Energy Efficient Data Collection with Mobile Robots in Heterogeneous Sensor Networks

Jared Goerner, Nilanjan Chakraborty, *Member, IEEE,* and Katia Sycara, *Fellow, IEEE*

Abstract—In this paper, we study the problem of constructing a path for a mobile data collecting robot such that the total data collection cost (i.e., sum of transmission energy of the sensor nodes and movement energy of the robot) in a sensor network is minimized. We assume that the sensor nodes can transmit within a certain region around their position, which is called the *communication set*. We model the communication set as a convex set to take into account asymmetric transmission systems (like directional antennas). We derive a necessary condition for the optimality of a mobile robot tour through the communication sets. Based on this condition, we design a three-step approach to compute a local minimum of the optimization problem. We prove that our solution is guaranteed to be within a constant factor of the global optimal solution. Our algorithm works for both 2-dimensional and 3-dimensional sensor networks where the sensor nodes are heterogeneous and can have directional communication properties. In contrast, existing algorithms for computing data collecting routes are for planar sensor networks and assume the communication sets to be discs. We also present simulation results depicting the performance of our algorithm.

Index Terms—Sensor network, Data collection, Path planning.

I. INTRODUCTION

Spatially distributed wireless sensor networks are increasingly finding applications in areas like environmental monitoring [18], [28] and surveillance [15]. In such sensor networks, aggregating the data collected by the sensor nodes and transferring it to a base station is an important problem. The use of mobile robots (or mobile nodes) to collect data from static sensor nodes has been proposed in the literature as it can be more energy efficient [27], [25]. Moreover, it also reduces *hot spot* problems [23] prevalent in static multi-hop sensor networks, wherein, sensor nodes close to the data sink deplete their batteries faster than other nodes. Thus, the overall network lifetime is reduced (since these nodes spend a significant fraction of their energy in communicating data to the sink). Since the sensor nodes can reliably communicate within a region surrounding their location (called *communication neighborhood* or *communication set* hereafter), the mobile robots can collect data by visiting the communication neighborhood of each sensor node. The total data collection energy cost is the sum of the transmission energy expended by the sensor nodes to transmit the data and the energy expended by the mobile node for visiting the sensor nodes. In this paper, we study the problem of *designing a route for a data collecting mobile robot* such that the sum of transmission energy and traveling energy is minimized.

Usually, in the literature, the sensor nodes are assumed to lie on a plane and the communication neighborhood of each node is assumed to be a disc of a given radius. However, in general, the sensor nodes can lie in three-dimensional ($3D$) space [16]. Moreover, the sensor communication set may not be a disc (or a sphere) and can possibly have directional communication properties [31] (e.g., in underwater acoustic communication). Anisotropic communication properties can arise (a) due to the presence of directional antennas and/or (b) due to environmental effects and hardware differences in case of radios that are theoretically identical [31]. To capture the directional communication property of the sensor nodes we generalize the disc model to be a convex set. Thus, we consider sensor networks where the sensor nodes are heterogeneous with respect to their communication capabilities.

The data transmission energy is usually a superlinear function of the distance [12], [13], [27], whereas the travel energy varies linearly with distance. If we neglect the transmission cost, the problem of designing a data collection tour for a single robot with least travel cost becomes a traveling salesman problem with neighborhoods (TSPN) [2]. Hence, the TSPN problem is a special case of our problem. Solving the TSPN problem is NP-hard and it is also NP-hard to approximate the TSPN problem within a factor of $(2 - \epsilon)$ [24]. Thus, the problem of finding the minimum energy data collection tour is also computationally intractable unless $P = NP$. We design an algorithm for computing minimum cost routes for mobile data collecting robots where the sensor nodes can be arranged in $3D$ space and the communication set of each node is assumed to be a convex set (which includes the disc model of communication as a special case).

We first present a necessary condition for the optimality of a tour of a mobile robot through the sensor neighborhoods. Based on our necessary condition, we design a three-step algorithm. First, we find a traveling salesman tour through the sensor nodes that gives us an order of visiting the sensor nodes. Second, given the order of visiting the nodes, we modify the tour such that the sum of the transmission cost and travel cost is minimized. The second step can be formulated as a convex optimization problem and can thus be solved optimally. The output of the second step is a sequence of points lying in the sensor neighborhoods. In the third step, we formulate and solve a TSP (where the metric used is a modification of the Euclidean distance) using the output of the second step. In this step, we reduce the movement energy, while keeping the

J. Goerner is with Disney Research, Pittsburgh. N. Chakraborty and K. Sycara are with the Robotics Institute, Carnegie Mellon University, Pittsburgh. Email: `jgg@andrew.cmu.edu`, {`nilanjan`, `katia`}`@cs.cmu.edu`

transmission energy constant. The solution that we obtain is a local minimum of the overall optimization problem since it satisfies the necessary conditions for optimality. We prove that the energy cost of our route is within a constant factor of the energy cost of the global optimal solution. We present simulation results for both planar and spatial sensor networks. The key contribution of this paper is that we give an algorithm with provable performance guarantees for data collection in both $2D$ and $3D$ sensor networks with sensor nodes having anisotropic communication sets.

II. RELATED WORK

There are different versions of the data collection problem that can be formulated depending on the objective to be optimized, the assumptions on the communication model, and the assumptions on the motion model of the mobile robots. Two natural objectives are to minimize the time required for data collection [5] or the energy consumed in data collection [8], [30], [4]. Two commonly used mobility models for mobile data collecting agents are random mobility and controlled mobility. Irrespective of the optimization objective and mobility assumptions, the sensor networks are assumed to be planar. In this paper, we are concerned with energy optimization in $2D$ or $3D$ sensor networks, with controlled mobility of the robots. Therefore, we will focus our review on the literature concerned with energy optimization and controlled mobile robots (for a more comprehensive and general review, see [17], [27]).

The total energy expended in data collection is the sum of the data transmission energy expended by the static sensor nodes and the energy expended in the movement of the mobile data collecting node. In [8], the authors consider the problem of designing a route for a mobile data collector such that the overall data transmission energy is minimized with a constraint on the distance traveled by the robot. They assume that the amount of data to be transmitted by each sensor node is different. On the other hand, Yuan et al. consider the problem of designing a minimum cost route for robots transferring data between nodes in large sensor networks [30]. Their problem is a TSPN with non-overlapping neighborhoods. Bari et al. consider the problem of minimizing the travel length of a mobile data collector in a hierarchical sensor network [4]. In [4], the mobile data collector visits static relay nodes (which collect data from the sensor nodes) instead of visiting the sensor nodes individually. Their problem is also a TSPN but with overlapping circular neighborhoods. In this paper, we consider the problem of simultaneously minimizing the data transmission energy and the movement energy of the mobile robot. Although [27] considers the energy of the moving nodes as well as transmission energy, they assume the mobile robots to be moving along fixed paths and regulate the velocity of the robot for improving transmission quality. All of the above works consider planar sensor networks with isotropic communication models (i.e., the communication set is a disc). We consider both planar and spatial sensor networks and also anisotropic communication sets (modeled as convex sets).

As stated above, some of the literature [30], [4] formulate the data collecting tour as a Euclidean TSPN problem defined as follows: *Given a collection of n sets in $\mathbb{R}^D$, $\mathcal{R}_i, i = 1, \ldots, n$, a tour through the sets is a sequence of n points, $\mathbf{q}_i \in \mathcal{R}_i$, such that each set is visited at least once. The traveling salesman problem with neighborhood is to find a tour of minimum Euclidean length through the sets.* Note that if each set $\mathcal{R}_i$ is a single point, we have the classical Traveling Salesman Problem (TSP). Thus, the Euclidean TSPN is a generalization of the Euclidean TSP, which is NP-complete [21], and is therefore hard to solve optimally. The TSP and different versions of its generalizations that are known as vehicle routing problems (VRP) have been studied in the operations research [29] as well as robotics literature (see [7], [22] and references therein). For the purposes of this paper, the TSPN is the most relevant and so we will limit our discussion to TSPN. In the computational geometry literature, a number of algorithms with approximation guarantees that depend on assumptions about the shape of the neighborhoods of the points and the overlap between the neighborhoods have been proposed [2], [9], [10], [14], [11], [19]. Note that the TSPN problem is a special case of our problem and we can use our algorithm presented to solve the TSPN problem.

III. PROBLEM FORMULATION

In this Section, we present our problem formulation for minimizing the total data collection energy and indicate the assumptions that we make. We assume that there is a convex set $\mathcal{R}_i$ around each sensor node within which it can communicate reliably, i.e., the packet loss is low enough such that error correcting codes can faithfully reconstruct the data. The energy consumed for data transfer from a sensor node at $\mathbf{p}_i$ to a point, $\mathbf{p}_j$, is a superlinear function of the distance d_{ij}.

$$E_i = \gamma_i + \beta_i (d_{ij})^{\alpha_i} \tag{1}$$

where β_i and γ_i are constant parameters that depend on the sensor and the amount of data to be transmitted, and α_i is the path loss exponent (usually, $2 \leq \alpha_i \leq 6$ [26]). The travel energy expended by the mobile robot is proportional to the total distance traveled. Let $\mathbf{p}_j \in \mathbb{R}^D, D = 2, 3, j = 1, \ldots, n$, be the positions of the sensor nodes and $\mathbf{q}_j \in \mathbb{R}^D, D = 2, 3, j = 1, \ldots, n$, be the positions inside the communication set $\mathcal{R}_j$ of node j, where the robot visits to collect the data. Let $\mathcal{Q}_n = \{\mathbf{q}_i, i = 1, \ldots, n\}$ be the set of points the robot visits. Let $\mathcal{P}_n$ be the set of all permutations of the n sensor nodes, i.e., the set of all possible sequences in which all the regions $\mathcal{R}_j$ can be visited. The data transmission energy for each node, j, is then proportional to $\|\mathbf{p}_j - \mathbf{q}_j\|^{\alpha_j}$ and for a given sequence $\Pi = \{i_1, i_2, \ldots, i_n\} \in \mathcal{P}_n$ of visiting the nodes the total energy expended can be written as

$$E = \sum_{j=1}^{n} (w_{0j} + w_{1j} \|\mathbf{p}_j - \mathbf{q}_j\|^{\alpha_j}) +$$
$$w_2 (\sum_{k=1}^{n-1} \|\mathbf{q}_{i_k} - \mathbf{q}_{i_{k+1}}\| + \|\mathbf{q}_{i_n} - \mathbf{q}_{i_1}\|) \tag{2}$$

where w_{0j}, w_{1j} are constants dependent on the sensor parameters and w_2 is a constant dependent on the mobile robot. We can also incorporate a weight within these constant factors that can take in the preference of the network designer on the importance of two components of the energy loss in the particular application. For example, if the energy loss for the data collector is not of concern, we can assign $w_2 = 0$. The first term in Equation (2) is the total transmission energy (E_{trans}) and the second term is the total motion energy (E_{move}).

The general combinatorial optimization problem that we are studying can be stated as follows:
We have a set of n points $\mathbf{P} = \{\mathbf{p}_i\} \subset \mathbb{R}^D, D = 2, 3, i = 1, \ldots, n\}$. With each point $\mathbf{p}_i$ we associate a convex set $\mathcal{R}_i$, such that $\mathbf{p}_i \in \mathcal{R}_i$. Find a tour of a mobile robot visiting each region $\mathcal{R}_i$ such that the total transmission energy and travel energy given by Equation (2) *is minimized*, i.e.,

$$\min \ E(\Pi, \mathcal{Q}_n) \ \text{s.t.} \ \mathbf{q}_i \in \mathcal{R}_i, i = 1, \ldots, n \qquad (3)$$

In Equation (3), we optimize over the points, $\mathbf{q}_i \in \mathcal{R}_i$, to be visited as well as the sequence, Π, in which the regions, $\mathcal{R}_i$, should be visited. Each region $\mathcal{R}_i$ is a convex set specified by a set of convex inequalities and let m be the total number of constraints specifying all the communication regions.

The problem in (3) is a generalization of the TSPN problem. If w_{0j} and w_{1j} are both zero, then the problem is a TSPN. Thus, the algorithm proposed in Section IV can also be used to find the solution of TSP with convex neighborhoods.

Remark 3.1: Equation 1 is a popular model for energy consumption in sensor networks. However, the discussion in this paper holds for more general energy consumption functions as long as the function is a convex function of the two points between whom data is transferred.

Remark 3.2: In our problem formulation, there are some standard system-level assumptions that we state here explicitly. We consider only the energy expended in transmitting the data by the sensor nodes and the travel energy expended by the robot. Consequently, we assume an ideal Medium Access Control (MAC) layer, such that negligible energy is wasted in collisions. Moreover, the cost of storing the data and other operations that a sensor node may carry out is also neglected. The sensor nodes and mobile robot are localized and a mobile data collector must stop in order to collect data without errors. The mobile robot has a radio with enough power so that it can establish a link with the sensor nodes (for data transmission to start) once it reaches near a sensor node. The sensor data is not considered to be time-sensitive.

IV. MINIMUM ENERGY ROUTING ALGORITHM

In this section, we present our algorithm for computing the minimum energy route (MER) for a data collecting mobile robot in a sensor network with heterogeneous communication capabilities. We first characterize the necessary conditions for the MER path to be an optimal solution. Based on this characterization, we design a three-step algorithm for computing the MER. The three steps are as follows.

1) Given the set of sensor node locations $\mathbf{P}$, find a TSP tour through these points. Let $\bar{\Pi} \in \mathcal{P}_n$, the sequence of visiting the sensor nodes, be the output of this step.
2) For the sequence $\bar{\Pi}$ solve the optimization problem in Equation 3. The output of this step is a set of points $\mathbf{Q} = \mathbf{q}_i, i = 1, \ldots, n$, all of which may not be distinct.
3) On this set of points form a metric between any two points $\mathbf{q}_i$ and $\mathbf{q}_j$ using the Euclidean distance between the two points and the transmission energy from the sensor nodes $\mathbf{p}_i$ to $\mathbf{q}_i$ and $\mathbf{p}_j$ to $\mathbf{q}_j$. Solve the TSP on the set $\mathbf{q}$ with the modified metric and output the solution. This sequence of the set $\mathbf{Q}$ is the final sequence in which the robot travels.

We will now describe the necessary condition and then elaborate on the steps of the algorithm. Without loss of generality, let the optimal sequence of visiting the points be $\Pi^* = \{1, \ldots, n\}$ and $q_i^*, i = 1, \ldots, n$, be the locations where the robots should visit for the minimum energy route. The transmission cost for sensor node i in the optimal solution is $E_i^* = w_{0i} + w_{1i} \|\mathbf{q}_i^* - \mathbf{p}_i^*\|^{\alpha_i}$. Let r_{ij} be the distance between $\mathbf{q}_i^*$ and $\mathbf{q}_j^*$ for any two sensor nodes i and j and $\bar{r}_{ij} = r_{ij} + \frac{1}{2}(E_i^* + E_j^*)$.

Lemma 4.1: The necessary condition for $\mathbf{q}_i^*, i = 1, \ldots, n$, to be an optimal solution for the optimization problem in Equation (3) is that the sequence $\Pi^* = \{1, \ldots, n\}$ is the minimum cost TSP tour on the set of points $\mathbf{q}_i^*, i = 1, \ldots, n$, with the cost of the edge between any two points $\mathbf{q}_i^*$ and $\mathbf{q}_j^*$, $i \neq j, i = 1, \ldots, n, j = 1, \ldots, n$, being $\bar{r}_{ij}$.

Proof: For the sequence Π^*, the total data collection cost is $E^* = E_{\text{trans}}^* + E_{\text{move}}^*$. Let us consider the complete weighted graph $G = (V, E, w)$, where V is a set of nodes corresponding to the points $\mathbf{q}_i^*$ and $w = \bar{r}_{ij}$. Now E^* is the cost of a tour on G that follows the sequence Π^*. Let $\Pi = \{i_1, \ldots, i_n\}$ be some other tour on G whose cost is $E < E^*$. Now,

$$\begin{aligned}
E &= \bar{r}_{i_1 i_2} + \bar{r}_{i_2 i_3} + \cdots + \bar{r}_{i_n i_1} \\
&= \frac{1}{2}(\|\mathbf{p}_{i_1} - \mathbf{q}_{i_1}^*\|^{\alpha_{i_1}} + \|\mathbf{p}_{i_2} - \mathbf{q}_{i_2}^*\|^{\alpha_{i_2}}) \\
&\quad + \frac{1}{2}(\|\mathbf{p}_{i_2} - \mathbf{q}_{i_2}^*\|^{\alpha_{i_2}} + \|\mathbf{p}_{i_3} - \mathbf{q}_{i_3}^*\|^{\alpha_{i_3}}) \\
&\quad + \cdots + \frac{1}{2}(\|\mathbf{p}_{i_n} - \mathbf{q}_{i_n}^*\|^{\alpha_{i_n}} + \|\mathbf{p}_{i_1} - \mathbf{q}_{i_1}^*\|^{\alpha_{i_1}}) \\
&\quad + \|\mathbf{q}_{i_1}^* - \mathbf{q}_{i_2}^*\| + \|\mathbf{q}_{i_2}^* - \mathbf{q}_{i_3}^*\| \\
&\quad + \cdots + \|\mathbf{q}_{i_n}^* - \mathbf{q}_{i_1}^*\| \\
&= E_{\text{trans}}^* + E_{\text{move}}
\end{aligned} \qquad (4)$$

where $E_{\text{move}} = \|\mathbf{q}_{i_1}^* - \mathbf{q}_{i_2}^*\| + \cdots + \|\mathbf{q}_{i_n}^* - \mathbf{q}_{i_1}^*\|$. The first term of Equation (4) is identical to E_{trans}^*. Thus, for E to be less than E^*, the second term of Equation (4) should be less than E_{move}^*. This implies that there is a solution to the MER problem with cost less than E^*. By assumption, E^* is the optimal tour and thus we arrive at a contradiction. Therefore, the sequence Π^* has to be the optimal TSP tour on G. ∎

Note that the condition stated in Lemma 4.1 is not a sufficient condition. A simple example of a sequence of points that satisfy the condition is a TSP tour through the sensor

nodes. Clearly, this solution is not the optimal MER.

In the first step of our algorithm to find the MER, we find a sequence in which the data collector should visit the sensor nodes and in the second step, we compute a minimum energy route for the sequence obtained in the first step. For finding the sequence of visiting the sensor nodes, we solve a Euclidean traveling salesman problem on the sensor node positions. The TSP solution gives us a minimum length route for the mobile robot where the transmission energy is as small as possible. In other words it gives us the optimal solution to Equation (3) if we assign $\mathbf{q}_j = \mathbf{p}_j$, whence the first term in Equation (2) corresponding to the transmission energy becomes constant. Although the Euclidean TSP problem is NP-complete [21], there are theoretical algorithms that can solve Euclidean TSP within an arbitrary small factor [3], [20] and also good implementations of heuristic solutions that give very good performance in practice [1].

After we obtain a sequence for visiting the sensor nodes by solving the TSP, we can solve the problem of finding the points to be visited within the communication set by solving a convex optimization problem. Without loss of generality let the sequence obtained for visiting the sensor nodes from the TSP in the first step be $\{1, \ldots, n\}$. The optimization problem in Equation (3) can then be written as

$$
\begin{aligned}
\min \quad & \sum_{j=1}^{n} (w_{0j} + w_{1j} \|\mathbf{p}_j - \mathbf{q}_j\|^{\alpha}) + \\
& w_2 (\sum_{k=1}^{n-1} \|\mathbf{q}_i - \mathbf{q}_{i+1}\| + \|\mathbf{q}_n - \mathbf{q}_1\|) \\
\text{s.t.} \quad & \mathbf{q}_i \in \mathcal{R}_i, i = 1, \ldots, n
\end{aligned}
\tag{5}
$$

Equation (5) is a convex optimization problem because each term in the objective function is a convex function in $\mathbf{q}_i$, and $\mathbf{q}_i$ is constrained to lie within a convex set since we have assumed $\mathcal{R}_i$ to be convex sets. Therefore, we can solve the problem optimally in polynomial time using interior point methods that are efficient in practice (i.e., in $O(m+n)^{3.5}$ time theoretically but much faster practically [6]). Note that the solution obtained in this step is optimal for a given sequence of visiting the nodes. However, there is no guarantee of optimality for the overall problem, since the sequence that we have obtained may not be the optimal sequence.

In order to ensure that the necessary condition is met, once we have created a path that minimizes the total energy cost, we solve the TSP again over the set of points in the path. In this TSP, the distance between two points $\mathbf{q}_i$ and $\mathbf{q}_j$ is given by $\bar{r}_{ij}$. Since the points do not move, the transmission energy cannot change. Then, there are two scenarios: in the first case, the TSP solution does not change the order that our path already had. In that case, the path already meets the necessary condition and is at least a local optimum to the optimization problem give by Equation (3). In the second case, the TSP solution does change the order, decreasing the energy cost due to movement and ensuring that the new path meets the necessary condition for an optimal solution given in Lemma 4.1.

Computing the TSPN route: The first two steps of the algorithm described above can also be used for solving the traveling salesman problem with neighborhoods. If we set $w_{0j} = w_{1j} = 0$, the objective function in Equation 5 corresponds to a scalar multiple of the total distance traveled and is still convex. Thus step 2 can be solved by solving a convex optimization problem. There are algorithms in the mobile data collection literature [30], [4], as well as in computational geometry [10] that have used such a two-step approach for solving the TSPN problem. Since we have posed the problem of finding the minimum length path as a convex optimization problem (once the sequence is fixed), we can solve the problem optimally. Hence, our approach will give the best solution to the TSPN that is achievable by the two-step approach of first finding a TSP tour through the points and then modifying the tour (while keeping the order fixed) to reduce the total distance traveled.

The TSPN solution can also be used as a data collection route for the mobile robot. The TSPN tour gives us a data collection tour where the travel energy alone is minimized. However, in this case, the transmission energy to the points should be added to obtain the total energy cost.

V. APPROXIMATION GUARANTEES OF MER ALGORITHM

In this section, we give bounds on the performance of the MER algorithm. We show that the energy cost of a TSP tour through the sensor nodes is within a constant factor of the optimal solution. Since our algorithm is guaranteed to give a solution with energy cost that is upper bounded by the cost of TSP tour through the sensor nodes, it follows that our solution is within a constant factor of the optimal solution. We will first present the results for the case where the communication sets are non-overlapping balls of different radii in $\mathbb{R}^d, d = 2, 3$. We will then extend the results to the case of overlapping ball-shaped communication sets and finally extend the results to general convex sets.

Let $r_i, i = 1, \ldots, n$, be the radii of the communication set of the ith sensor node and $r_{\max}, r_{\min}$ be the maximum and minimum communication set radii respectively. Let L_c be the tour length through the sensor nodes used in our MER algorithm, and L_c^* be the optimal TSP tour length through the sensor nodes. Let $\alpha = \frac{L_c}{L_c^*}$ be the approximation ratio of the TSP tour used in the MER algorithm and $\beta = \frac{r_{\max}}{r_{\min}}$. Let L^* be the tour length for the minimum energy tour. Let E be the energy cost along the MER and E^* be the energy cost for the optimal solution. Then we have the following:

Theorem 5.1: Let the communication sets of the sensor nodes be disjoint balls. The energy cost of the route given by the MER algorithm is within a constant factor of the optimal energy cost, i.e.,

$$
E \leq (1 + \frac{8\beta^2}{\pi})\alpha E^* + C_1 \quad \text{for} \quad d = 2, \tag{6}
$$

$$
E \leq (1 + 6\beta^3)\alpha E^* + C_2 \quad \text{for} \quad d = 3, \tag{7}
$$

where $C_1 = 8\alpha\beta^2 r_{\max}$ and $C_2 = 16\alpha\beta^3 r_{\max}$ are constants.

Proof: The tour of length L^* can be converted to a center tour by a detour of at most $2r_i$ for each node i. Since the optimal TSP tour through the center nodes will have length lesser than this modified tour, we have

$$L_c^* \leq L^* + 2\sum_{i=1}^n r_i \leq L^* + 2nr_{\max}. \tag{8}$$

Now, we have to find an upper bound for n in terms of the tour length. Let us consider a ball of radius $2r_{\max}$ that moves with its center on the optimal route. For $d = 2$, the area swept by this ball is $4r_{\max}L^* + 4\pi r_{\max}^2$. For $d = 3$, the volume swept by the ball is $4\pi r_{\max}^2 L^* + \frac{32}{3}\pi r_{\max}^3$ This area (volume) is greater than the area (volume) of the communication sets of all the sensor nodes, i.e., for $d = 2$

$$\pi \sum_{i=1}^n r_i^2 \leq 4r_{\max}L^* + 4\pi r_{\max}^2$$
$$\Rightarrow \ \pi \sum_{i=1}^n r_{\min}^2 \leq 4r_{\max}L^* + 4\pi r_{\max}^2 \tag{9}$$
$$\Rightarrow \ n \leq \frac{4\beta^2}{\pi r_{\max}}L^* + 4\beta^2$$

For $d = 3$, we have

$$\frac{4}{3}\pi \sum_{i=1}^n r_i^3 \leq 4\pi r_{\max}^2 L^* + \frac{32}{3}\pi r_{\max}^3$$
$$\Rightarrow \ n \leq \frac{3\beta^3}{r_{\max}}L^* + 8\beta^3 \tag{10}$$

Now, substituting for n in Equation (8), and using $L_c \leq \alpha L_c^*$ we have

$$L_c \leq (1 + \frac{8\beta^2}{\pi})\alpha L^* + 8\alpha\beta^2 r_{\max} \quad \text{for} \quad d = 2,$$
$$L_c \leq (1 + 6\beta^3)\alpha L^* + 16\alpha\beta^3 r_{\max} \quad \text{for} \quad d = 3, \tag{11}$$

The movement cost is proportional to the tour lengths, and for motion through the sensor nodes, the transmission cost is zero. Moreover, the coefficient of L^* in Equation (11) is greater than 1. Therefore, for $d = 2$, $E_c \leq C_0(L^* + T^*) + C_1$, where $C_0 = (1 + \frac{8\beta^2}{\pi})\alpha$, E_c is the energy cost of motion through the sensor nodes, and T^* is the optimal transmission energy. Now, by construction, the energy cost of MER tour is less than the energy cost of center tour. Therefore, $E_c \leq C_0 E^* + C_1$. An analogous argument can be made for $d = 3$, where, only the value of the constant coefficients are different. ∎

Theorem 5.2: Let the sensor communication sets be balls that can overlap with other sets. The energy cost of the MER is within a constant factor of the optimal route, i.e.,

$$E \leq (1 + 8\beta^2(1 + \frac{1}{\pi}))\alpha E^* + C_1 \quad \text{for} \quad d = 2, \tag{12}$$
$$E \leq (1 + 6\beta^3(1 + \pi))\alpha E^* + C_2 \quad \text{for} \quad d = 3, \tag{13}$$

where C_1 and C_2 are as defined above.

Proof: As before, let L^* be the length of the optimal tour. The optimal tour consists of straight line segments, where the turning points are the points where the mobile robot can stop to collect data (note that if the tour is a straight line through some nodes, the turning angle is π radians). The tour divides the sensor nodes into (say γ) clusters, where all the data in a cluster can be downloaded from one point on the tour. Note that if all the sensor nodes have disjoint communication sets, then $\gamma = n$. The center nodes of all the nodes in the clusters are within a ball of radius $r_{\max}$. Therefore, in a cluster, we can make a detour of at most $2r_{\max} + 2\pi r_{\max}$ to obtain a tour through the sensor nodes. Therefore, we have

$$L_c^* \leq L^* + 2(\pi + 1)\gamma r_{\max} \tag{14}$$

Now, we have to find an upper bound for γ. Using the same arguments as for finding the bound of n in the proof of Theorem 5.1 (only change being that the sums in Equations (9) and (10) are over γ), we obtain identical expressions for γ. Substituting them in Equation (14), we obtain

$$L_c \leq (1 + 8\beta^2(1 + \frac{1}{\pi}))\alpha L^* + 8\alpha\beta^2 r_{\max} \quad \text{for} \quad d = 2,$$
$$L_c \leq (1 + 6\beta^3(1 + \pi))\alpha L^* + 16\alpha\beta^3 r_{\max} \quad \text{for} \quad d = 3, \tag{15}$$

Now, using the same arguments as in the proof of Theorem 5.1, we obtain Equations (12) and (13).

For generalizing the results to any convex set, one should note that the radius of a convex set is the radius of the smallest ball circumscribing the convex set. In the proofs above, the shape of the sets were used to derive a bound on γ. Now, the Equations (9) and (10) hold for general convex sets, with r_i being their radii. Therefore, the Theorems 5.1 and 5.2 are valid for general convex sets.

VI. SIMULATION RESULTS

We now present simulation results of our algorithm. We generated random sensor networks with node density 1 and 2 both in $\mathbb{R}^2$ and $\mathbb{R}^3$. For the results presented here, we assumed the communication sets to be either discs (spheres) or cones. The communication set of each sensor node is a cone or a disc with probability 0.5. The radius for circles or spheres ranges between 0.8 and 1.2, and the radius for cones ranges between 1.3 and 1.7. The differences in radius are to ensure that the communication sets take up approximately the same area. The cone angles range between $\pi/8$ and $3\pi/8$. We set $w_{0j} = 0, w_{1j} = 1, w_2 = 1$, and $\alpha_j = 3, \forall j$ in our simulations.

We present the cost of the TSP tour through the sensor nodes and the the TSPN tour through the communication neighborhoods of the sensor nodes as baseline results. Note that the TSP tour gives the minimum movement cost tour when the transmission cost is 0 (in practice it will be a constant $\sum_j w_{0j}$). The TSPN tour gives the tour with minimum movement cost. To compute the total cost of this tour, we add the energy cost for transmission from the sensor node to the point where the mobile node visits the communication neighborhood. The TSPN route is computed using the method in Section IV. The MER route is computed using the three-step approach presented in Section IV. For computing the TSP tour we have used the Lin-Kerninghan heuristic (as implemented in the Concorde TSP solver).

We ran simulation for $2D$ and $3D$ networks of sizes ranging from 20 to 800. Tables I and II are for randomly generated planar sensor networks with node density 1 and 2 respectively. For $3D$ sensor networks the results are given in Tables III and IV. In each table, the second column gives the motion cost of TSP tour. Since in this case the transmission cost is 0, this is also the total data collection cost for the TSP tour. The third, fourth, and fifth columns gives the motion cost, transmission cost, and total cost of the TSPN tour. The sixth, seventh, and eighth column gives the motion cost, transmission cost, and total cost of the minimum energy route. The last column gives the time required for computing the MER in seconds (not including the time needed to get the ordering of nodes from the TSP solver). For a given number of nodes, each data point is an average over 20 different randomly generated sensor networks. All runs were on a computer with 12 GB RAM, quad core 2.6GHz processor, and an 8 MB cache.

From all the tables, we can see that the cost of the MER route is always less than the TSPN cost and the TSP cost. Although the MER motion cost is more than the TSPN route, the reduction of transmission cost makes the total cost less for MER. The reduction of the MER cost is quite consistent across different node densities, network sizes, and the dimensionality of space. For example the ratio of the MER cost to TSPN cost is consistently between 0.7 and 0.8 across all the runs. When comparing the TSPN cost to the TSP cost, we observe an interesting phenomenon. As Table I shows, the TSPN tour has a higher cost for planar sensor networks with node density 1, irrespective of the number of nodes. When node density is increased in a planar network, all energy costs (total, motion, and transmission for all paths) drop, and the TSPN path becomes more efficient than the TSP path (see Table II).

VII. SUMMARY AND FUTURE WORK

In this paper, we presented an algorithm to compute a route for a mobile data collecting robot in planar or spatial sensor networks with heterogeneous communication capabilities. The communication set of each sensor is modeled as a convex set, which is a generalization of the disc communication models usually used in the literature. We derived a necessary condition for optimality of a route and presented a three-step algorithm for computing the MER route which gives a local minimum for our problem. We first compute a sequence of visiting the sensor nodes by finding a traveling salesman tour through them. Given the order of visiting the nodes, we formulate the problem of minimizing data collection cost as a convex optimization problem and solve it optimally. The output of the second step is a sequence of points lying in the sensor neighborhoods. In the third step, we formulate and solve a TSP (where the metric used is a modification of the Euclidean distance) on the output of the second step, thereby reducing the movement energy, while keeping the transmission energy constant. We proved that our algorithm is within a constant factor of the optimal solution.

Future Work: Although we provide approximation guarantees for our algorithm, the approximation factors are quite weak. One avenue of future research would be to improve the analysis of the algorithm in this paper or find approximation algorithms with tighter guarantees. Extending the algorithm for data collection with multiple robots is also of future interest.

ACKNOWLEDGMENTS

REFERENCES

[1] D. L. Applegate, R. E. Bixby, V. Chvatal, and W. J. Cook. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, 2006.

[2] E. Arkin and R. Hassin. Approximation algorithms for the geometric covering salesman problem. *Discrete Applied Math*, 55(3):197–218, 1994.

[3] S. Arora. Nearly linear time approximation schemes for Euclidean TSP and other geometric problems. *Journal of ACM*, 45(5):1–30, 1998.

[4] A. Bari, Y. Chen, D. Roy, A. Jaekel, and S. Bandopadhyay. Energy aware trajectory computation of mobile data collectors in hierarchical sensor networks. In *Proceedings of IEEE ICC 2010*, May 2010.

[5] D. Bhadauria and V. Isler. Data gathering tours for mobile robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009.*, pages 3868 –3873, October 2009.

[6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, U.K., 2004.

[7] F. Bullo, E. Frazzoli, M. Pavone, K. Savla, and S. Smith. Dynamic vehicle routing for robotic systems. *Proceedings of the IEEE*, 99(9):1482 –1504, sept. 2011.

[8] D. Ciullo, G. D. Celik, and E. Modiano. Minimizing transmission energy in sensor networks via trajectory control. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*, pages 132 –141, June 2010.

[9] M. de Berg, J. Gudmundsson, M. Katz, C. Levcopoulos, M. Overmars, and A. van der Stappen. TSP with neighborhoods of varying size. *Journal of Algorithms*, 57(1):22–36, 2005.

[10] A. Dumitrescu and J. Mitchell. Approximation algorithms for TSP with neighborhoods in the plane. *Journal of Algorithms*, 48(1):135–159, 2003.

[11] K. Elbassioni, A. Fishkin, N. Mustafa, and R. Sitters. Approximation algorithms for euclidean group TSP. In *Proceedings of 32nd International Colloquium Automata, Languages, and Programming*, pages 1115–1126, Lisboa, Portugal, July 2005.

[12] A. Ephremides. Energy concerns in wireless networks. *IEEE Wireless Communications*, 9:48–59, August 2002.

[13] D. K. Goldenberg, J. Lin, A. S. Morse, B. E. Rosen, and Y. R. Yang. Towards mobility as a network control primitive. In *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '04, pages 163–174, 2004.

[14] J. Gudmundsson and C. Levcopoulos. Fast approximation algorithm for TSP with neighborhoods. *Nordic Jnl. of Computing*, 6(4):469–488, 1999.

[15] T. He, S. Krishnamurthy, J. A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, L. Gu, J. Hui, and B. Krogh. Energy-efficient surveillance system using wireless sensor networks. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, MobiSys '04, pages 270–283, 2004.

[16] W. J. Kaiser, G. J. Pottie, M. Srivastava, G. S. Sukhatme, J. Villasenor, and D. Estrin. Networked infomechanical systems (NIMS) for ambient intelligence. Technical Report 31, Center For Embedded Networked Sensing, UCLA, 2003.

[17] J. Ma, C. Chen, and J. P. Salomaa. mWSN for large scale mobile sensing. *J. Signal Process. Syst.*, 51:195–206, May 2008.

[18] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson. Wireless sensor networks for habitat monitoring. In *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, WSNA '02, pages 88–97, 2002.

TABLE I
ENERGY COST FOR THE TSP ROUTE, TSPN ROUTE AND MINIMUM ENERGY ROUTE (MER) FOR PLANAR SENSOR NETWORKS (NODE DENSITY = 1).

Number of nodes	TSP Cost Motion	TSPN Cost			MER Cost			MER Computation time
		Motion	Transmission	Total	Motion	Transmission	Total	
20	17.6	11.7	9.1	20.8	12.9	1.8	14.7	2.8
160	130.4	86.1	54.5	140.6	93.7	8.6	102.3	22.4
320	264.1	169.1	108.5	277.6	185.1	16.5	201.6	47.2
480	385.4	248.6	159.4	408.0	272.5	23.4	295.9	74.7
640	515.4	328.6	218.3	546.9	360.8	32.2	393.0	105.3
800	647.6	412.7	267.5	680.2	452.0	39.2	491.2	138.7

TABLE II
ENERGY COST FOR THE TSP ROUTE, TSPN ROUTE AND MINIMUM ENERGY ROUTE (MER) FOR PLANAR SENSOR NETWORKS (NODE DENSITY = 2).

Number of nodes	TSP Cost Motion	TSPN Cost			MER Cost			MER Computation time
		Motion	Transmission	Total	Motion	Transmission	Total	
20	15.3	7.9	5.2	13.1	8.9	0.7	9.6	2.8
160	114.7	59.6	30.4	90.0	65.5	4.2	69.7	22.5
320	225.5	115.6	63.3	178.9	127.4	7.7	135.2	47.4
480	335.9	172.7	93.2	265.9	190.1	11.5	201.6	74.6
640	450.9	229.9	113.3	343.2	251.4	14.4	265.9	104.6
800	565.8	286.6	140.3	426.9	313.5	18.3	331.8	137.5

TABLE III
ENERGY COST FOR THE TSP ROUTE, TSPN ROUTE AND MINIMUM ENERGY ROUTE (MER) FOR SPATIAL SENSOR NETWORKS (NODE DENSITY = 1).

Number of nodes	TSP Cost Motion	TSPN Cost			MER Cost			MER Computation time
		Motion	Transmission	Total	Motion	Transmission	Total	
20	17.4	7.8	6.1	13.9	9.3	1.7	11.0	2.1
160	122.9	58.4	40.3	98.7	67.8	10.8	78.7	16.3
320	242.0	114.7	79.8	194.5	133.1	21.3	154.3	34.3
480	358.2	168.7	116.4	285.1	195.3	31.1	226.5	54.1
640	474.1	225.2	151.2	376.4	259.8	41.3	301.1	75.4
800	589.0	278.8	193.3	472.0	322.7	50.8	373.5	98.4

TABLE IV
ENERGY COST FOR THE TSP ROUTE, TSPN ROUTE AND MINIMUM ENERGY ROUTE (MER) FOR SPATIAL SENSOR NETWORKS (NODE DENSITY = 2).

Number of nodes	TSP Cost Motion	TSPN Cost			MER Cost			MER Computation time
		Motion	Transmission	Total	Motion	Transmission	Total	
20	13.7	5.0	4.7	9.6	6.1	1.5	7.6	2.1
160	98.7	41.0	32.8	73.7	48.4	9.3	57.7	16.9
320	192.3	82.3	63.5	145.9	96.5	17.3	113.9	35.4
480	285.5	122.8	92.1	214.9	143.3	25.5	168.8	55.5
640	376.4	158.5	125.8	284.2	186.4	34.2	220.5	78.0
800	468.0	198.3	152.7	351.1	232.1	42.2	274.3	101.7

[19] C. Mata and J. Mitchell. Approximation algorithms for geometric tour and network design problems. In *Proceedings of 11th ACM Annual Symposium of Computational Geometry (SoCG '95)*, pages 360–369, Vancouver, Canada, June 1995.

[20] J. S. B. Mitchell. Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric tsp, k-mst, and related problems. *SIAM Journal of Computing*, 28:1298–1309, March 1999.

[21] C. H. Papadimitriou. The Euclidean Traveling Salesman Problem is NP-complete. *Theoretical Computer Science*, 4(3):237–244, 1977.

[22] M. Pavone, E. Frazzoli, and F. Bullo. Adaptive and distributed algorithms for vehicle routing in a stochastic and dynamic environment. *Automatic Control, IEEE Transactions on*, 56(6):1259 –1274, june 2011.

[23] M. Perillo, Z. Cheng, and W. Heinzelman. An analysis of strategies for mitigating the sensor network hot spot problem. In *Proceedings of the The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 474–478, 2005.

[24] S. Safra and O. Schwartz. On the complexity of approximating TSP with neighborhoods and related problems. *Computational Complexity*, 14(4):281–307, 2006.

[25] R. C. Shah, S. Roy, S. Jain, and W. Brunette. Data mules: Modeling a three-tier architecture for sparse sensor networks. In *IEEE SNPA Workshop*, pages 30–41, 2003.

[26] K. Sohrabi, B. Manriquez, and G. J. Pottie. Near ground wideband channel measurement in 800 − 1000 MHz. In *IEEE 49th Vehicular Technology Conference*, pages 571–574, May 1999.

[27] A. A. Somasundara, A. Kansal, D. D. Jea., D. D. Estrin, and M. B. Srivastava. Controllably mobile infrastructure for low energy embedded networks. *IEEE Trans. on Mobile Computing*, 5:958–973, August 2006.

[28] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin. Habitat monitoring with sensor networks. *Commun. ACM*, 47:34–40, June 2004.

[29] P. Toth and D. Vigo. *The vehicle routing problem*. SIAM monographs on discrete mathematics and applications. Society for Industrial and Applied Mathematics, 2002.

[30] B. Yuan, M. Orlowska, and S. Sadiq. On the optimal robot routing problem in wireless sensor networks. *IEEE Transactions on Knowledge and Data Engineering*, 19(9):1252–1261, 2007.

[31] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. Impact of radio irregularity on wireless sensor networks. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, MobiSys '04, pages 125–138, 2004.