20] that minimize the number of upper-tier relay nodes, ensuring desired connectivity (e.g. $k_r = 1, 2$) and coverage ($k_s = 1, 2$), without any energy constraints. We achieve this by setting $e_{max} = \infty$ for the ILP formulation. The intersection based approach considers all intersecting points as potential relay node locations, as discussed in section 3.6. For grid based approach, we varied the number of potential relay node locations from 48 (for coarse grid) to 165 (fine grid), which are indicated as 48-Grid, 88-Grid and 165-Grid. We only consider up to 50 sensor nodes in this experiment so that all relay nodes in intersection based approach can be included without applying any reduction heuristic.

| | | | Placement Strategy | | | |
|---|---|---|---|---|---|---|
| # of Sensors | $k_s$ | $k_r$ | 48 Grid | 88 Grid | 165 Grid | Intersection |
| 20 | 1 | 1 | 11 | 10 | 9 | 8 |
| | 2 | 2 | 22 | 20 | 18 | 15 |
| 30 | 1 | 1 | 11 | 11 | 9 | 8 |
| | 2 | 2 | 22 | 20 | 18 | 15 |
| 40 | 1 | 1 | 13 | 12 | 10 | 9 |
| | 2 | 2 | 27 | 23 | 21 | 17 |
| 50 | 1 | 1 | 14 | 12 | 12 | 10 |
| | 2 | 2 | 29 | 24 | 24 | 20 |

**Table I No. of Relay nodes required by various placement schemes**

As shown in Table I, the quality of the solutions improves with higher number of potential relay locations in grid based approach, but the intersection based approach consistently outperforms the grid based approach in all cases. The underlying reason relies on the fact that grid based approach covers the sensing field with the same imaginary grid (could be either coarse or fine) while ignoring the distribution information of sensor nodes. However, the potential relay locations in our intersection based approach are generated with respect to the distribution of sensor nodes, which is more accurate.

Unlike many existing solutions for relay nodes placement ([8], [16], [17]), our formulation does not require the same value for both $k_r$ and $k_s$. These two values can be adjusted independently. For example, it is quite possible to have $k_r = 1$, $k_s = 2$ or $k_r = 3$, $k_s = 1$ depending on user preference or application requirements.

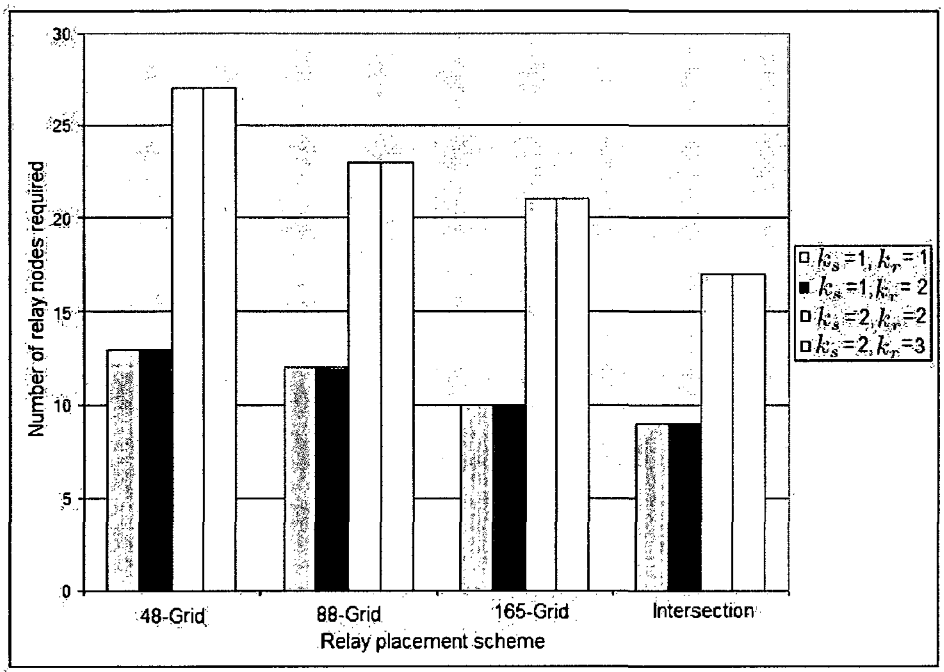


**Figure 4.1 Grid vs intersection based placement approach with different $k_s$ and $k_r$ values**

The results for different values of $k_r$ and $k_s$, on 40 nodes sensor networks, is given in Figure 4.1 (the legend follows the convention of Table 1). For the intersection based approach, we have considered all potential relay positions without reduction. In Fig. 4.1, with respect to all pairs of $k_s$ and $k_r$ values (e.g. $k_s = 2$ and $k_r =3$), the quality of the solution, in grid based approach, is enhanced as the grid is more and more finely formed (e.g. varying form 48-grid to 165 grid). But our proposed intersection based approach outperforms all cases of grid based scenarios. As also shown in this figure, with respect to all cases (e.g. Grid 48, Grid 88 , Grid 165 etc.), the required number of relay nodes significantly increases with the higher value of desired coverage ($k_s$) while fixing the value of connectivity ($k_r$). This relies on the fact that more relay nodes have to be included in the upper tier relay nodes network in order to ensure that each sensor node can communicate with at least $k_s$ number of relay nodes. However, while fixing the values of $k_s$ for those cases, increasing the value of connectivity ($k_r$) does not require more relay nodes to be included. This is because the connectivity value ($k_r$) is used to ensure each selected relay node is able to communicate with at least $k_r$ other relay nodes. In real application, the transmission range of a relay node is much longer than a regular sensor node. Therefore, the connectivity constraint ($k_r$) among upper tier relay nodes network can be easily satisfied without the necessity to include more relay node.

In the previous experiments, we only considered relatively small-size sensor networks which only contain up to 50 sensor nodes. Given sensor networks with relatively small number of sensor nodes deployed (e.g. 40 sensor nodes in the previous experiment), the number of potential relay locations generated by our intersection based approach is fairly small and can be directly supplied to the proposed ILP formulation. However, unlike grid based approach where the number of candidate relay positions is fixed no matter how many sensor nodes are deployed, potential relay node positions in our intersection based approach increase dramatically with the number of deployed sensor nodes. Therefore, given sensor networks with hundreds of sensor nodes, intersection based approach generates too many relay positions to make the ILP formulation tractable. To deal with this limitation, a reduction heuristic of relay nodes (e.g. the one presented in Appendix I) can be applied to ensure that the remaining positions, even after removal, can still satisfy the pre-specified coverage and connectivity requirements. After reducing the relay positions generated in intersection based approach, we conduct the same experiment, as shown in Fig. 4.1, on networks with hundreds of sensor nodes for different coverage and connectivity requirements. The experimental results are presented in Table II (the legend follows the convention of Table I).

| | | | Placement Strategy | | | | |
|---|---|---|---|---|---|---|---|
| # of Sensors | $k_s$ | $k_r$ | 48 Grid | 88 Grid | 165 Grid | Intersection | |
| | | | | | | Result | # of potential |
| 200 | 1 | 1 | 22 | 19 | 17 | 14 | 57 |
| | 1 | 2 | 22 | 19 | 17 | 14 | 57 |
| | 2 | 2 | 42 | 35 | 31 | 28 | 57 |
| | 2 | 3 | 42 | 35 | 31 | 28 | 57 |
| 300 | 1 | 1 | 26 | 20 | 17 | 15 | 61 |
| | 1 | 2 | 26 | 20 | 17 | 15 | 61 |
| | 2 | 2 | NA | 39 | 33 | 29 | 61 |
| | 2 | 3 | NA | 39 | 33 | 29 | 61 |
| 400 | 1 | 1 | 25 | 21 | 18 | 15 | 66 |
| | 1 | 2 | 25 | 21 | 18 | 15 | 66 |
| | 2 | 2 | NA | 39 | 34 | 31 | 66 |
| | 2 | 3 | NA | 39 | 34 | 31 | 66 |

**Table II Grid vs intersection based placement approach with different $k_s$ and $k_r$ values**

We, in this experiment, focused on sensor networks containing 200, 300 and 400 sensor nodes respectively. According to the results in Table II, intersection base approach still consistently outperforms grid based approach with respect to all cases of coverage and connectivity requirements even though not all relay positions generated in intersection based approach are included. Results in Table II, on the other hand, also demonstrate the limitation of the grid based approach. Given large number of deployed sensor nodes, a coarsely-formed grid, for example 48-Grid, cannot satisfy a high level of coverage and connectivity requirement even if all 48 potential relay nodes are included. Therefore, there is no solution, marked as "NA" in table II, can be produced (e.g. 400 sensor nodes with $k_s$ = 2 and $k_r$ =3). This limitation is resolved in the intersection based approach because potential relay nodes, in this approach, are generated with respect to the sensor locations. The significance of intersection based approach is also demonstrated in the last column of Table II, which represents the number of potential relay nodes after applying relay reduction heuristic on the initial set of relay node positions generated by our intersection based approach. By comparing with Grid 165 which use 165 potential relay positions to yield solutions of highest quality, our intersection based approach utilizes less than half of 165 potential relay positions but produces better results than Grid 165. This scenario applies to all cases in Table II.

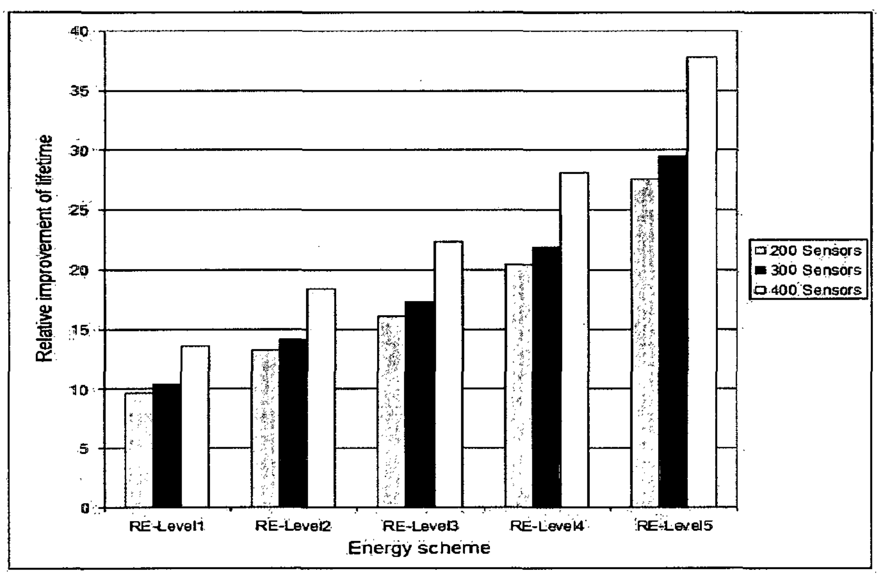


**Figure 4.2 Relative lifetime improvements using different energy constraint levels**

In the previous experiments, we have demonstrated that our ILP formulation can handle relay placement of various coverage and connectivity requirements. Compared to different scenarios of grid based approach, our intersection based approach consistently outperforms grid based approach while using much less number of potential relay positions. However, in addition to optimizing the placement of relay nodes, our proposed ILP formulation also considers the energy dissipation during the relay placement phase. Our ILP formulation is able to determine not only the optimal placement of relay nodes, satisfying the pre-specified level of coverage and connectivity, but also a routing scheme which guarantees the network lifetime. We achieved this by setting up the maximum per round energy dissipation constraint ($e_{max}$) in our proposed ILP formulation. In our next experiment, we varied the number of sensor nodes from 200 to 400 and computed potential relay node positions using the intersection based strategy combined with the relay reduction heuristic. We used 5 predefined levels for $e_{max}$, varying from RE-Level 1 (Restricted Energy – Level 1) with $e_{max} = 400000nJ$, RE-Level 2 with $e_{max} = 300000nJ$, RE-Level 3 with $e_{max} = 250000nJ$, RE-Level 4 with $e_{max} = 200000nJ$ and RE-Level 5 with $e_{max} = 150000nJ$. The lifetimes corresponding to each RE-Level were calculated based on the maximum allowed energy dissipation per round for each relay node. The lifetime $L_{min}$, obtained by setting $e_{max} = \infty$, corresponds to existing placement schemes that simply minimize the number of relay nodes without considering the per round energy dissipation.

Fig. 4.2 shows the relative improvement of lifetime compared to $L_{min}$ using different levels of energy constraints (RE-Level 1 – RE-Level 5), for 200, 300 and 400 sensors respectively. In this experiment, the relative lifetime is calculated as a ratio *of* $L_{RE\text{-}Level\ i}$ to $L_{min}$. From Figure 4.2, we can see that: for each level of energy constraint *RE-Level i*, the more sensor nodes were deployed, the more relative improvement of lifetime is obtained. As we mentioned previously, the relative lifetime improvement is calculated with respect to $L_{min}$ which refers to the case of finding the minimum number of relay nodes without considering the per round energy dissipation. In such a case, it is highly possible that the resultant data gathering scheme may not be energy-efficient, which means the data forwarding pattern is not optimized and unbalanced. The unbalanced distribution of data flow will result in a situation where some relay nodes are responsible for transmitting much more data volume than the others, and hence dissipate energy at a much higher rate. As more sensor nodes are deployed, the unbalanced load of data will introduce more severe negative effect on the network lifetime simply because more data packets need to be forwarded to the base station.

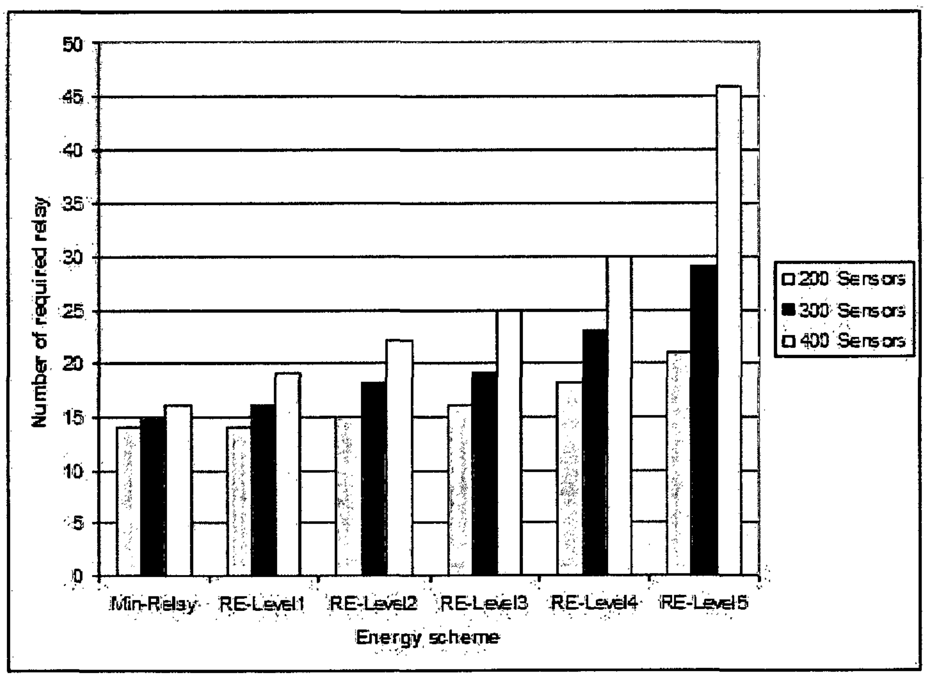


**Figure 4.3 No. of required relay nodes with different energy constraint levels**

On the other hand, by setting the maximum per round energy dissipation constraint ($e_{max}$) to a specific level, our ILP formulation will enforce some sort of balanced data loading among selected relay nodes to ensure the energy dissipation of each selected relay node is within the constraint ($e_{max}$). However, these improvements come at the cost of introducing some additional relay nodes in the network. Figure 4.3 shows the number of relay nodes required for each scenario investigated in the experiments shown in Fig. 4.2. The *Min-Relay* indicates the minimum number of relay nodes required, without restriction on energy dissipation, to obtain a lifetime of $L_{min}$. As expected, the required number of relay nodes increases as the value of $e_{max}$ is more and more strictly constrained, but it can be seen from Fig 4.2 and Fig. 4.3 that, by using our approach, the network lifetimes can be significantly improved while allowing only very few of extra relay nodes. For example, for 400 nodes sensor networks by adding at most 30 extra relay nodes, the network lifetime can be increased as much as 38 times.

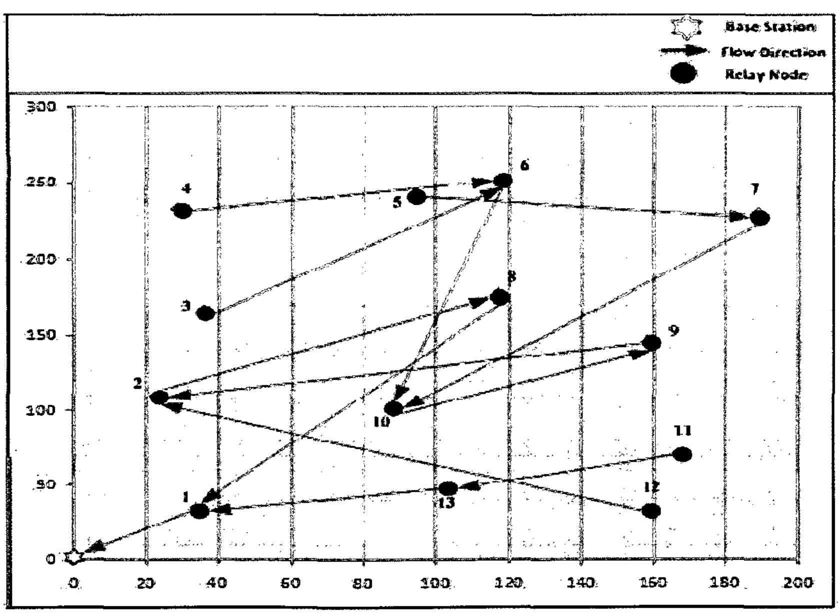


**Figure 4.4 Example of routing scheme of 200 sensors with $e_{max} = \infty$**

As we claimed previously, our ILP formulation is capable of finding a suitable routing scheme which provides a guaranteed network lifetime while jointly optimizing the relay node placement problem. This routing scheme is not directly generated by our ILP formulation. However, it can be reconstructed from the solution of the ILP formulation. In other words, for a given problem, the data gathering pattern is implied in the solution of the ILP formulation.

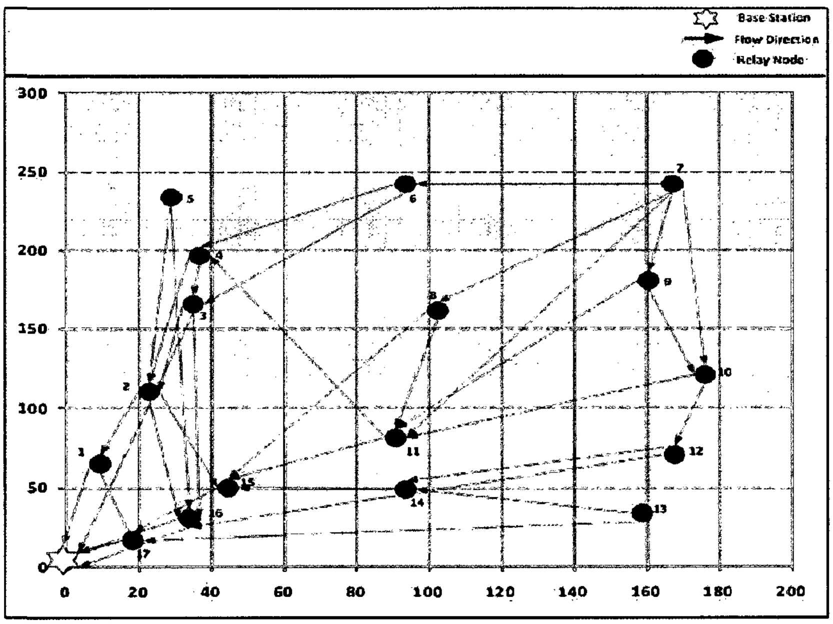


**Figure 4.5 Example of routing scheme of 200 sensors with $e_{max}$ = 200k *nJ***

Figure 4.4 presents an example of data gathering pattern for a 200 sensor nodes network, which only focuses on minimizing the number of placed relay nodes without considering the energy dissipation. Figure 4.5, on the other hand, depicts the data gather scheme for the same 200 sensor nodes network where the maximum per-round energy dissipation is constrained at a level of 200k *nJ*. Fig. 4.4 employs 13 relay nodes serving as cluster heads to relay data from 200 sensor nodes to the base station. However, all data flows are collected at relay node 1 before sending to the base station. This kind of centralized data gathering will lead to relay node 1 dissipating energy at a much higher rate than the rest of other relay nodes and quickly becomes "dead"(e.g. running out of power). Fig. 4.5, on the other hand, utilizes 17 relay nodes but the data flows are distributed at relay node 1, 2, 16, 17 before transmitting to the base station. Moreover, unlike Fig. 4.4 in which there is only single outgoing data flow from each relay node, routing scheme in Fig. 4.5 applies some flow splitting on some relay nodes (e.g. relay node 6, 7, 10 etc.) to reduce the energy dissipation at those nodes. This kind of balanced data gathering contributes to the significant network lifetime improvement through evenly distribute data flow among selected relay nodes.

Applying per-round energy constraint also results in an optimized data routing scheme. As we can see from Fig. 4.4, the data from relay node 4 to the base station follows the path "4 → 6 → 10 → 9 → 2 → 8 → 1 → Base". With respect to the *first order radio model* [2], energy dissipation is directly related to the distance between the source and destination. An alternative better routing path, such as "4 → 2 → 1 →Base", would save large amount of energy for those excluded relay nodes. A more energy-efficient routing scheme is presented in Fig. 4.5, where energy dissipated at each relay node is meant to convey the data closer to the base station, for example, relay node 5 sending data through "5 → 2 → 1 → Base".

# Chapter 5 Conclusions and Future Work

## 5.1 Conclusions

In two-tiered, cluster based sensor networks using relay nodes as cluster heads, conventional approaches solve the relay nodes placement problem that ensures connectivity and coverage, and the routing separately. In this thesis, we have solved these problems jointly, using an ILP formulation.

Our formulation determines the number and positions of the relay nodes such that each sensor nodes is covered by at least $k_s$ relay nodes, and the relay node network is $k_r$-*connected*, while ensuring that a specified network lifetime is achieved by constraining the energy dissipation of all relay nodes to be below a given value. Our approach also determines an appropriate routing scheme that reduces the energy dissipation of the critical relay node(s). Moreover, we have proposed an intersection based approach for preparing the initial set of potential relay positions. The simulation results demonstrate that our intersection based approach consistently outperforms grid based approach [20]. We, through simulation results, also demonstrated that our ILP approach can significantly increase the network lifetime, as well as can provide desired level of fault tolerance at the cost of a few additional relay nodes. We show that our ILP formulation is able to generate optimal solutions for networks with hundreds of sensor nodes.

## 5.2 Future Work

As a direction of future work, we are currently working on developing a distributed approach that can be used for even large sensor networks.

# Appendix I Relay Reduction Heuristic

```
RelayReduction_Heuristic(Input: S_sensor, k_s, k_r, P_base ; Output: S_relay)
begin
        S_intersecs = Obtain circle intersection for every pair of point in S_sensor;
        Tab_relay_sensor = Construct relay covering sensor table from S_intersecs, and S_sensor;
        Foreach( Point p in S_sensor)
        do
            if ( CheckCovering(p, Tab_relay_sensor ) < k_s )
            then
                Tab_relay_sensor = AddMorePostions(Tab_relay_sensor , p, k_s);
            endif
        end
        Tab_relay_sensor = RowWiseAssendingSort(Tab_relay_sensor, P_base);
        Tab_relay_sensor = RemoveDuplicate(Tab_relay_sensor , k_s);
        S_relay = Φ;
        While (not all points in S_sensor is covered)
        do
            Find a relay R in Tab_relay_sensor covering the maximum number of sensors;
            Add R to S_relay;
            Update relay-covering-sensor table Tab_relay_sensor ;
        end
        Foreach( Point p in S_relay)
        do
            if ( CheckConnectivity(p, S_relay) < k_r )
            then
                S_relay = AddMorePostions(S_intersects, k_r);
            endif
        end
        S_relay = ConnectingToBase(S_relay, P_base);
        Return S_relay;
end
```

# Bibliography


[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey", *Computer Networks,* volume 38, pages 393-422, 2002.

[2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy Efficient Communication Protocol for Wireless Micro-sensor Networks", *In HICSS,* pages 3005-3014, 2000.

[3] J. Pan, Y.T. Hou, L. Cai., Y. Shi and S.X. Shen, "Topology Control for Wireless Sensor Networks", *In International Conference on Mobile Computing and Networking,* pages 286-299, 2003.

[4] X. Cheng, D-Z. Du, L. Wang and B.B. Xu, "Relay Sensor Placement in Wireless Sensor Networks", *IEEE Transaction on Computers,* 2001.

[5] E. Falck, P. Floren, P. Laski, J. Kohonen and P. Orponent, "Balanced Data Gathering in Energy-Constrained Sensor Networks", *In ALGOSENSORS 2004,* LNCS volume 3121 pages 59-70, Springer, 2004.

[6] G. Gupta and M. Younis, "Load-balanced Clustering of Wireless Sensor Networks", *In IEEE ICC 03,* volume 3, pages 1848-1852, 2003

[7] G. Gupta and M. Younis, "Fault-tolerant Clustering of Wireless Sensor Networks", *In Proceeding of IEEE WCNC,* pages 1579-1584, 2003.

[8] J. Tang, B. Hao and A. Sen, "Relay Node Placement in Large Scale Wireless

Sensor Networks", *Computer Communications,* volume 29(4), pages 490-501, 2006.

[9] A. Bari, A. Jaekel and S. Bandyopadhyay, "Optimal Load Balanced Clustering in Two-Tiered Sensor Networks", *In IEEE Basenets,* 2006.

[10] A. Bari, A. Jaekel and S. Bandyopadhyay, "Integrated Clustering and Routing Strategies for Large Scale Sensor Networks", *IFIP NETWORKING 2007,* Springer LNCS 4479, pages 143-154, 2007.

[11] A. Bari, A. Jaekel and S. Bandyopadhyay, "Maximizing the Lifetime of Two-Tiered Sensor Networks", *In IEEE EIT,* pages 222-226, 2006.

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

[13] G. Lee, M. Lee, W. Seok, J. Kong and O. Byeon. "A Base Station Centralized Simple Clustering Protocol for Sensor Networks", *In IFIP 06,* LNCS 4096, pages 682-691, 2006.

[14] K. Kalpakis, K. Dasgupta and P. Namjoshi, "Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks", *In IEEE International Conference on Networks,* 2002.

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

[16] B. Hao, J. Tang and G. Xue, "Fault-tolerant Relay Node Placement in Wireless Sensor Networks: formulation and approximation", *In HPSR,* pages 246-250, 2004.

[17] H. Liu, P. Wan and W. Jia, "Fult-tolerant Relay Node Placement in Wireless Sensor Networks", *In COCOON,* 2005.

[18] S.C. Ergen and P. Varaiya, "Optimal Placement of Relay Nodes for Energy Efficiency in Sensor Networks", *In ICC,* 2006.

[19] M. Yarvis, N. Kushalnagar, H. Singh, A. Rangarajan, Y. Liu and S. Singh, "Exploiting Heterogeneity in Sensor Networks, *In INFOCOM 2005,* volume 2, pages 878-890, 2005.

[20] A. Bari, A. Jaekel and S. Bandyopadhyay, "Optimal Placement of Relay Nodes in Two-Tiered, Fault-Tolerant Sensor Networks", *IEEE ISCC,* 2007.

[21] W. Ye, J. Heidemann and D. Estrin, "Medium Access Control with Coordinated Adaptive Sleeping for Wireless Sensor Networks", *IEEE/ACM Transactions on Networking,* volume 12, pages 493-506, 2006.

[22] A. Mainwarning, J. Polastre, R. Szewczyk, D. Culler and J. Anderson, "Wireless Sensor Networks for Habitat Monitoring", *In the Proceedings of ACM WSNA,* 2002.

[23] N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan and D. Estrin, "A Wireless Sensor Network for Structural Monitoring", *In the*

*Proceedings of ACM Sensys,* 2004.

[24] S. Coleri and P. Varaiya, "Optimal Placement of Relay Nodes in Sensor Networks", *In the Proceedings of ICC,* 2006.

[25] R. Subramanian and F. Fekri, "Sleep Scheduling and Lifetime Maximization in Sensor Networks: fundamental limits and optimal solutions", *In the proceedings of IEEE IPSN,* pages 218-225, 2006.

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

[27] 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.

[28] C.-Y. Chong and S.P. Kumar, "Sensor Networks: Evolution, Opportunities, and Challenges", *Proc. of the IEEE,* volume 8, pages 1247-1256, 2003.

[29] K. Akkaya and M. Younis, "A Survey on Routing Protocols for Wireless Sensor Networks", *IEEE Transactions on Mobile Computing,* volume 3, pages 325-349, 2005.

[30] J. N. Al-Karaki and A. E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey", *In IEEE Wireless Communications,* 2005.

[31] V. Rai and R. Mahapatra, "Lifetime Modeling of a Sensor Network", *In IEEE International Conference on Design, Automation and Test in Europe (DATE).,*

volume 1, pages 202-203, 2005.

[32] G. Lin and G. Xue, "Steiner Tree Problem with Minimum Number of Steiner Points and Bounded Edgelength", *Information Processing Letters,* volume 69, pages 53-57, 1999.

[33] K. Dasgupta, M. Kukreja and K. Kalpaki, "Topology-aware Placement and Role Assignment for Energy-efficient Information Gathering in Sensor Networks", *In Proceedings of Eighth IEEE International Symposium on Computer and Communication,* pages 341-348, 2003

[34] B. Krishnamachari and F. Ordonez, "Analysis of Energy-Efficient, Fair Routing in Wirelsess Sensor Networks through Non-Linear Optimization", *In Proceedings of IEEE 58$^{th}$ Vehicular Technology Conference,* volume 5, pages 2844-2848, 2003.

[35] J. Suomela, "Relay Placement in Sensor Networks", *Master Thesis, University of Helsinki, Department of Computer Science,* 2005.

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

[37] A. Manjeshwar and D.P. Agarwal, "TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks", *In Proceedings 15$^{th}$ International Parallel and Distributed Processing Symposium,* pages 2009-2015, 2001.

[38] Y. Ossama and F. Sonia, "HEED: A Hybrid, Energy-efficient, Distributed

Clustering Approach for Ad Hoc Sensor Networks", *IEEE Transactions on Mobile Computing,* volume 3, pages 366-379, 2004.

[39] A. Bari, "Optimization Strategies for Two-tiered Sensor Networks", *Master Thesis, School of Computer Science, University of Windsor,* 2006.

[40] A. Bari, Y, Xu, X. Wu and A. Jaekel, "Design of Sensor Networks with Guaranteed Connectivity and Lifetime", *In Proceedings of the 3rd International Conference on Wireless Internet,* Austin, Texas, 2007

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

[42] A. Bari, Y. Xu, and A. Jaekel, "Integrated Placement and Routing of Relay Nodes for Fault-tolerant Hierarchical Sensor Networks", *In Proceedings of 17th International Conference on Computer Communication and Networks*, pages 1-6, 2008

# VITA AUCTORIS

NAME: Yufei Xu

EDUCATION:

Master of Science, Computer Science
University of Windsor, Windsor, Canada
2006 – 2009

Bachelor of Science, Computer Science (Honor)
University of Windsor, Windsor, Canada
2003 – 2006

Bachelor of Engineering, Industrial Automation
Nanjing University of Science & Technology, Nanjing, China
1998 – 2002