Optimal Relay Node Placement and Trajectory Computation in Sensor Networks with Mobile Data Collector*

Ataul Bari, Fangyun Luo, Will Froese, and Arunita Jaekel

University of Windsor, Windsor, ON, N9B 3P4, Canada
{bari1,luoh,froesew,arunita}@uwindsor.ca

Abstract. Most sensor network architectures typically assume that nodes are stationary after deployment. However, a number of recent papers have shown that the use of mobile nodes or mobile data collectors (MDC) can significantly improve the performance of a network. In this model, the network can be viewed as a three-tier architecture, where the lowest-tier consists of a set of sensor nodes. The middle-tier contains a number of higher powered relay nodes, each acting as a cluster head for a number of sensor nodes in the tier below, and one or more mobile data collector(s), constitute the upper-tier. For such hierarchical architectures, there are a number of important design problems such as determining the number of relay nodes that are needed and their locations, determining the appropriate buffer capacities in the relay nodes to ensure there is no data loss due to buffer overflow and calculating a suitable trajectory for each MDC. In this paper, we first propose an integrated integer linear program (ILP) formulation that calculates the optimal number and positions of the relay nodes in the middle-tier, along with the requisite buffer sizes. We then present an algorithm for calculating the trajectory of the MDC, based on the relay node locations and the load on each individual relay node, in a way that minimizes the maximum energy dissipation of the relay nodes. Experimental results demonstrate that our approach is feasible for networks with hundreds of sensor nodes.

1 Introduction

The performance benefits of hierarchical sensor networks [1], [6] has been well investigated in the literature. In such networks, sensor nodes are grouped into clusters and transmit their data to their respective cluster heads. The cluster heads collect the data from all nodes in their own cluster and transmit their data to the base station(s), using an appropriate routing scheme. Since the cluster heads are required to transmit large amounts of data over longer distances, compared to individual sensor nodes, the use of specialized nodes for cluster heads has gained considerable support in recent years [1], [3], [5], [9]. These specialized nodes, often called *relay nodes*, are typically equipped with enhanced capabilities in terms of energy provisioning, buffer capacities and transmission

---

* A. Jaekel has been supported by discovery grants from the Natural Sciences and Engineering Research Council of Canada.

J. Zheng, D. Simplot-Ryl, V.C.M. Leung (Eds.): ADHOCNETS 2010, LNICST 49, pp. 400–415, 2010.

ranges. The resulting architecture, where a large number of low-power, sensor nodes with limited capabilities form the lower-tier and relatively fewer relay nodes with enhanced capabilities form the upper-tier, has been shown to improve network performance in a number of areas including network lifetime, load-balanced routing and fault-tolerance [1] – [7].

For the above two-tier architecture, the lifetime of the network is primarily determined by the lifetime of the upper-tier relay node network. Each relay node is responsible for receiving (and possibly aggregating) the data from all sensor nodes in its cluster and then transmitting the data to (or towards) the base station, using either single-hop or multi-hop paths [6], [8], [9]. The energy dissipation of the relay nodes increases rapidly with the distance between the sender and the receiver (either another relay node or the base station), and has a significant impact on the lifetime of the network. A number of energy-aware routing strategies have been proposed to extend the lifetime of the relay node network [6], [8], [9]. However, such strategies are of limited use for relay nodes that are far away from other nodes and must therefore transmit over a large distance, or for nodes near the base station that must transmit data from many other nodes, in case of multi-hop routing. In fact, for sparse networks it is even possible that no feasible routing scheme exists, since the distance to the nearest neighbor may be greater than the transmission distance of a node.

A number of recent papers have shown that the use of some mobile nodes or mobile data collectors (MDC) can significantly improve the performance of a network in terms of lifetime, coverage, and connectivity [15], and techniques for effectively utilizing the unique capabilities of mobile nodes have been attracting increasing research attention in the past few years [19] – [25]. In this paper, we consider a new network model that extends the traditional two-tier architecture, by adding a third tier consisting of one (or possibly more) mobile data collector(s) (MDC), above the relay node network (which now constitutes the middle-tier). The MDC, which is not power constrained, visits all relay nodes in the middle-tier, following a fixed trajectory [15], collects data from them, and delivers the collected data to the base station. Thus, the relay nodes are relieved from the burden of "routing" data towards the base station, possibly over long distances, resulting in considerable energy savings at these nodes.

In such a three-tier architecture, we assume that the lower-tier sensor nodes have already been deployed, and the number and locations of these sensor nodes have been determined by the monitoring needs of the specific application. Therefore, our goal is to design the middle and the upper tiers of the network. In this context, we address the following design problems in this paper:

i) Find the locations of the relay nodes constituting the middle-tier, such that each sensor node is *covered* by at least one relay node (i.e. there is at least one relay node within the transmission range of each sensor node), and the number of the relay nodes is minimized. This is the *relay node placement* problem.
ii) Determine the set of sensor nodes assigned to the cluster of each relay node, such that the overall buffer requirements for the relay nodes is minimized.

iii) Determine a suitable trajectory for the MDC(s), such that the energy dissipation rate of the relay nodes is minimized (or is below a specified level). Restricting the energy dissipation allows the network to remain functional for a specified minimum period of time.

In this paper we first propose an integer linear program (ILP) formulation that, given a set of potential locations of relay nodes in a network, optimally solves the relay node placement problem. Our formulation also computes the buffer requirements for the relay nodes, so that data generated by the lower-tier sensor nodes can be stored at the relay nodes and delivered to the MDC without any loss of data (i.e. without buffer overflow). We also provide a modification of our ILP that, given the maximum number of relay nodes to be used in the middle-tier, finds the locations of the relay nodes and minimizes their buffer requirements. Once the locations of the relay nodes has been determined, we present an algorithm for calculating the trajectory of the MDC, such that the energy dissipation of the relay nodes due to data transmission is minimized.

The remainder of the paper is organized as follows. In section 2, we review the relevant work. In sections 3 and 4, we present our ILP formulations and algorithm for trajectory computation respectively. We discuss our experimental results in section 5 and conclude in section 6.

2 Review

2.1 Relay Nodes Placement, Clustering and Routing

The relay node placement problem deals with placing a minimum number of relay node such that each sensor node can send its data to at least one relay node. This is an NP-hard problem [3], and has been widely investigated in the literature [3], [7], [10], [11], [13] and [14]. In [3], the authors solve the placement problem by dividing the entire region into cells, finding an optimal solution for each cell, and then combining the solutions. They consider *single* and *double* connectivity problem, both at the sensor level, as well as, at the relay node network level. An approximation algorithm to achieve single and double connected network is proposed in [10]. In [11], the authors propose 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. A mixed-integer *non-linear* program, and a heuristic algorithm in proposed in [7], that focuses on prolonging the lifetime of sensor networks by deploying relay nodes within the networks.

The network connectivity and the fault-tolerance is also addressed in [13], where the authors propose an ILP-based approach for the appropriate placement of the relay nodes in a sensor network. This approach optimally selects the minimum number of relay nodes from a given set of “potential positions” of the relay nodes, obtained from the area of the network deployment. Noting that the number of potential positions in a real plane can be infinite, the authors use a heuristic to limit such locations to a level so that the ILP becomes computationally tractable. This heuristic views the entire networking area as an imaginary *grid*, and the potential relay node locations are taken at the center and the corner positions of each cell. The authors shows that such an approach provides good solutions, especially when the network area is small and the sensor nodes are densely deployed. They also shows that a finer grid leads to a better solution.

The clustering and the routing schemes, for two-tiered networks, are addressed in a number of recent papers, including those in [1], [2], [6], [9], [12]. In [1], the authors focus on load balanced clustering and propose a heuristic solution for the problem. Fault tolerant clustering is addressed in [2], and in [12], the authors investigate the problem of maximizing network lifetime by appropriately placing nodes, which are not energy constrained (e.g., connected to a wall outlet). A *Tenet* architecture for tiered sensor networks is proposed in [4] that simplifies application development and reuses mote-tier software. In [5], [6] and [9], the authors focus on clustering and routing schemes that maximize the network lifetime. These approaches do not consider the problem of minimizing the number of relay nodes and finding their locations.

2.2 Mobility in Sensor Networks

Improving the performance of sensor networks, exploiting some mobile entities, has been investigated in [19] – [25]. In [19], multiple mobile base stations, which might be required to move periodically, is considered to prolong the network lifetime. The authors propose an ILP formulation to compute the new locations of the mobile base stations. In [20], multiple mobile base stations, which move in parallel straight paths, are considered, and is shown that such mobility can achieve scalability and load-balancing. In [26], the authors focus on optimal data collection, and propose a protocol that extends the lifetime of the network. Their approach takes into account both the base station mobility and the multi-hop routing. In [22], the authors focus on fault-tolerance. They propose a scheme that uses K-means clustering and TSP-based shortest path approaches to achieve resiliency. In [23], a three-tiered network is considered where mobile data collectors, lie in the middle-tier, move randomly within the network and pick-up data from the sensor nodes. A sensor node transmits data only when a MDC enters the direct communication range of the node. In [24], the authors propose a queuing theory based mathematical model that analyzes the performance and trade-offs of the three-tier architecture. Partitioning Based Scheduling (PBS) heuristic that computes the trajectory of MDC is used in [25]. The focus here is to reduce sensor buffer overflow. They propose a solution that has two parts. In the first part, nodes are partitioned into groups based on their locations and the data generation rates. Then, a node visiting schedule is generated within a group. Finally, the solutions of these groups are combined to obtain the final path of the MDC. The problem of delivering data to multiple mobile sinks is addressed in [26], [27] and [28].

A hierarchical sensor network architecture, using higher powered relay nodes as cluster heads, which utilizes a MDC to collect data from the cluster heads has been discussed in [17], and [18]. The objective, in both works, is to extend the overall lifetime of the network. In [17], an ILP formulation is proposed, which optimally selects the relay nodes, from a set of potential relay nodes positions, such that i) the number of relay nodes are minimal, and ii) the length of the trajectory of the MDC is as short as possible. In [18], the authors focus on reducing the length of the trajectory of the MDC, by allowing the MDC to visit the *neighborhood* of each relay nodes, instead of visiting their exact locations. The authors propose two heuristic solutions for the problem.

2.3 Network Power Model

The power needed for data communication is the dominant factor in power consumption in sensor networks. We have considered the first-order radio model [8] to account for the energy consumption due to communication where receive (transmit) circuitry consumes $\alpha_1\ nJ/bit$ ($\alpha_2\ nJ/bit$) 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*, $2 \leq q \leq 4$ [8] 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 value of path-loss exponent, $q = 2$.

3 Network Design Exploiting Node Mobility

3.1 Network Model

We consider a three-tiered wireless sensor network, where the lower-tier consists of a set $\mathcal{S}$ of n sensor nodes, i.e., $|\mathcal{S}| = n$. We assume that the deployment of the sensor nodes ensures the appropriate coverage of the sensing area. We consider a set $\mathcal{R}$ of m potential locations of relay nodes, i.e., $|\mathcal{R}| = m$. A subset $\mathbb{R}$ of $\mathcal{R}$ will constitute the middle-tier network. Each element of $\mathbb{R}$ would act as a cluster head. We also consider a MDC, lying in the upper-tier of the network. The MDC visits all relay nodes in $\mathbb{R}$, collects their data, and delivers the data to the base station. Each relay node has a buffer size $\mathcal{B}$. We assign, to each node, a unique label as follows:

i) for each sensor node, a label $i, 1 \leq i \leq n$,
ii) for each possible location of the relay nodes, a label $j, n+1 \leq j \leq n+m$, and
iii) for the MDC, a label $n+m+1$.

If a sensor node $i \in \mathcal{S}$ is covered by a relay node at location j (we shall refer to such relay node as *relay node* j), then i can transmit its data directly to j. A sensor node i may be covered by more than one relay node, however, our objective is to design the relay node network such that each sensor node belongs to exactly one cluster, $\mathcal{C}^j$, corresponding to a relay node j. Our proposed formulation

i) determines the minimum number and the positions of the relay nodes, which will act as the cluster heads, and form the middle-tier network, and
ii) assign sensor nodes to clusters such that the relay nodes buffer requirements are minimized.

We assume that the positions of the sensor nodes are known (or can be determined, e.g. using GPS), and the relay nodes can be placed at the computed locations. This approach is feasible for many applications (e.g., monitoring industrial environments, road condition, and habitat). We have used a grid based approach [13] to generate $\mathcal{R}$. However, our ILP formulation does not depend on how $\mathcal{R}$ is generated and other approaches such as approach given in [3] can easily be used.

We consider that a sensor node $i \in \mathcal{S}$ generates data at a rate of b_i bits per unit time, and transmits to the corresponding cluster head. The value of $b_i, \forall i \in \mathcal{S}$ can either be the same, or may vary. In our model, each relay node j, receives data from the sensor nodes belonging to its own cluster $\mathcal{C}^j$, and buffers them until j can transmit buffered data to the MDC, while it is visiting j. Data buffering is essential for applications where it is important not to lose any data generated by the sensor nodes. In our model, the MDC visits each relay node j, periodically, at fixed time intervals. Once transmitted, the buffer of j is cleared so that the buffer can be reused to store data until the next visit by the MDC. A relay node j transmits its buffered data only when the MDC is closest to j, in its trajectory. The MDC traverses at a constant speed following a predetermined trajectory, and it needs T_r time units to complete the trajectory. That is, the time interval between any two successive visits by a MDC to a relay node j is known and is equal to T_r.

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 label $i, 1 \leq i \leq n$.
- m: The total number of possible positions of relay nodes, each position having a unique label $j, n+1 \leq j \leq n+m$.
- j: The relay node at location j, $n+1 \leq j \leq n+m$.
- $n+m+1$: The label of the MDC.
- r_{max}: The transmission range of each sensor node.
- $d_{i,j}$: The Euclidean distance between node i and node j.
- b_i: Number of bits generated by sensor node i in unit time.
- $\mathcal{C}^j$: The set of sensor nodes belonging to the cluster of relay node j.
- W_1, W_2: Positive constants.
- T_r: Time required by the MDC between two successive visits at any relay node j.
- y_{max}: Maximum allowable number of relay nodes.

We also define the following variables:

- $X_{i,j}$: Binary variable defined as follows:
  $$X_{i,j} = \begin{cases} 1 \text{ if the sensor node } i \text{ selects} \\ \quad \text{relay node } j \text{ as its cluster head,} \\ 0 \text{ otherwise.} \end{cases}$$
- Y_j: Binary variable defined as follows:
  $$Y_j = \begin{cases} 1 \text{ if relay node at location } j \text{ is included} \\ \quad \text{in the middle-tier network,} \\ 0 \text{ otherwise.} \end{cases}$$
- R_j: Continuous variable indicating the total number of bits generated (during the period T_r) by the sensor nodes belonging to the cluster $\mathcal{C}^j$. of the relay node j.
- $\mathcal{B}_{max}$: A Continuous variable such that $R_j \leq \mathcal{B}_{max}, \forall j,\ n{+}1 \leq j \leq n{+}m$.

3.3 ILP Formulation for Minimizing the Number of Relay Nodes

In this section, we propose our ILP formulation. Our formulation

i) ensures that each sensor node is covered by at least one relay node,
ii) minimizes the number of relay nodes, and
iii) minimizes the maximum buffer capacity of any relay node j, which is included in the middle-tier, as a secondary objective.

Using the notations discussed in section 3.2, we present our formulation as follows:

$$\textbf{Minimize } W_1 \cdot \sum_{j=n+1}^{n+m} Y_j + W_2 \cdot \mathcal{B}_{max} \tag{1}$$

Subject to:

a) A sensor node i can transmit data 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) The relay node at location j must be included in the middle-tier network, if it is selected as a possible cluster head by at least one sensor node i.

$$Y_j \geq X_{i,j} \qquad \forall i, 1 \leq i \leq n, \quad \forall j, n+1 \leq j \leq n+m \tag{3}$$

c) A sensor node transmits data to exactly one relay node.

$$\sum_{j=n+1}^{n+m} X_{i,j} = 1 \qquad \forall i, 1 \leq i \leq n \tag{4}$$

d) Compute the total number of bits buffered at the relay node j during the interval T_r.

$$R_j = T_r \cdot \sum_{i=1}^{n} b_i \cdot X_{i,j} \qquad \begin{array}{l} \forall i, 1 \leq i \leq n, \\ \forall j, n+1 \leq j \leq n+m \end{array} \tag{5}$$

e) The total amount of data to be buffered by any relay node in one round of data gathering cannot exceed $\mathcal{B}_{max}$.

$$R_j \leq \mathcal{B}_{max} \qquad \forall j, n+1 \leq j \leq n+m \tag{6}$$

3.4 Justification of the ILP Equations

Equation (1) is the objective function for the formulation, and consists of two terms. The primary goal (represented by the first term) is to minimize the total number of relay nodes used to form the middle-tier network. As mentioned earlier, a relay node j is included in the middle-tier network (i.e., $Y_j = 1$), only if j selected as a cluster head by at least one sensor node i. Therefore, by counting the number of relay nodes selected to be the cluster heads, we can determine the number of relay nodes needed in the middle-tier network. This is exactly the value calculated by the first term in the objective function. The second term is used to minimize the maximum buffer capacity of the relay nodes, which is the secondary objectives. By choosing appropriate values for W_1 and W_2, we can select the relative importance of the two objectives being minimized. For example, if we set $W_2 = 0$, then the *only* parameter we are interested in minimizing is the number of relay nodes.

a) Constraint (2) specifies that a sensor node can communicate with a relay node j, only if j is within the transmission range of the sensor node.
b) Constraint (3) ensures that if a relay node j is chosen as a cluster head by at least one sensor node, then j must be included in the middle-tier network. If a relay node j is not selected to be a cluster head by any sensor node, normally it should not be selected. 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 other constraints.
c) Constraint (4) requires that each sensor node belongs to exactly one cluster and transmits data to the corresponding cluster head.
d) If a relay node j is selected to be included in the middle-tier and a sensor node i belongs to its cluster $\mathcal{C}^j$, then $X_{i,j} = 1$. Constraint (5) thus calculates the total number of bits, R_j, buffered in j during the interval T_r, by summing the data transmitted to it from all the sensor nodes belonging to the cluster $\mathcal{C}^j$ and then multiplying this value by the interval T_r.
e) Constraint (6) ensures that if a relay node j is selected, then the total bits buffered at j during the interval T_r do not exceed the buffer size $\mathcal{B}_{max}$. The left hand side of constraint (6) is the total amount of bits buffered at j during the interval T_r. The right hand side $\mathcal{B}_{max}$, of constraint (6), must be greater

than or equal to the maximum of the amount of bits buffered by the relay nodes. Since the objective function is to minimize $\mathcal{B}_{max}$, constraint (6) forces $\mathcal{B}_{max}$ to be the maximum buffer capacity required by any relay node.

3.5 ILP Formulation for Minimizing the Buffer Size of Relay Nodes

The ILP formulation given in Section 3.3 minimizes the total number of relay nodes required to form the middle-tier network, and considers the buffer size as a secondary objective. This gives the lower bound for the number of relay nodes. However, it some cases, the total number of relay nodes may be given, and the problem is to find out the placement of the given number of relay nodes such that the maximum size of the buffer is minimized. This can be easily achieved by the following modification of the ILP formulation, given in Section 3.3.

$$\textbf{Minimize } \mathcal{B}_{max} \tag{7}$$

Subject to:

a) - e) Constraint (2) - Constraint (6).
b) Maximum number of relay nodes cannot exceeds y_{max}.

$$\sum_{j=n+1}^{n+m} Y_j \leq y_{max} \qquad \forall j, n+1 \leq j \leq n+m \tag{8}$$

Equation (7) is the objective function that minimizes the maximum buffer capacity requirement of any relay node. Constraint (8) enforces the limit on the maximum number of relay nodes.

4 Computation of Trajectory

In this section, we present a heuristic approach to compute a trajectory for the MDC, such that the maximum energy dissipated by any relay node is minimized. A number of papers have considered the use of complex trajectories, where the MDC visits each node individually [25], [17], [18]. However, in this paper the goal is to use a very simple trajectory that can be easily traversed by the MDC. So, we consider the case where the MDC travels back and forth along a straight line. A straight line trajectory has been shown to be a practical and useful option for mobile nodes in [20].

Given the positions and the expected loads of the relay nodes, our approach finds a straight line, to be used by the MDC as the path to collect the data from the relay nodes, such that the energy dissipation by the relay nodes are minimized. We note that the energy minimization can be accounted as the sum of the energy dissipated by the relay nodes. In such case, any standard weighted-regression analysis method, using the communication cost as discussed in Section 2.3, can be applied to compute the best fitting trajectory. However, in our network model, if any relay node depletes power, then all sensor nodes belong to the relay node become inaccessible, and the network may fail to meet the reliability standard. In this case, to extend the lifetime, it is important to minimize the

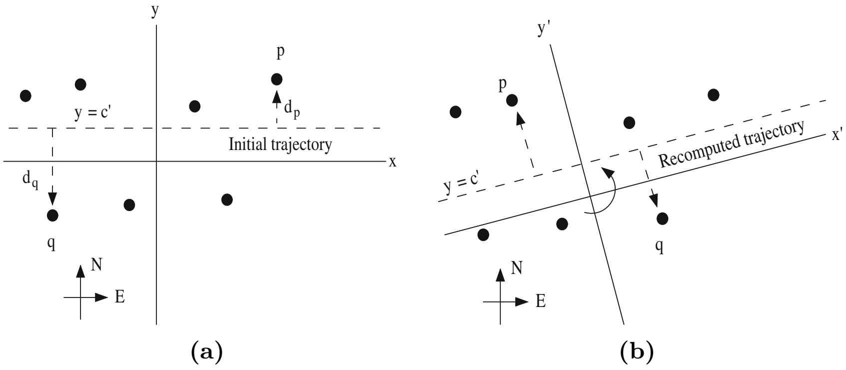

Fig. 1. Computation of the (a) initial trajectory, (b) improved trajectory by rotating the orientation of the axis

maximum energy dissipated by any relay node in the network. We propose a simple approach that can be used to compute a suitable straight line trajectory.

Let $\mathbb{R}$ be the set of relay nodes, where each relay node is given an unique label $j, 1 \leq j \leq |\mathbb{R}|$. Also, let the coordinate (j_x, j_y) specify the position of relay node $j, \forall j \in \mathbb{R}$. We start by assuming that, given the area of the network, the trajectory is a horizontal line. Let the line be given by the equation $y = c$, where c is a constant. We set the initial value of c as the midpoint of the y-coordinate values of the uppermost and the lowermost relay nodes in the network. Since a relay node j transmits to the MDC when the MDC is closest to j, the transmission distance of j is the vertical distance (i.e., y-axis distance) of j, projected on the trajectory line $y = c$. Using this initial trajectory, we find the relay node p (q) that dissipates the maximum amount of energy, among all nodes located in the above (below) the initial line. To minimize the maximum energy of the relay nodes, we need to find a new value $c', q_y \leq c' \leq p_y$, for the constant c, so that the energy dissipation of nodes p and q is balanced (Fig. 1(a)). We achieve this by setting the energy dissipation of nodes p and q (computed using the model discussed in Section 2.3), corresponding to the trajectory $y = c'$, as equal. Let the vertical distance of the node p and the node q, from the new trajectory be d_p and d_q, respectively. Also, let the vertical distance between nodes p and q be λ. Then, using the equation given in Section 2.3, and the notation given in Section 3.2, we have:

$$d_p^2 - \gamma d_q^2 = \xi \tag{9}$$

$$d_p + d_q = \lambda \tag{10}$$

Where $\gamma = \frac{R_q}{R_p}$, and $\xi = \frac{1}{\beta}(\alpha_1 + \alpha_2)(\gamma - 1)$. The values of α_1, α_2 and β are obtained from network power model discussed in Section 2.3. We obtain the new value of $c' (= q_y + d_q)$ by solving the above two equations. Using the new line, we recompute p and q, and apply the process, in an iterative manner, until the difference between the energy dissipation by p and q becomes less than a small preset value.

Based on the actual layout of the networking area and the distribution of the sensor nodes, a different orientation, rather than strictly horizontal, for the trajectory may be beneficial. Once we obtain the initial trajectory, we compute the best orientation of the trajectory by rotating it in the range $0° - 180°$. We rotate the line by a small angle ψ, at a time, and get a new orientation, as shown in Fig. 1(b). At each orientation, we recompute the value of E_{max}, using the approach described above. After the rotation is complete, we select the orientation that gives the minimum among all orientations. As shown in the Section 5.2, this rotation can substantially improve the solution, based on the actual layout of the network.

5 Results

5.1 Simulation of ILP Formulation

In this section, we present the simulation results of our formulation for selecting the relay nodes in the middle-tier of the network. We have conducted different sets of experiments by setting different values for the parameters in our formulation. In the first set of experiments, our objective is to jointly optimize the number of relay nodes required to form the middle-tier relay node network, and the maximum buffer requirement of the relay nodes. The relative importance of each term is determined by the value of the constant W_1 and W_2, used in equation (1). Since our primary goal is to minimize the number of relay nodes, while the secondary objective is to reduce the buffer requirement of each node, we set $W_1 = 8000$ and $W_2 = 0.1$ for our simulations. We have used an experimental setup, where the sensor nodes are randomly distributed over a $200 \times 280 m^2$ area. We have assumed that $r_{max} = 40m$. The results are obtained by CPLEX 9.1 solver.

We have simulated our scheme with different number of sensor nodes, ranging from 100—600. For each size of the sensor node network, we randomly generate five different sets for the locations of the sensor nodes in the network, and compute the results using each set. The results reported in the tables and figures in this section reflect the *averages* of all the different runs for each network size. As in [13], we have used a grid based approach to compute the initial potential positions of the relay nodes. The number of potential relay node locations were set to 48 (for coarse grid) and to 165 (for fine grid), indicated as 48-Grid and 165-Grid respectively in the following discussions of our results. We have also assumed that each sensor node generates data at a rate of 100 bits/unit-time, i.e., $b_i = 100, \forall i, 1 \leq i \leq n$ and $T_r = 5$.

Fig. 2 shows the number of relay nodes needed in the middle-tier for different number of sensor nodes distributions. We note that for the same distribution, using 165-grid (fine grid) consistently leads to better solutions compared to 48-grid. It is also interesting to note that, although the number of relay nodes required increases with the number of sensor nodes, the rate of increase is not very high. For example using the 165-grid only a few additional relay nodes are required to cover 600 sensor nodes, as compared to 100 sensor nodes.

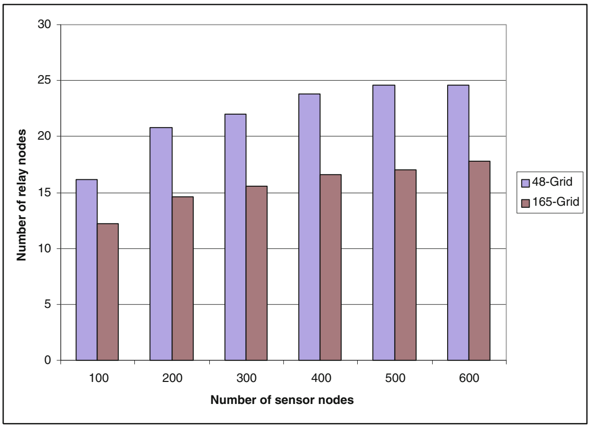

Fig. 2. The number of relay nodes required to form the middle-tier network for different number of sensor nodes

Next we consider the buffer requirements of the relay nodes selected for the middle-tier. Fig. 3 shows the value of the maximum buffer size ($\mathcal{B}_{max}$) calculated by our ILP using 48-grid and 165-grid configurations. At first glance, the figure seems to indicate that 48-grid produces better results (i.e. lower buffer size) compared to 165-grid. However, we must remember that the 48-grid configuration requires a higher number of relay nodes. This means that the same amount of data is distributed over more relay nodes, resulting in a lower buffer requirement *per node*. But, when we compare the total buffer requirements (as shown in Fig. 4), we see that the 165-grid generates better results, both in terms of the number of relay nodes and the total buffer size.

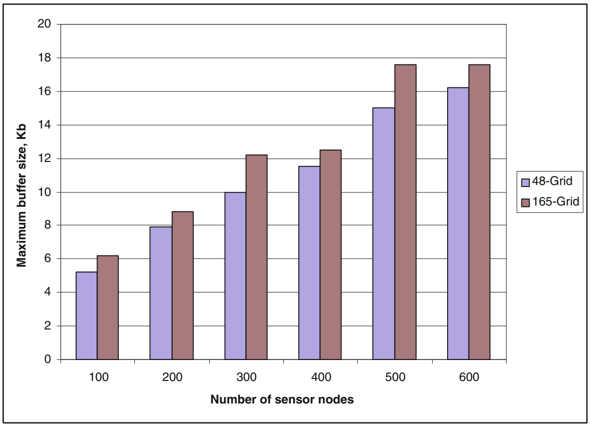

Fig. 3. The maximum buffer capacity per node required for different number of sensor nodes

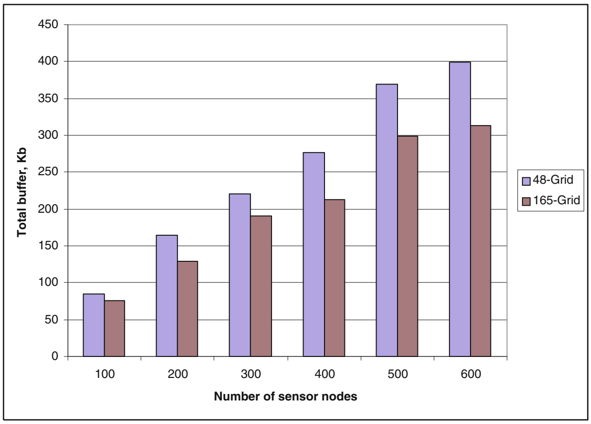

Fig. 4. The total buffer capacity of the relay nodes required for different number of sensor nodes

5.2 Simulation of Trajectory Computation Algorithm

The goal of our trajectory computation algorithm is to calculate the trajectory of the MDC (along a straight line), such that the maximum energy dissipation (E_{max}) of any relay node is minimized. Fig. 5 shows the average value of E_{max}, for different size of sensor node networks, corresponding to the trajectory that minimizes the value of E_{max}, for each configuration. As before, we note that although the value of E_{max} appears to be lower for 48-grid, this is because it requires more relay nodes resulting in lower energy dissipation per node. As expected, the value of E_{max} increases steadily with the number of sensor nodes

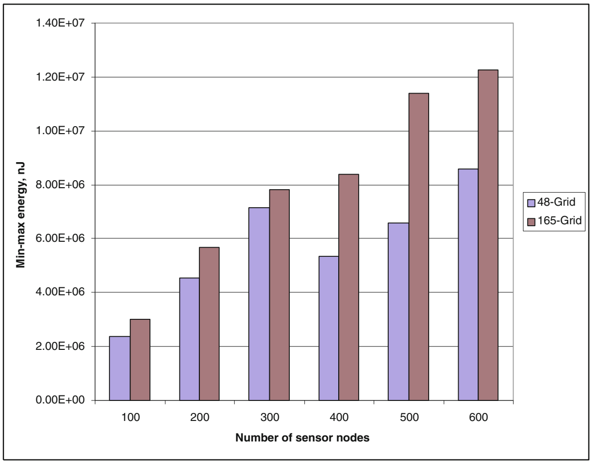

Fig. 5. The minimum of the maximum energy dissipation by the relay nodes in the networks with different number of sensor nodes

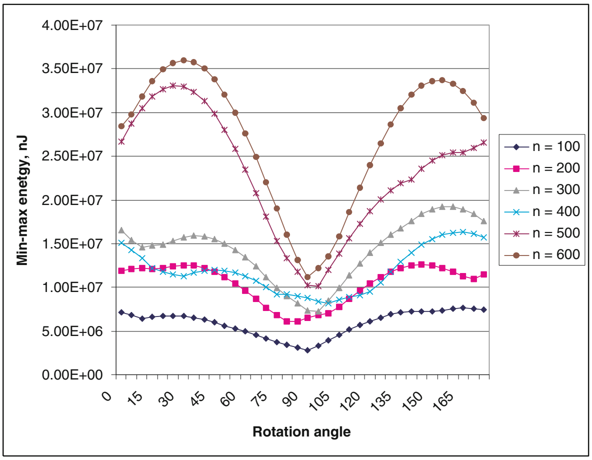

Fig. 6. Variation of the minimum of the maximum energy dissipation by the relay nodes with the rotation of the axis, in the networks with different number of sensor nodes with grid setting 165-Grid

for 165-grid case. However, for the 48-grid case, we notice an anomalous case, where the value of E_{max} for 300 sensor nodes is actually higher than that for both 400 and 500 sensor node distributions. This is because, the performance of coarse grid configurations is not always reliable and may sometimes fail to find a good solution (e.g. for the 48-grid and 300 sensor node case). On the other hand, when we use finer grids (e.g. 165-grid), the computational complexity increases but the we get more consistent and reliable solutions.

Finally, Fig. 6 shows how the value of E_{max} varies with the angle of the straight line trajectory for the MDC. In general, the angle at which the value of E_{max} is minimized will depend on the distribution of the sensor node and the shape of sensing area. In our experiments the sensing area was a rectangular shape ($200m$ along x-axis and $280m$ along y-axis), and the sensor nodes were randomly distributed in the sensing field. Therefore, we can expect that the best trajectory will be a (nearly) vertical line. This is exactly what we find in Fig. 6, where the minimum value of E_{max} is obtained at an angle of about 90° for each sensor node distribution. We also note that the value of E_{max} varies widely with the angle for higher values of n, but as n decreases, these variations are greatly reduced.

6 Conclusions

In this paper, we have proposed a new formulation that, given a set of potential locations of relay nodes, optimally determines the minimum number of relay nodes, along with their locations, in a hierarchical sensor network, which includes a MDC that travels along a fixed trajectory. The placement is done in such a way that i) each sensor node is covered by at least 1 relay node, ii) no relay node suffers from the buffer overflow, and iii) maximum buffer requirement of the relay nodes is minimized. Our ILP is able to generate optimal solutions for networks with hundreds of sensor nodes. We have also proposed a new heuristic for calculating a straight line trajectory for the MDC that minimizes the maximum energy dissipation of the relay nodes.

Although, we have focussed on straight line trajectories in this paper, our approach can be adapted to consider other simple trajectories, such as a rectangular or circular path. We are currently extending our approach to consider such trajectories and to determine the most suitable path, based on different layouts of sensing areas and different node distributions.

References

1. Gupta, G., Younis, M.: Load-balanced clustering of wireless sensor networks. In: IEEE ICC, vol. 3, pp. 1848–1852 (2003)
2. Gupta, G., Younis, M.: Fault-tolerant clustering of wireless sensor networks. In: IEEE WCNC, pp. 1579–1584 (2003)
3. Tang, J., Hao, B., Sen, A.: Relay node placement in large scale wireless sensor networks. Computer Communication 29(4), 490–501 (2006)
4. Gnawali, O., Greenstein, B., Jang, K.-Y., Joki, A., Paek, J., Vieira, M., Estrin, D., Govindan, R., Kohler, E.: The Tenet Architecture for Tiered Sensor Networks. In: The Proceeding of SenSys (2006)
5. Bari, A., Jaekel, A., Bandyopadhyay, S.: Clustering Strategies for Improving the Lifetime of Two-Tiered Sensor Networks. Computer Communications 31(14), 3451–3459 (2008)
6. Bari, A., Jaekel, A., Bandyopadhyay, S.: Integrated Clustering and Routing Strategies for Large Scale Sensor Networks. In: Akyildiz, I.F., Sivakumar, R., Ekici, E., de Oliveira, J.C., McNair, J. (eds.) NETWORKING 2007. LNCS, vol. 4479, pp. 143–154. Springer, Heidelberg (2007)
7. Hou, Y.T., Shi, Y., Sherali, H., Midkiff, S.F.: On Energy Provisioning and Relay Node Placement for Wireless Sensor Networks. IEEE Transactions on Wireless Communications 4(5), 2579–2590 (2005)
8. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: Energy effcient communication protocol for wireless micro-sensor networks. In: HICSS, pp. 3005–3014 (2000)
9. Bari, A., Jaekel, A., Bandyopadhyay, S.: Optimal placement and routing strategies for resilient two-tiered sensor networks. In: Wireless Communications and Mobile Computing. Wiley, Chichester (2008), doi:10.1002/wcm.639
10. Hao, B., Tang, J., Xue, G.: Fault-tolerant relay node placement in wireless sensor networks: formulation and approximation. In: HPSR, pp. 246–250 (2004)
11. Liu, H., Wan, P., Jia, W.: Fault-Tolerant Relay Node Placement in Wireless Sensor Networks. In: Wang, L. (ed.) COCOON 2005. LNCS, vol. 3595, pp. 230–239. Springer, Heidelberg (2005)
12. Yarvis, M., Kushalnagar, N., Singh, H., Rangarajan, A., Liu, Y., Singh, S.: Exploiting heterogeneity in sensor networks. In: INFOCOM 2005, vol. 2, pp. 878–890 (2005)
13. Bari, A., Jaekel, A., Bandyopadhyay, S.: Optimal Placement of Relay Nodes in Two-Tiered, Fault Tolerant Sensor Networks. In: IEEE ISCC (2007)

14. Bari, A., Xu, Y., Jaekel, A.: Integrated Placement and Routing of Relay Nodes for Fault-Tolerant Hierarchical Sensor Networks. In: IEEE ICCCN - SN, pp. 1–6 (2008)
15. Bari, A., Jaekel, A.: Techniques for Exploiting Mobility in Wireless Sensor Networks, Handbook of Research on Mobile Business: Technical, Methodological and Social perspective, 2nd edn. IGI Global, Information Science Reference, pp. 445–455 (2008) ISBN: 978-1-60566
16. Pazzi, R.W.N., Boukerche, A.: Mobile data collector strategy for delay-sensitive applications over wireless sensor networks. Computer Communications 31(5), 1028–1039 (2008)
17. Bari, A., Chen, Y., Jaekel, A., Bandyopadhyay, S.: A New Architecture for Hierarchical Sensor Networks with Mobile Data Collectors. In: ICDCN, Networking track (2010)
18. Bari, A., Chen, Y., Roy, D., Jaekel, A., Bandyopadhyay, S.: Energy Aware Trajectory Computation of Mobile Data Collectors in Hierarchical Sensor Networks. In: IEEE ICC (to appear, 2010)
19. Gandham, S.R., Dawande, M., Prakash, R., Venkatesan, S.: Energy efficient schemes for wireless sensor networks with multiple mobile base stations. In: IEEE Global Telecommunication Conference, GLOBECOM 2003, vol. 1, pp. 377–381 (2003)
20. Jea, D., Somasundara, A., Srivastava, M.: Multiple controlled mobile elements (Data Mules) for data collection in sensor networks. In: Prasanna, V.K., Iyengar, S.S., Spirakis, P.G., Welsh, M. (eds.) DCOSS 2005. LNCS, vol. 3560, pp. 244–257. Springer, Heidelberg (2005)
21. Luo, J., Hubaux, J.-P.: Joint mobility and routing for lifetime elongation in wireless sensor networks. In: INFOCOM, vol. 3, pp. 1735–1746 (2005)
22. Nakayamaa, H., Ansarib, N., Jamalipourc, A., Katoa, N.: Fault-resilient sensing in wireless sensor networks. Computer Communications 30(11-12), 2375–2384 (2007)
23. Shah, R.C., Roy, S., Jain, S., Brunette, W.: Data mules: Modeling a Three-Tier Architecture For Sparse Sensor Networks. In: IEEE Workshop on Sensor Network Protocols and Applications (2003)
24. Jain, S., Shah, R., Brunette, W., Borriello, G., Roy, S.: Exploiting Mobility for Energy Efficient Data Collection in Sensor Networks. Mobile Networks and Applications 11(3), 327–339 (2006)
25. Gu, Y., Bozdag, D., Ekici, E., Ozguner, F., Lee, C.-G.: Partitioning-Based Mobile Element Scheduling in Wireless Sensor Networks. In: IEEE Conf. Sensor and Ad Hoc Communication and Network (2005)
26. Luo, H., Ye, F., Cheng, J., Lu, S., Zhang, L.: Ttdd: Two-tier data dissemination in large-scale wireless sensor networks. Wireless Networks 11, 161–175 (2005)
27. Kim, H.S., Abdelzaher, T.F., Kwon, W.H.: Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks. In: Proc. of the 1st Int. Conf. on Embedded Networked Sensor Systems (SENSYS), pp. 193–204 (2003)
28. Hwang, K., In, J., Eom, D.: Distributed dynamic shared tree for minimum energy data aggregation of multiple mobile sinks in wireless sensor networks. In: Römer, K., Karl, H., Mattern, F. (eds.) EWSN 2006. LNCS, vol. 3868, pp. 132–147. Springer, Heidelberg (2006)
29. Pan, J., Hou, Y.T., Cai, L., Shi, Y., Shen, S.X.: Topology Control for Wireless Sensor Networks. In: International Conference on Mobile Computing and Networking, pp. 286–299 (2003)