Design of Sensor Networks with Guaranteed Connectivity and Lifetime

Ataul Bari
School of Computer Science
University of Windsor
401 Sunset Ave., Windsor
ON N9B 3P4, Canada
bari1@uwindsor.ca

Yufei Xu
School of Computer Science
University of Windsor
401 Sunset Ave., Windsor
ON N9B 3P4, Canada
xu1t@uwindsor.ca

Xin Wu
School of Computer Science
University of Windsor
401 Sunset Ave., Windsor
ON N9B 3P4, Canada
wu11f@uwindsor.ca

Arunita Jaekel
University of Windsor
401 Sunset Ave., Windsor
ON N9B 3P4, Canada
arunita@uwindsor.ca

ABSTRACT

Nodes in sensor networks are often prone to failure, particularly when deployed in hostile territories, where chances of damage/destruction are significantly higher. In many applications it is necessary to have some guarantees on the coverage, connectivity and lifetime of the sensor network. The network should also be able to adapt to single and/or multiple node failures as well as disruptions due to the inherent limitations of the wireless communication medium. In hierarchical sensor networks using *relay nodes*, sensor nodes are arranged in clusters and higher-powered relay nodes can be used as cluster heads. In this paper, we propose an integer linear program (ILP) for determining the minimum number of relay nodes, along with their locations and a suitable communication strategy such that the network is able to meet specified performance guarantees with respect to coverage, connectivity and lifetime. To the best of our knowledge, this is the first formulation that jointly optimizes energy-aware placement and routing of relay nodes in two-tiered sensor networks.

Categories and Subject Descriptors

C.2 [**Computer-communication networks**]: Network Protocols

General Terms

Network design

1. INTRODUCTION

The scalability and the lifetime of *sensor* networks [1] are affected by the limited transmission range and the battery power of sensor nodes. In a two-tiered sensor network, the individual sensor nodes, comprising the lower tier, are grouped into clusters, and transmit data to their respective cluster heads. The cluster heads form the upper tier of the network, and are responsible for collecting and forwarding data toward the base station. Recently, *relay nodes* have been proposed for balanced data gathering, reduction of transmission range, connectivity and fault tolerance [3], [4], [5], [6]. These nodes can be provisioned with higher power and added functionality, as compared to the sensor nodes, and are ideally suited to serve as cluster heads in a hierarchical, two-tier sensor network [7], [8], [9], [10], [11], [12], [13], [14], [15], [16].

The major source of power consumption in a sensor network is due to the wireless communication, which increases rapidly with the distance between the source and the destination of the communication. The lifetime of a sensor network is typically determined by the battery power of the "critical node(s)" in the networks [2], [3]. Therefore, it is extremely important to devise strategies that extend the lifetime of the critical nodes and consequently the lifetime of sensor network as a whole.

In this paper, we consider a two-tiered network architecture, where, in the upper tier, higher powered relay nodes are used as cluster heads. The sensor nodes lie in the lower tier and transmit their data directly to their respective cluster heads. Therefore, individual sensor nodes are relieved from the burden of routing and forwarding, which reduces the energy consumption of the nodes. Each relay node collects data from the sensor nodes belonging to its own cluster and forwards the collected data to the base station (or sink). Data communication from relay nodes to the base station is generally *multi-hop* (where each relay node, in addition to the forwarding the data it receives from its own cluster, also forwards data it receives from other relay nodes, towards the base station, using multi-hop paths) [9], [10], [16], [18]. *Single-hop* communication can be considered as a special case, where each relay node receives data only from its own cluster, and sends this data directly to the base station, provided that the base station lies within the transmission

WICON 2007, October 22-24, Austin, USA

DOI 10.4108/pwsn.2007.2273

range of each relay node [2], [17]. An example of *multi-hop data transmission model* (MHDTM) is shown in Fig. 1. 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.

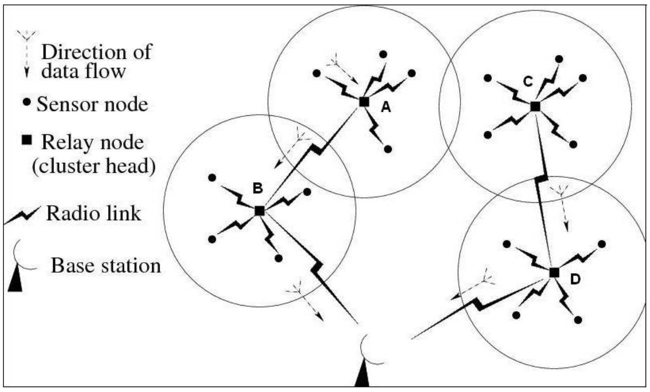

Figure 1: An example of two-tiered sensor network

Under fault-free conditions, it is sufficient for a sensor node to be covered by a single relay node and the relay node network to be 1-connected. But in this scenario, the failure of a single relay node results in data loss from all sensor nodes belonging to the cluster of the failed relay node. Such failure of a relay node may also prevent information flow of other relay nodes, which are using the failed node for forwarding data towards the base station. In order to ensure adequate performance of the sensor network, it is important to have a placement strategy with some redundancy, so that each relay node can forward its data to multiple relay nodes (or directly to the base station), so the routing strategy is able to adapt to node/link failures. Similarly, to ensure adequate coverage, each sensor node should be able to communicate with multiple relay nodes. The desired level of redundancy will depend on the intended application, and a generalized formulation should be capable of handling this.

In this paper we present an integer linear program (ILP) for designing two-tiered sensor networks with some specified performance guarantees. We assume that sensor nodes have been placed at required positions, and we determine

i) the locations of the relay nodes for the upper-tier network, and

ii) an optimal routing strategy for the relay node network such that performance requirements in the following areas are always satisfied:

   1) coverage requirements: each sensor node is *covered* by (i.e. can communicate with) at least $k_s, k_s = 1, 2...$ relay node(s).
   2) connectivity requirements: the relay node network is k_r-connected, $k_r = 1, 2...$, and
   3) energy requirements: the energy dissipation of a relay node cannot exceed a maximum allowed value e_{max}.

The parameters k_s and k_r are determined by the application, and specified as inputs to the ILP. Mission-critical applications will typically use higher values of k_s and k_r. The energy requirements are used to specify the desired network lifetime, assuming initial energy of the relay nodes is given. The objective is to achieve the desired lifetime and level of fault tolerance, with as few relay nodes as possible.

Recently, some heuristic placement strategies have been proposed in the literature for the special case where $k_s, k_r = 1, 2$, i.e. for single and double coverage of sensor nodes, as well as single and double connectivity of the relay node network [5], [10], [20], [21]. An optimal placement scheme, which minimizes the number of relay nodes is presented in [27]. However, all of these approaches only consider coverage and connectivity, and do not take into account the energy dissipation of the relay nodes. Furthermore, they only determine the *placement* of the relay nodes and do not address the routing problem at all. To the best of our knowledge, this is the first formulation that jointly optimizes energy-aware placement and routing of relay nodes in two-tiered sensor networks.

The remainder of the paper is organized as follows. In Section 2, we review previous placement and routing strategies for relay nodes in sensor networks. In Section 3, we present our ILP formulation for optimal relay node placement and routing and discuss and analyze our experimental results in Section 4. Finally, we conclude with a critical summary and some directions for future work in Section 5.

2. REVIEW

In recent years, a number of papers have considered the issue of placement of relay nodes in sensor networks. In [4], [5], [22] and [23], the authors have considered the placement problem of relay nodes in a *flat* sensor network architecture. In [4], the authors have focused on maximizing the lifetime of a sensor network, under the constraint that each point in the sensing region is covered by at least one sensor node, and proposed an algorithm for finding the location of nodes, along with their roles, to achieve the objective. In their model, any node can assume the role of a sensor node or the relay node. In [5], the authors have formulated an optimization problem, that places relay nodes to ensure that the resulting network is connected. They have focused on special class of sensors, e.g., biomedical sensor networks and solved the problem based on an approach that uses the well known concept of *Steiner Minimum Tree with minimum number of steiner points*. In [22], the authors have formulated the relay node placement problem, with the objective to maximize the lifetime of the network, as nonlinear program and proposed approximation algorithm. In [23], authors have addressed the placement problem of the sensor nodes, the relay nodes and the base station in flat sensor networks, and proposed a number of ILP formulations to achieve different objectives, such as minimizing the number of sensor nodes to be deployed while maintaining the coverage and connectivity, minimizing the cost and the energy consumption, and maximizing the lifetime and the utilization of the resource in sensor networks.

The problem of relay node placement in hierarchical sensor network architecture is addressed in [3], [10], [16], [20], and [21]. In [3], the authors have proposed strategies that maximize the topological lifetime of a sensor network by arranging the relay nodes (they have called them *Application Nodes (ANs)*) and finding the optimal location of the base station. In [20], the authors have proposed an approximation algorithm to achieve single and double connectivity of the sensor and relay nodes in a network. In [21], authors have proposed a two-step approximation algorithm to obtain 1-connected (in the first step) and 2-connected (in the second step, by adding extra back-up nodes to the result of the first step) sensor and relay node network. The works in [20] and [21] have not addressed the general case of k-connectivity for fault tolerance. In [16], author have focused on prolonging the lifetime of sensor networks with energy provisioning to the existing nodes and deploying relay nodes within the networks and proposed MILP formulation and heuristic to solve the problem. This work does not consider the fault tolerance. In [10], a hierarchical network architecture is considered where the entire region is divided into cells, and an optimal solution is determined for each cell. The authors have considered relay node networks, with each cell having a length $2r.l$, where l is an integer and r is the communication range of each sensor node. The *P-positions* for a pair of sensor nodes at locations $\boldsymbol{x}$ and $\boldsymbol{y}$ are defined as the point(s) of intersection (if any) of two circles of radius r with centers at $\boldsymbol{x}$ and $\boldsymbol{y}$ in the same cell. An optimal placement of relay nodes for each cell is computed from $\wp$, the set of P-positions for all pairs of sensor nodes within the cell, by checking *all* subsets of $\wp$ of size four or less. Their method is able to find a solution if the transmission range of the relay nodes, $R \geq 4r$, and is not able to handle the general case of k-connectivity.

A number of routing schemes for two-tiered networks have been proposed in the literature [7], [8], [9], [11], [16], [18]. Most of these adopt the *flow-splitting* (also called *multi-path* routing) model. In contrast, in a *single-path* routing model, a node is not allowed to split the traffic, and forwards all its data to a single neighbor. This model avoids many limitations of the flow splitting model [9].

The problem of routing in wireless sensor networks, under the "flow-splitting" model, has been extensively covered in the literature. In [16], Hou et al. have attempted to maximize the lifetime of a sensor network by provisioning relay and sensor nodes with additional energy using a mixed-integer *non-linear* program and have proposed a heuristic. In [18], the authors have formulated the lifetime optimization problem, under the flow-splitting model. In [6], Falck et al. have addressed the issue of balanced data gathering in sensor networks and have proposed a LP formulation that enforces some balancing constraints in the data gathering schedule. In [7], Gupta and Younis have focused on load balanced clustering and have proposed a heuristic solution for the optimization problem. Routing without flow splitting (i.e., single-path routing) has been studied in [9], [11], [13], [14], [24], and [25]. In [9], the authors have presented a transformation algorithm to convert a multiple outgoing flow routing model to a single outgoing flow routing model. In [25], the authors have investigated the problem of maximizing network lifetime by appropriately placing nodes which are not energy constrained (e.g., connected to a wall outlet). In [24], the authors propose a formulation for constructing minimum-energy data-aggregation trees, for a *flat* architecture. In [13], [11] and [14], minimizing the number of relay nodes and finding there locations were not considered.

3. NETWORK DESIGN WITH PERFORMANCE GUARANTEES

3.1 Network Model

For our model, we consider a two-tiered wireless sensor network, where the lower tier consists of n sensor nodes, randomly distributed in 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 prespecified 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 $k_s = 1, 2, ..$, and each relay node can forward its data to k_r, $k_r = 1, 2, ..$, 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 relay 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 the 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. We are also given a set of *potential* locations for the relay nodes. For this paper, we have assumed that the set of potential positions are situated on an imaginary grid covering the entire sensing area. The grid may be made as fine or as coarse as desired. A finer grid increases the number of potential locations and typically results in better solutions. However, this increases the complexity of the formulation, and hence the time required to obtain a solution. It is also possible to choose the potential locations based on the P-positions [10], determined by the intersections of the circles representing the transmission areas of the sensor nodes. In either case, once the set of potential locations of the relay nodes are given, our ILP can be used to generate the upper-tier network, with desired coverage and connectivity.

Let $\mathcal{S}$ be the set of all sensor nodes. We assign each node a unique label as follows:

1. for each sensor node, a label $i, 1 \leq i \leq n$,
2. for each possible location of relay node, a label $j, n < j \leq n + m$ and
3. for the base station, a label $n + m + 1$.

In our model, at any given point of time, each sensor node communicates with only one relay node and in each cluster, one relay node acts as a cluster head. Data gathering is *proactive*, i.e., data are collected and forwarded to the base station periodically, following a schedule, determined separately. We will refer to each period of data gathering as a *round* [18]. The dominant factor in power consumption in sensor networks is the power needed for communication. In the first-order radio model [2], receive (transmit) circuitry consumes α_1 nJ/bit (α_2 nJ/bit) of energy. The total energy to receive b bits is given by, $E_{R_x}(b) = \alpha_1 b$ while the total energy needed to transmit b bits over a distance d is given by $E_{T_x}(b, d) = \alpha_2 b + \beta b d^q$, where q is the *path loss exponent*, $2 \leq q \leq 4$ [3] and β 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, $q = 2$.

3.2 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 lying between 1 and n.
- m: The total number of possible positions of relay nodes, with each position having a unique index lying between $n+1$ and $n+m$.
- 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 number of relay nodes covering each sensor node.
- k_r: Desired connectivity of the relay node network.
- α_2 (α_1): Energy coefficient for transmission (reception).
- β: Energy coefficient for amplifier.
- q: Path loss exponent.
- $\mathcal{D}$: A large constant.
- 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:

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

$$X_{i,j} = \begin{cases} 1 & \text{if the sensor node } i \text{ transmits to} \\ & \text{the relay node } j, \\ 0 & \text{otherwise.} \end{cases}$$

- $P_{j,k}$: Binary variable defined as follows:

$$P_{j,k} = \begin{cases} 1 & \text{if the rely node } j \text{ transmits to} \\ & \text{the relay node } k, \\ 0 & \text{otherwise.} \end{cases}$$

- Y_j: Binary variable defined as follows:

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

- 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 node j.
- G_j: Continuous variable indicating the amount of energy needed 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 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 the 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).

3.3 ILP Formulation for network design

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 a relay node 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 value for k_s and k_r, this formulation can ensure fault tolerance.

Given the network as described in Section 3.1, the objective of this formulation is to minimize the number of relay nodes, such that each sensor node can communicate with at least one relay node. The formulation is given below.

$$\textbf{Minimize} \sum_{j=n+1}^{n+m} Y_j \tag{1}$$

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.

$$X_{i,j} \cdot d_{i,j} \leq r_{max} \qquad \forall i, 1 \leq i \leq n, \quad \forall j, n+1 \leq j \leq n+m \tag{2}$$

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.

$$P_{j,k} \cdot d_{j,k} \leq d_{max} \qquad \forall j,k: \ j \neq n+m+1 \tag{3}$$

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

$$Y_j \geq X_{i,j} \qquad \begin{array}{l}\forall i, 1 \leq i \leq n, \\ \forall j, n+1 \leq j \leq n+m\end{array} \tag{4}$$

d) A sensor node must be connected to at least k_s relay nodes.

$$\sum_{j=n+1}^{n+m} X_{i,j} \geq k_s \qquad \forall i, 1 \leq i \leq n \tag{5}$$

e) Constraint that determines the number of the relay nodes that the relay node j can use to route data towards the base station.

$$C_j = \sum_{w:(d_{j,w} \leq d_{max}) \ AND \ (d_{w,n+m+1} < d_{j,n+m+1})} Y_w \tag{6}$$

Constraint (6) has to be repeated for all j, $n < j \leq n+m$.

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.

$$C_j \geq k_r \cdot Y_j \qquad \forall j: \ d_{j,n+m+1} \geq d_{max} \tag{7}$$

f) Non flow-splitting constraint.

$$\sum_k P_{j,k} = Y_j \qquad \forall j,k: \ j \neq n+m+1 \tag{8}$$

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

$$w_j = \sum_i b_i \cdot X_{i,j} \qquad \begin{array}{l}\forall i, 1 \leq i \leq n, \\ \forall j, n+1 \leq j \leq n+m\end{array} \tag{9}$$

h) Flow constraint.

$$\sum_k f_{j,k} - \sum_k f_{k,j} = w_j \tag{10}$$

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

$$T_j = \sum_j f_{j,k} \qquad \forall j,k, j \neq n+m+1 \tag{11}$$

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

$$G_j = \beta \sum_k f_{j,k} \cdot d_{j,k}^q \qquad \forall j,k, j \neq n+m+1 \tag{12}$$

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

$$R_j = \sum_k f_{k,j} \qquad \forall j, n < j \leq n+m+1 \tag{13}$$

l) Base station does not transmit.

$$f_{n+m+1,k} = 0 \qquad \forall k, 1 \leq k \leq n+m+1 \tag{14}$$

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

$$f_{j,k} \leq \mathcal{D} \cdot P_{j,k} \qquad \forall j,k, j \neq n+m+1 \tag{15}$$

n) Calculate the energy dissipated by relay node j.

$$\alpha_1(R_j + w_j) + \alpha_2 T_j + G_j = E_j, \ \forall j: \ j \neq n+m+1 \tag{16}$$

o) Constraint for maximum energy dissipation.

$$E_j \leq e_{max} \qquad \forall j: \ j \neq n+m+1 \tag{17}$$

3.4 Justification of the ILP Equations

Equation (1) is the objective function for the formulation that minimizes the total number of relay nodes. The minimization of the number of relay nodes is obtained after ensuring the required connectivity and coverage of all the individual sensor and relay nodes in the area of interest as well as ensuring the desired lifetime of the relay node networks.

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 the sensor node.

b. Constraint (3) enforces the restriction that a relay node can only transmit to another relay node (or to the BS), if the destination node is within the transmission range of the transmitting relay node.

c. Constraint (4) ensures that if the relay node r_j at location j is chosen as a 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 cluster head for any sensor node, normally it should not be selected (unless it is needed to maintain required connectivity). This is not specifically enforced by any constraint, but is taken care of by the objective function, which will set $Y_j = 0$, if this does not violate any of the other constraints.

d. Constraint (5) requires that each sensor node be covered by at least k_s relay nodes, instead of a single relay node. The actual value of k_s, can be chosen based on the intended application. For most applications $k_s = 2$ or 3 should 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 select another cluster head from the remaining $k_s - 1$ nodes.

e. Constraints (6) and (7) determine the connectivity of the relay node network.

f. Constraint (8) prevents flow-splitting by specifying that a relay node j can transmit to only one other node k.

g. Constraint (9) calculates the total number of bits, w_j, generated in cluster j, by summing the data transmitted to it by all the sensor nodes belong to the cluster j.

h. Constraint (10) corresponds to the standard flow constraints [26], and states that the total data flowing from 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 in cluster j, (w_j).

i. Constraint (11) 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 k.

j. Constraint (12) calculates the amplifier energy, G_j, by summing the energy required for each link. In the actual solution, only one outgoing link will have non-zero data flow.

k. Constraint (13) specifies the total number of bits received at node j from other relay node(s), by summing the data flow on all incoming links.

l. Constraint (14) specifies that the base station $n+m+1$, does not transmit to any other node.

m. Constraint (15) specifies that data can be sent from node j to node k, only if link (j,k) is selected as the single outgoing link by constraint (8), i.e. $Y_j = 1$. If $P_{j,k} = 0$, then constraint (15) forces $f_{j,k} = 0$. The constant $\mathcal{D}$ is needed since the value of $f_{j,k}$ may be greater than 1. The value of $\mathcal{D}$ should be large enough to allow the maximum possible data flow on link (j,k). We have set $\mathcal{D} = \sum_i b_i, 1 \le i \le n$.

n. Constraint (16) 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 receive energy $\alpha_1(R_j + w_j)$,

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

   iii) the transmit amplifier energy G_j.

o. Constraint (17) ensures that the total energy dissipated by a relay node cannot exceed e_{max}, which is supplied as data to the formulation.

Theorem 1: Constraints (6) and (7) guarantee that the relay node network can survive $k_r - 1$ faults.

Proof: For each relay node r_j in the upper tier network, constraint (6) 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 (7) ensures that there are at least k_r such nodes, for any relay node which cannot transmit to the base station directly. This means that even if up to $k_r - 1$ relay nodes fail, 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 (7) 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 node network has the desired connectivity.

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.

If the relay nodes have transmitter (receiver) capacity constraint for the maximum number of bits that can be transmitted (received), it can be easily handled by adding two more constraints to our formulation, as follows:

$$T_j \le T_{max}, \qquad \forall j,\ n < j \le n+m \tag{18}$$

$$R_j \le R_{max}, \qquad \forall j,\ n < j \le n+m \tag{19}$$

where T_{max} (R_{max}) is the capacity of the transmitter (receiver).

4. EXPERIMENTAL RESULTS

In this section, we present the simulation results for our placement strategy. Our objective is to minimize the number of relay nodes required to form the upper tier relay node network, with a specified connectivity (k_r), coverage (k_s) and maximum energy dissipation (e_{max}). We compare our results to the existing placement strategies [27] that attempt to minimize the number of relay nodes, without considering the routing strategy and corresponding energy dissipation of the nodes.

We have used an experimental setup similar to [10], where the sensor nodes are randomly distributed over a $200 \times 280m^2$ area. The communication range of each sensor node is assumed to be $r_{max} = 40m$, and the communication range of a relay node is set to $d_{max} = 200m$. The initial energy of the relay nodes is taken as $5J$ and all relay nodes are assumed to have same initial energy. We experimented with different sensor node distributions ranging from 50 nodes to 100 nodes. We also varied the number of possible relay node locations from 48 to 165 possible locations. Finally, we experimented with different values of k_r and k_s.

In Fig. 2 and Fig. 3, we show the results of our experiments with 50 sensor nodes networks. The results with the 75 and 100 sensor node networks are similar. Fig. 2 shows the amount of lifetime (in rounds) that are achieved by the network for different number of initial positions of the relay nodes, with $k_s = k_r = 1$. The values of the lifetime, for each number of initial positions, are shown in the sequence, *Min-relay, Re-Level 1, Re-Level 2* and *Re-Level 3.* The Min-Relay indicates the lifetime achieved by minimizing *only* the number of relay nodes (i.e., e_{max} is set to ∞). This corresponds to existing approaches that do not consider the combined problem of routing and energy dissipation. The other three results indicates the lifetime obtained using different values of e_{max}, the maximum allowed energy dissipation per round for each relay node. In our experimental set-up, RE-Level 1 (Restricted Energy - Level 1) is the most relaxed with $e_{max} = 100000nJ$ and Level 3 is the most constrained with $e_{max} = 50000nJ$. Fig. 3 shows the number of relay nodes required for each scenario investigated in the experiments from Fig. 2. As expected, the required number of relay nodes is increased as the value of e_{max} is constrained, but it can be seen from these two figures that, by using our approach, the network lifetimes can be significantly improved (up to 10 times) by allowing only a very few extra relay nodes. Fig. 3 also shows that the quality of the solution improves slightly, in terms of the required number of relay nodes to cover the network, as higher number of initial potential positions of relay nodes are considered.

Fig. 4 shows the number of relay nodes required to cover the network with different number of initial relay nodes positions for 50, 75 and 100 sensors network with $k_s = k_r = 1$ and $k_s = k_r = 2$. In this figure, the legend 50-1-1 (others follow the same convention) indicates a network with 50 sensor nodes, where each sensor nodes are connected with 1 relay node and the relay node network is 1-connected. As shown in the figure, for a given number of possible initial positions and a given connectivity requirement, the number of relay nodes required to cover the network increases with the number of sensor nodes in the distribution.

Unlike existing solutions, our formulation does not require the same value for both k_r and k_s. The 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. The results for different values of k_r and k_s, on 100 sensor nodes network, is given in Fig. 5 (the legend follows the convention of Fig. 4). In our experiment, for the 48 initial relay positions, no solutions exists for a desired value of $k_s = 3, k_r = 2$. As shown in the figure, the required number of relay nodes increases with the higher value of desired connectivity and coverage, which was expected.

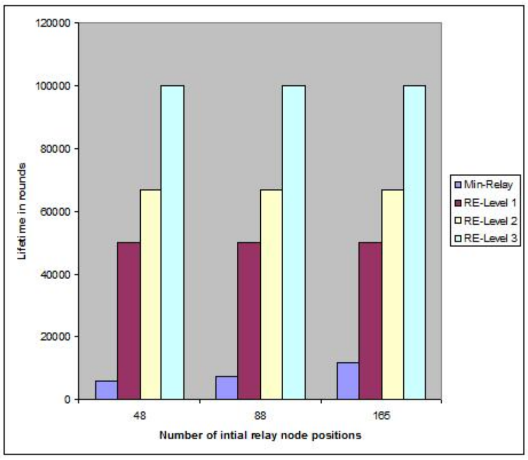

Figure 2: Variation of the network lifetime with the # of initial potential positions of relay nodes.

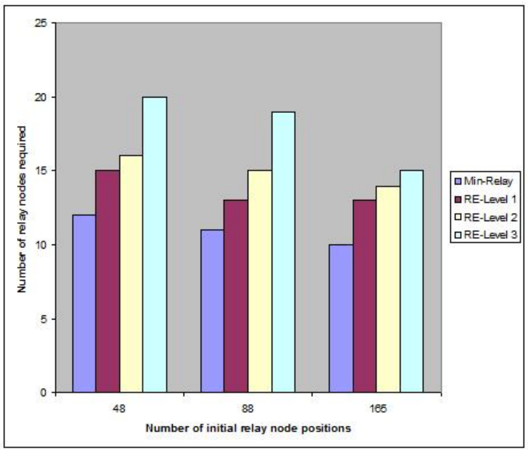

Figure 3: Variation of the # of required relay nodes with the # of initial potential positions of relay nodes.

5. CONCLUSIONS

In this paper we have investigated the problem of designing sensor networks with specified performance guarantees with respect to coverage, connectivity and energy dissipation. We have presented a novel ILP formulation that determines the number and positions of the relay nodes to ensure each sensor node is covered by at least k_s relay nodes, and the relay node network is k_r-connected. We further ensure that a specified network lifetime can be achieved by constraining the energy dissipation of all relay nodes to be below a given value. Unlike previous formulations, which focus primarily on finding a suitable *placement* of relay nodes, our approach also determines an appropriate routing scheme that reduces the energy dissipation of the critical node(s). Experimental results demonstrate that our approach can be used to significantly increase network lifetime, and improve fault tolerance at the cost of a few additional relay nodes. Our ILP is able to generate optimal solutions for networks with hundreds of sensor nodes. We are currently extending our approach so that it can be used in a distributed environment with much larger networks consisting of thousands of sensor nodes.

6. ACKNOWLEDGMENTS

The work of A. Jaekel has been supported by research grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

7. REFERENCES

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

[2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan. Energy effcient communication protocol for wireless micro-sensor networks. In *HICSS*, pp. 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*, pp. 286-299, 2003.

[4] K. Dasgupta, M. Kukreja and K. Kalpaki.

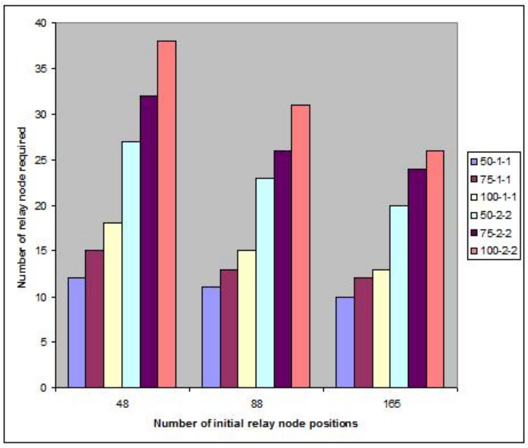

Figure 4: Variation of the # of relay nodes with the # of initial positions of the relay nodes, for $k_s = k_r = 1$ and $k_s = k_r = 2$.

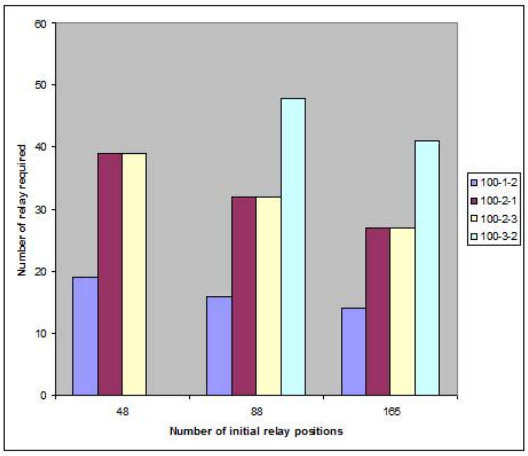

Figure 5: Variation of the # of relay nodes with the # of initial positions of the relay nodes, for different values of k_s and k_r.

Topology-aware placement and role assignment for energy-efficient information gathering in sensor networks, In *IEEE ISCC*, pp. 341-348, 2003.

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

[6] E. Falck, P. Floren, P. Kaski, J. Kohonen, and P. Orponen. Balanced data gathering in energy-constrained sensor networks. *In ALGOSENSORS 2004*, LNCS vol. 3121 pp. 59-70, Springer, 2004.

[7] G. Gupta and M. Younis. Load-balanced clustering of wireless sensor networks. *In IEEE ICC 03*, vol. 3, pp. 1848-1852, 2003.

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

[9] Y. T. Hou, Y. Shi, J. Pan, and S. F. Midkiff. Lifetime-optimal data routing in wireless sensor networks without flow splitting. In *IEEE Basenets*, 2004.

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

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

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

[13] A. Bari, A. Jaekel and S. Bandyopadhyay. Placement and Routing of Relay Nodes in Two-Tiered Sensor Networks, In *ISBAT*, 2006.

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

[15] G. Gupta and M. Younis. Performance evaluation of load-balanced clustering of wireless sensor networks. In *ICT*, vol. 2, pp. 1577-1583, 2003.

[16] 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.

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

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

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

[20] 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.

[21] H. Liu, and P. Wan and W. Jia. Fault-Tolerant Relay Node Placement in Wireless Sensor Networks. In *COCOON*, 2005.

[22] S. C. Ergen and P. Varaiya. Optimal placement of relay nodes for energy efficiency in sensor networks. In *ICC*, 2006.

[23] M. Patel, R. Chandrasekaran and S. Venkatesan. Energy Efficient Sensor, Relay and Base Station Placements for Coverage, Connectivity and Routing. In *IEEE IPCCC*, pages 581 - 586, 2005.

[24] M. Borghini, F. Cuomo, T. Melodia, U. Monaco and F. Ricciato. Optimal data delivery in wireless sensor networks in the energy and latency domains. *WICON*, pp. 138-145, 2005.

[25] M. Yarvis, N. Kushalnagar, H. Singh, A. Rangarajan, Y. Liu nad S. Singh. Exploiting heterogeneity in sensor networks. In *INFOCOM* 2005, vol. 2, pp. 878-890, 2005.

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

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