

Aggregate Node Placements in Sensor Networks

Weifa Liang
DCS
ANU
Canberra, Australia

Guanjun Ma
DCS
USTC
Hefei, China

Yinlong Xu
DCS
USTC
Hefei, China

Jiugen Shi
School of Comp.
Hefei Polytech Uni.
Hefei, China

Abstract—In this paper we consider the design issue of sensor networks by placing a few powerful aggregate nodes into a dense sensor network such that the network lifetime is significantly prolonged when performing data gathering. Specifically, the problem is to place K aggregate nodes into a dense sensor network of n sensor nodes with $K \ll n$ such that the lifetime of the resulting network is maximized, subject to the constraints that both the maximum transmission range of an aggregate node and the maximum transmission delay between an aggregate node and a sensor node covered by the aggregate node are met. Clearly, this is a joint optimization problem of aggregate node placement and the communication structure, which is NP-hard. We approach the problem by devising a fast and scalable heuristic algorithm. We also conduct experiments by simulation to evaluate its performance, and the experimental results show that the proposed algorithm outperforms a commonly equal distance placement schema significantly.

I. INTRODUCTION

Wireless sensor networks have been receiving significant attention due to their potential applications in environmental monitoring, surveillance, military operations and other domains. While these applications revealed tremendous potential for capturing important environment phenomenon, they have also posed certain associated limitations. One of the major limitations is that sensor nodes are powered by energy-limited batteries, the sensor lifetime imposes a severe constraint on network performance, so energy conservation in such networks is of paramount importance. As data gathering is one of the fundamental and frequent operations in most sensor networks, the energy gain obtained through the optimization of data gathering operation can help prolong the lifetime of a sensor network substantially. Therefore, energy efficient data gathering is one major research focus. In this paper we consider a strategy of prolonging the network lifetime through the deployment of a few powerful aggregate nodes into an existing dense sensor network when performing data gathering.

A. Related works

Given a sensor network consisting of homogeneous sensor nodes, data gathering that aims to minimize the total energy consumption and/or the maximum energy consumption has been extensively studied in the literature [7], [8], [10], [5], [9]. Unlike these mentioned works that focus on the construction of routing trees, there are several studies on placing sensor nodes into a region of interest to form a sensor network and building a routing tree for data gathering at the same time [3], [1], [4].

For example, Dasgupta *et al.* [3] considered the sensor placement problem so as to maximize the network lifetime under the constraint that certain points of interest are within the coverage radius of at least one sensor. Ganesan *et al.* [4] studied the joint optimization of sensor placement and transmission structure so as to minimize the total power consumption by the network under the maximum and average distortion constraints, using different coding schemes and correlation models. Cheng *et al.* [1] considered the node placement problem with an objective either to maximize the network lifetime or minimize the total energy consumption by formulating the problem into a non-linear programming problem. In contrast with these node placement algorithms for homogeneous sensor networks, we consider a node placement problem with aiming to place a few powerful aggregate nodes into the network strategically such that the lifetime of the resulting network is maximized when performing data gathering, subject to the constraints that both the maximum transmission range of aggregate nodes and the maximum transmission delay between an aggregate node and a sensor node covered by the aggregate node are met.

B. Contributions

Our major contributions are as follows. We first introduce an aggregate node placement problem in heterogeneous sensor networks, by placing a few powerful aggregate nodes into a dense sensor network such that the lifetime of the resulting network is further maximized when performing data gathering. Due to the NP hardness of the problem, we then devise a fast, scalable heuristic algorithm. We finally conduct extensive experiments by simulation to evaluate the performance of the proposed algorithm. The experimental results show that the proposed heuristic significantly outperforms another commonly equal distance node placement schema.

II. PRELIMINARIES

A. System model

We consider a heterogeneous sensor network consisting of two types of wireless devices: lots of resource-constrained, cheap *sensor nodes* and a few resource-rich, expensive *aggregate nodes*. The cheap sensor node has limited battery power, a short fixed transmission range, a low data rate and a low duty cycle. The main tasks performed by a sensor node are sensing, data processing, and data transmission and relay. In comparison with the cheap sensor node, the expensive aggregate node has more power reserve, an adjustable transmission range, a

higher data transmission rate, and better data processing and storage capabilities. The main tasks of an aggregate node are to aggregate and/or process the sensed data from sensor nodes and to transmit the collected data to the other aggregate nodes. Since the cost associated with aggregate nodes is not cheap, the number of them in a sensor network is very limited, and they must be placed carefully in order to maximize the lifetime of the network when performing data gathering.

To perform data gathering in such a heterogeneous sensor network, in-network processing paradigm can be applied [11]. That is, a routing tree rooted at the base station and spanning aggregate nodes and sensor nodes is required. The tree actually is a 2-tier cluster routing tree, in which the aggregate nodes serve as the cluster heads and the sensor nodes serve as the members in clusters. Each sensor node has only a cluster head. A sensor node is a *member* of its cluster head, and the cluster head *covers* the sensor node. Within a cluster each member can forward its sensing data to the cluster head through multiple-hop member relay. Once the cluster head collects all the sensing data from its members and descendant aggregate nodes, it can process and transmit the collected data to its parent aggregate node. All data will be collected at the base station ultimately through multi-hop aggregate nodes relay. Clearly, each sensor node consumes the same amount of transmission energy by transmitting a unit-length message, since it has an identical transmission range. However, the transmission energy consumption of an aggregate node is various, depending on the distance between the node and its receiver. For two aggregate nodes u and v with distance $d_{u,v}$, the transmission energy at u is modeled to be proportional to $d_{u,v}^\kappa$ if a unit-length message is transmitted from u to v , where κ is a path-loss exponent parameter that typically takes on a value between 2 and 4, depending on the characteristics of the communication medium.

B. Aggregate node placement problem

Given a sensor network of n sensor nodes deployed randomly in a region of interest, a base station, and a few expensive aggregate nodes K ($K \ll n$), the *aggregate node placement problem* is to place the K aggregate nodes into the network such that the lifetime of the resulting heterogeneous sensor network is maximized when performing data gathering, subject to the constraints that both the maximum transmission range R_{max} of an aggregate node and the maximum transmission delay D_{max} between an aggregate node and a sensor node covered by the aggregate node are met.

C. Network lifetime

Unless otherwise specified, in this paper we assume that the lifetime of a heterogeneous sensor network is fully determined by the lifetime of the subnetwork induced by the aggregate nodes only. The rationale behind our assumption is that the cheap sensor nodes are densely (redundantly) deployed in the monitored region, they must be highly correlated with each other in terms of the sensed data. Thus, the set of cheap sensor nodes in area s_i can be partitioned into several disjoint subsets

such that the sensor nodes in each subset cover the area s_i and the communication subgraph induced by the sensor nodes is connected. A routing subtree rooted at v_i spanning the sensor nodes in the subset can then be obtained [6]. Although the lifetime of each subtree is limited, the sum of the lifetimes of all the subtrees rooted at v_i are comparable to the lifetime of aggregate node v_i . Consequently, the lifetime of the routing tree consisting of aggregate nodes will determine the lifetime of the heterogeneous sensor network.

D. Data volume vs the number of sensor nodes in an area

With the assumption that sensor nodes in the sensor network are densely deployed, we assume that the total volume of sensed data generated by the sensor nodes in a subregion is proportional to the area of that subregion. We further assume that each sensor node generates the same volume of data at any given time interval. Therefore, in the rest of this paper we use the area of a subregion to represent the total volume of sensed data generated by the sensor nodes as well as the number of sensor nodes in the subregion.

III. AGGREGATE NODE PLACEMENT IN A FAN

We assume that the monitored region is a fan of angle θ with fan radius L , in which the sensor nodes are densely deployed, $0 < \theta < \pi$. We now assign k aggregate nodes to the fan to collect the sensed data generated by the sensor nodes in the fan area, with an objective to maximize the lifetime of the resulting network. Finding an optimal solution for this problem is very challenging, because it involves searching through the spaces of all possible configurations of aggregate node placement, and all possible routing trees rooted at the base station for each configuration. In fact, the problem is NP hard, we instead propose a heuristic by placing the k aggregate nodes in the middle ray line of the fan. The aggregate nodes are indexed from k to 1 in decreasing order, started from the side of the base station (see Fig. 1), where the base station is at the center of the fan. Let r_i be the transmission distance between aggregate node v_i and aggregate node v_{i+1} in this placement schema, we will determine the value of each r_i such that the lifetime of the resulting network is maximized, when performing data gathering, $1 \leq i \leq k$. Since an aggregate node near to the base station transmits more data than another aggregate node far away from it, to prolong the network lifetime by balancing the energy consumption among the aggregate nodes, we expect that each aggregate node consumes the same amount of energy roughly per data gathering session. Thus, the network lifetime is maximized when all aggregate nodes run out their energy and die at the same time. Clearly, $r_1 > r_2 > \dots > r_{k-1} > r_k$, because the data volume transmitted by aggregate node v_{i+1} is larger than that by aggregate node v_i , $1 \leq i \leq k-1$. Let r_0 be the distance between another endpoint of the middle ray line and aggregate node v_1 . There is a simple solution to the problem by formulating it into the following multiple-variable r_i non-

linear programming problem, $0 \leq i \leq k$.

$$\sum_{i=0}^k r_i = L. \quad (1)$$

Let F_i be the total volume of data transmitted from aggregate node v_i to its parent aggregate node v_{i+1} , following the assumption in Section II-D, we have

$$F_i = \frac{\theta}{2\pi}(\pi * L^2 - \pi * (\sum_{j=i}^k r_j)^2) = \frac{\theta}{2}(L^2 - (\sum_{j=i}^k r_j)^2), \quad (2)$$

$$r_{i+1}^2 F_{i+1} = r_i^2 F_i, \quad (3)$$

where $r_i^2 F_i$ is the transmission energy consumption of aggregate node v_i , $1 \leq i \leq k$. For this non-linear programming problem, it is very time consuming to find a solution even for a moderate problem size. Therefore, this simple algorithm essentially is inapplicable in practice.

A. Heuristic algorithm

We now propose a fast, scalable heuristic algorithm for the problem. Following $r_i < r_{i+1}$, we assume that $r_{i+1} = \sigma_{i+1} r_i$ with $0 < \sigma_{i+1} < 1$. To find solutions to all r_i , we need to figure out σ_i first by the following lemma, $1 \leq i \leq k$.

Lemma 1: Given a linear deployment of aggregate nodes in the middle ray line of a fan of angle θ with radius L , let s_i be the number of sensor nodes that propagate their sensing data to aggregate node v_i directly through multi-hop sensor node relay and R_i the trapezium of the area s_i covered by aggregate node v_i . Then, $\frac{s_{i+2}}{s_{i+1}} = \frac{R_{i+2}}{R_{i+1}} < \sigma_{i+1}$, $0 \leq i \leq k-1$.

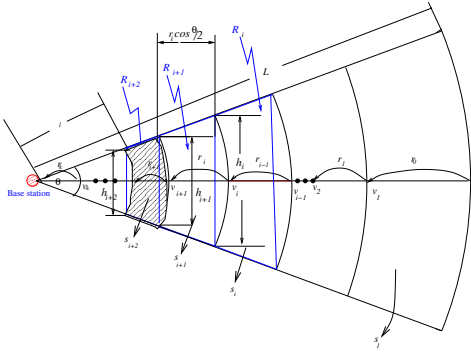


Fig. 1. A fan of angle θ and radius L

Proof: Following the illustration in Fig. 1,

$$\begin{aligned} \frac{s_{i+2}}{s_{i+1}} &= \frac{R_{i+2}}{R_{i+1}} = \frac{\frac{h_{i+2}+h_{i+1}}{2}(r_{i+1} \cos \frac{\theta}{2})}{\frac{h_{i+1}+h_i}{2}(r_i \cos \frac{\theta}{2})} \\ &< \frac{\frac{h_{i+1}+h_{i+1}}{2}(r_{i+1} \cos \frac{\theta}{2})}{\frac{h_{i+1}+h_{i+1}}{2}(r_i \cos \frac{\theta}{2})} = \frac{r_{i+1}}{r_i} = \sigma_{i+1}. \end{aligned} \quad (4)$$

For all i , $1 \leq i \leq k$, we have

$$F_i = \theta/(2\pi)[\pi L^2 - \pi(L - r_0)^2], \quad (5)$$

$$\begin{aligned} F_i &= s_i + F_{i-1} \leq \sigma_{i-1} s_{i-1} + F_{i-1} \\ &\leq \prod_{j=1}^{i-1} \sigma_j F_1 + F_{i-1} = \sum_{l=1}^{i-1} (\prod_{j=1}^l \sigma_j F_1). \end{aligned} \quad (6)$$

$$\sigma_1 = 1. \quad (7)$$

For convenience, in the rest of the paper we treat Ineq. (4) and (6) as equations. Let

$$\rho_i = \frac{s_i}{F_i} = \frac{\prod_{j=1}^{i-1} \sigma_j F_1}{\sum_{l=1}^{i-1} (\prod_{j=1}^l \sigma_j F_1)} = \frac{\prod_{j=1}^{i-1} \sigma_j}{\sum_{l=1}^{i-1} (\prod_{j=1}^l \sigma_j)}, \quad (8)$$

since $\sigma_i \sigma_{i-1} \dots \sigma_1 < \sigma_{i-1} \dots \sigma_1$. Clearly, $\rho_1 = 1$ and $\rho_2 = \frac{1}{2}$. Given $r_{i+1} = \sigma_{i+1} r_i$ and Eq. (3), we have

$$\begin{aligned} r_{i+1}^2 F_{i+1} &= (\sigma_{i+1} r_i)^2 F_{i+1} = (\sigma_{i+1} r_i)^2 (s_{i+1} + F_i) \\ &= (\sigma_{i+1} r_i)^2 (1 + \frac{s_{i+1}}{F_i}) F_i = (\sigma_{i+1} r_i)^2 (1 + \frac{\sigma_i s_i}{F_i}) F_i \\ &\leq (\sigma_{i+1} r_i)^2 (1 + \sigma_i \rho_i) F_i. \end{aligned} \quad (9)$$

Combined with Eq. (3), we have

$$\sigma_{i+1}^2 (1 + \sigma_i \rho_i) \leq 1; \sigma_{i+1} \leq \sqrt{\frac{1}{1 + \sigma_i \rho_i}}. \quad (10)$$

Given σ_2 and ρ_2 , the rest of σ_i and ρ_i can be computed by Eq. (8) and Eq. (10), $i \geq 3$. To determine the value of r_i , $2 \leq i \leq k$, we have

$$r_i = \sigma_i r_{i-1} = \prod_{j=1}^i \sigma_j r_1. \quad (11)$$

The problem then is reduced to find a solution for both r_1 and r_0 . Combined with Eq. (11), Eq. (1) is rewritten as follows.

$$\sum_{i=1}^k \prod_{j=1}^i \sigma_j r_1 = L - r_0. \quad (12)$$

Meanwhile, Eq. (3) implies that aggregate node v_1 and aggregate node v_k consume the same amounts of energy per data gathering session, then

$$r_1^2 F_1 = r_k^2 F_k, \quad (13)$$

Eq. (13) can be rewritten below.

$$2r_0 L - r_0^2 = (\prod_{i=1}^k \sigma_i)^2 [L^2 - (\prod_{i=1}^k \sigma_i)^2 r_1^2]. \quad (14)$$

The solution to variables r_1 and r_0 then can be obtained, by solving Eq. (12) and Eq. (14).

IV. AGGREGATE NODE PLACEMENTS

Following the same assumption as in [1], [4] that the monitored region is a circle of radius L , suppose that the sensor nodes have been redundantly deployed in the circle already. We propose a heuristic solution to the problem by partitioning the circle into a number of fans of angle θ and allocating each fan with k ($= \lfloor \frac{K\theta}{2\pi} \rfloor$) aggregate nodes. Assume that the k aggregate nodes assigned to each fan are arranged into the middle ray line of the fan and the distance between aggregate node v_i and aggregate node v_{i+1} is r_i , $1 \leq i \leq k-1$, the problem now is reduced to minimize the maximum energy

consumption among aggregate nodes in the middle ray line, which can be expressed as follows.

$$\text{minimize } \max_{1 \leq i \leq k} \left\{ \frac{\theta}{2\pi} r_i^\kappa (\pi L^2 - \pi (\sum_{j=i}^k r_j)^2) \right\},$$

where $\frac{\theta}{2\pi} r_i^\kappa (\pi L^2 - \pi (\sum_{j=i}^k r_j)^2)$ is the transmission energy consumption of aggregate node v_i , κ is the path loss exponent which is typically between 2 and 4. For convenience, we assume $\kappa = 2$. As a result, the problem is to find r_i and θ such that $\max_{1 \leq i \leq k} \left\{ \theta r_i^2 \frac{L^2 - (\sum_{j=i}^k r_j)^2}{2} \right\}$ is minimized, $1 \leq i \leq k$.

A. Algorithm without any constraints

Theorem 1: Given a few powerful aggregate nodes K , assume that a circle is partitioned into f fans of angle θ and each fan is assigned with $k = \lfloor \frac{K}{f} \rfloor$ aggregate nodes, which are placed in the middle ray line of the fan. To maximize the lifetime of the resulting network, the number of fans in the fan partition is no more than 4, i.e., $f \leq 4$.

Proof: Let f be the number of fans in the current partition with $f > 4$. To show the claim, we need to show that there is always another better fan partition which partitions the circle into f' fans such that the network lifetime is longer than that by the current fan partition with $f' < f$. Let $T(\theta, k, L, f)$ be the maximum energy consumption among aggregate nodes delivered by the proposed algorithm in the previous section through partitioning the circle of radius L into f fans of angle θ , and k aggregate nodes are deployed into the middle ray line of each fan. There is another fan partition, which partitions the circle into f' fans and each fan is assigned $2k$ aggregate nodes at least, where $f' = f/2$ and the fan angle is 2θ if f is even; otherwise, $f' = \frac{f-1}{2}$ and the fan angle is $\frac{4\pi}{f-1}$ ($= 2\theta + \frac{2\theta}{f-1}$). We claim that this latter fan partition has a longer network lifetime. In other words, we claim: (i) $\frac{T(\theta, k, L, f)}{T(2\theta, 2k, L, f/2)} \geq 4/3 = 1.33$ if $f > 2$ is even; otherwise, (ii) $\frac{T(\theta, k, L, f)}{T(\frac{4\pi}{f-1}, 2k, L, \frac{f-1}{2})} \geq 16/15 \approx 1.05$ when $f \geq 5$.

We first show claim (i) through the following digram (see Fig. 2). Assume that there are two fans of angle θ . Each has

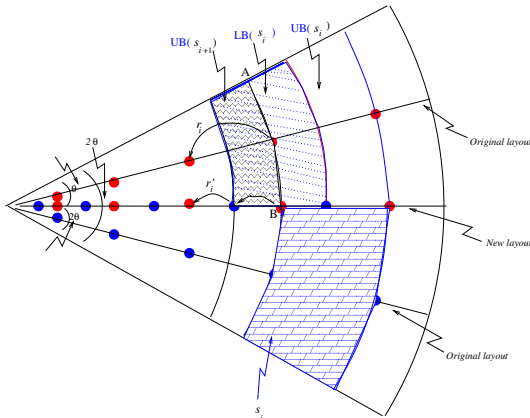


Fig. 2. Fan merging with each being angle θ and k aggregate nodes

been assigned k aggregate nodes to aggregate the sensed data

from its fan area. Now, a fan of angle 2θ is formed by merging them. We reallocate the $2k$ aggregate nodes in the middle ray line of the merged fan as follows.

The first k aggregate nodes are placed in the line exactly at the same positions as they were placed in the middle ray line of the fan of angle θ . For each of the remaining k aggregate nodes, insert it into the half way of the two corresponding placed neighboring nodes in the line. As a result, for each aggregate node v_i in this new placement schema, the transmission distance between v_i and its parent is a half of that in the original placement. Meanwhile, the total amount of data transmitted by v_i is various, which is analyzed as follows. Let r_i and r'_i be the transmission distances between aggregate node v_i and its parent v_{i+1} before and after fan merging, and F_i and F'_i the volume of data transmitted by v_i to its parent v_{i+1} before and after fan merging, respectively. Following the fact that the transmission energy consumption at a node is super-linear to the transmission distance between the node and its receiver, we have two cases to deal with.

Case 1: the distance of v_i to the base station in the new placement is identical to its original one, but its transmission distance to its parent is a half of its original one, the amount of data transmitted by v_i is twice as much as its original one. The energy consumption at v_i thus is

$$\begin{aligned} T(2\theta, 2k, L, f', v_i) &= r_i'^2 F'_i \leq \left(\frac{r_i}{2}\right)^2 (2F_i) = r_i^2 F_i / 2 \\ &= T(\theta, k, L, f, v_i) / 2. \end{aligned} \quad (15)$$

Case 2: the distance of v_i to the base station in the new placement is equal to its original distance minus $r_i/2$ (v_i is placed in the half way of two neighboring nodes), i.e., $r'_i = r_i/2$. Let $UB(s_i)$ and $LB(s_i)$ be the larger and the smaller curve areas of s_i by partitioning the area using its middle curve line \widehat{AB} (see Fig. 2). Clearly, $UB(s_i) > \frac{s_i}{2}$ while $LB(s_i) < \frac{s_i}{2}$, $2 \leq i \leq k$. To compute F'_i , s'_i needs to be computed first, where s'_i is the number of sensor nodes in an area that are covered by v_i in the new placement. However, it is obvious that s'_i is no greater than $2UB(s_{i+1})$, while $UB(s_{i+1}) \leq LB(s_i) \leq \frac{s_i}{2}$, illustrated by Fig. 2, we have

$$\begin{aligned} F'_i &= s'_i + 2F_i \leq 2UB(s_{i+1}) + 2F_i \\ &\leq s_i + 2F_i \leq 3F_i, \quad s_i \leq F_i. \end{aligned} \quad (16)$$

Having inequality (16), we have

$$\begin{aligned} T(2\theta, 2k, L, f', v_i) &= r_i'^2 * F'_i \leq r_i^2 / 4 * 3F_i \\ &= 3T(\theta, k, L, f, v_i) / 4. \end{aligned} \quad (17)$$

Thus, when f is even, we have

$$\begin{aligned} \frac{T(\theta, k, L, f)}{T(2\theta, 2k, L, f/2)} &= \frac{\max_{1 \leq i \leq k} \{T(\theta, k, L, f, v_i)\}}{\max_{1 \leq j \leq 2k} \{T(2\theta, 2k, L, f/2, v_j)\}} \\ &\geq \min\{2, 4/3\} = 1.33. \end{aligned} \quad (18)$$

We then show claim (ii) with similar arguments as for claim (i). We have the following observation. Let $S(\theta)$ be the fan area of a fan. The fan area $S(\frac{4\pi}{f-1})$ in the new fan partition of circle into f' fans of angle $\frac{4\pi}{f-1}$ ($= 2\theta + \frac{2\theta}{f-1}$) is $(2 + \frac{2}{f-1})S(\theta)$

when $f \geq 2$. We now place the $2k$ aggregate nodes into each of the f' fans by two cases.

Case 1: the distance of v_i to the base station in the new placement is identical to its original one, but the transmission distance between v_i and its parent is a half of its original one, the amount of data transmitted by v_i is $2 + \frac{2}{f-1}$ times as much as its original one. The energy consumption at v_i for the new node placement thus is

$$\begin{aligned} T\left(\frac{4\pi}{f-1}, 2k, L, \frac{f-1}{2}, v_i\right) &= r_i'^2 F'_i \leq \left(\frac{r_i}{2}\right)^2 \left(2 + \frac{2}{f-1}\right) F_i \\ &= \frac{1}{2} \left(1 + \frac{1}{f-1}\right) r_i^2 F_i = \frac{1}{2} \left(1 + \frac{1}{f-1}\right) T(\theta, k, L, f, v_i) \\ &= \frac{f}{2(f-1)} T(\theta, k, L, f, v_i) < T(\theta, k, L, f, v_i), \quad \text{if } f \geq 3. \end{aligned} \quad (19)$$

Case 2: the distance of v_i to the base station in the new placement is equal to its original distance minus $r_i/2$ (v_i is placed in the half way of two neighboring nodes), i.e., $r'_i = r_i/2$. To compute F'_i , s'_i needs to be computed first. However, it is obvious that s'_i is no greater than $(2 + \frac{2}{f-1})s_{i+1}$, we thus have

$$\begin{aligned} F'_i &= s'_i + \left(2 + \frac{2}{f-1}\right) F_i \\ &\leq \left(2 + \frac{2}{f-1}\right) UB(s_{i+1}) + \left(2 + \frac{2}{f-1}\right) F_i \\ &\leq \left(1 + \frac{1}{f-1}\right) 2LB(s_i) + \left(2 + \frac{2}{f-1}\right) F_i \\ &\leq \left(1 + \frac{1}{f-1}\right) s_i + 2\left(1 + \frac{1}{f-1}\right) F_i \\ &\leq \frac{3f}{f-1} F_i, \quad s_i \leq F_i. \end{aligned} \quad (20)$$

We have

$$\begin{aligned} T\left(\frac{4\pi}{f-1}, 2k, L, \frac{f-1}{2}, v_i\right) &= r_i'^2 * F'_i \leq \frac{r_i^2}{4} \left(\frac{3f}{f-1}\right) F_i \\ &= \frac{3f}{4(f-1)} r_i^2 F_i = \frac{3f}{4(f-1)} T(\theta, k, L, f, v_i) \\ &< T(\theta, k, L, f, v_i), \quad \text{if } f \geq 5. \end{aligned} \quad (21)$$

To summarize Cases 1 and 2, we have

$$\frac{T(\theta, k, L, f)}{T\left(\frac{4\pi}{f-1}, 2k, L, \frac{f-1}{2}\right)} = \frac{\max_{1 \leq i \leq k} \{T(\theta, k, L, f, v_i)\}}{\max_{1 \leq j \leq 2k} \{T\left(\frac{4\pi}{f-1}, 2k, L, \frac{f-1}{2}, v_j\right)\}} > 1, \quad \text{if } f \geq 5. \quad (22)$$

Theorem 1 implies, the larger the value of fan angle θ , the longer the network lifetime will be, given the number of aggregate nodes K . To maximize the network lifetime, the maximum number of fans obtained by the circle partition is no more than 4. We thus have the following heuristic algorithm.

Algorithm Lifetime_without_const(G, L, K)

1. $net_lifetime \leftarrow \infty$; $f \leftarrow 5$; $\theta \leftarrow \frac{2\pi}{f}$; $k \leftarrow \lfloor \frac{K}{f} \rfloor$;
/* k is the number of nodes assigned to each fan area */
2. $control \leftarrow 'true'$; $\rho_2 \leftarrow 1/2$; $\sigma_2 \leftarrow \frac{1}{\sqrt{2}}$;
3. while $control$ do

4. for $i \leftarrow 3$ to k do
5. $\sigma_i \leftarrow \sqrt{\frac{1}{1 + \sigma_{i-1} \rho_{i-1}}}$; $\rho_i \leftarrow \frac{\prod_{j=1}^{i-1} \sigma_j}{\sum_{l=1}^{i-1} (\prod_{j=1}^l \sigma_j)}$;
6. endfor;
7. compute r_0 and r_1 , using Eqs. (12) and (14)
8. for $i \leftarrow 2$ to k do
9. $r_i \leftarrow \prod_{j=1}^i \sigma_j r_1$;
10. endfor;
11. $T(\theta, k, L, f) \leftarrow \max_{1 \leq i \leq k} \{\theta r_i^2 \frac{L^2 - (\sum_{j=i}^k r_j)^2}{2}\}$
12. if $net_lifetime > T(\theta, k, L, f)$ then
13. $net_lifetime \leftarrow T(\theta, k, L, f)$;
14. $\theta_0 \leftarrow \theta$; $k_0 \leftarrow k$; $f_0 \leftarrow f$;
15. $f \leftarrow \lfloor f/2 \rfloor$; $\theta \leftarrow \frac{2\pi}{f}$; $k \leftarrow \lfloor \frac{K}{f} \rfloor$;
16. else $control \leftarrow 'false'$;
17. endif;
18. endwhile;
19. return f_0, θ_0 and k_0 .

It can be seen that this approach suffers a serious drawback by considering a scenario where the circle is partitioned into two fans of fan angle π . Within each fan, the $k = K/2$ aggregate nodes are arranged into its middle ray line, and the lifetime of the resulting network will be maximized when performing data gathering, following Theorem 1. However, under such a placement schema, the maximum transmission delay between a sensor node and the aggregate node that covers the sensor node is as large as $\sqrt{L^2 + (L - r_0)^2} \approx \sqrt{2}L$, which is unacceptable for many real applications, because the sensor node takes $\lceil \frac{\sqrt{2}L}{t_x} \rceil$ hops sensor node relays to reach its cluster head. As each sensor node is severely energy constrained and the message length transmitted by the sensor node is at least proportional to the number of hops, to perform message transmission to the cluster head, a member sensor node near to the cluster head will run out of its battery very quickly, where t_x is the fixed transmission radius of a sensor node.

B. Algorithm with constraints of R_{max} and D_{max}

In this subsection we consider a realistic communication model where the maximum transmission delay between a sensor node and an aggregate node that covers the sensor node is bounded by D_{max} . In addition, the maximum transmission range of an aggregate node is bounded by R_{max} . Under the constraints of both R_{max} and D_{max} , we deal with the aggregate node placement problem. To meet these two constraints, the minimum number k_{min} of aggregate nodes in the middle ray line of a fan is at least $\lceil \frac{L - D_{max}}{R_{max}} \rceil$, since the following inequality must be met.

$$k_{min} R_{max} + D_{max} \geq L. \quad (23)$$

Otherwise, the communication graph induced by the aggregate nodes is disconnected. Following Theorem 1, if the aggregate nodes deployed in the circle are organized as a star structure centered at the circle center, then, the larger the θ value, the longer the network lifetime will be. On the other hand, with the growth of the value of θ , the transmission delay becomes

bigger and bigger, which will be beyond the maximum transmission delay D_{max} eventually. Thus, the value of θ cannot be arbitrarily large. The rest is to address the choice of a proper θ to meet the constraints.

Following placing aggregate nodes into the middle ray line of a fan, given a fan area with fan angle θ , the sensing area s_1 , in which the sensor nodes are covered by aggregate node v_1 , is the largest one among the subareas covered by the k aggregate nodes in the fan area. We observe that, if the transmission delay (distance) between the farthest corner sensor node in s_1 and aggregate node v_1 is no greater than D_{max} , then, there is no any other sensor node in s_1 whose distance to v_1 is larger than D_{max} . Furthermore, for any other area s_i , the maximum transmission delay between a sensor node in s_i and v_i covering the sensor node is no more than D_{max} too, $2 \leq i \leq k$.

Having this observation, we can determine θ and each r_i to meet the maximum transmission delay constraint, by solving the following equations and inequalities. Fig. 3 illustrates the constraints.

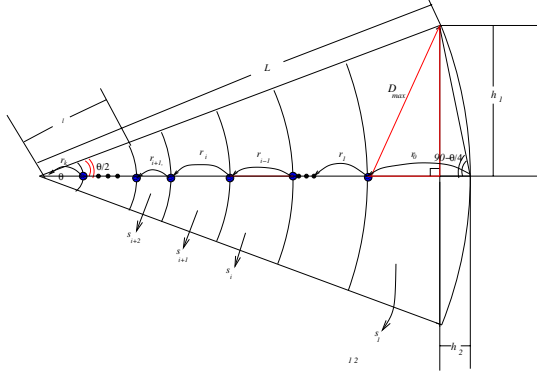


Fig. 3. A circle is partitioned into a number of fans of angle θ with the maximum transmission delay D_{max} .

$$h_1 = L \sin \theta / 2; h_2 = h_1 / \tan(90^\circ - \theta / 4), \quad (24)$$

and

$$(r_0 - h_2)^2 + h_1^2 \leq D_{max}^2. \quad (25)$$

We thus have

$$r_0 \leq \frac{L \sin \frac{\theta}{2}}{\tan(90^\circ - \frac{\theta}{4})} + \sqrt{D_{max}^2 - L^2 (\sin \frac{\theta}{2})^2}, \quad (26)$$

$$r_0 \leq D_{max}; k = \lfloor \frac{K\theta}{2\pi} \rfloor, \quad (27)$$

Given $\rho_2 = 1/2$ and $\sigma_2 = \frac{1}{\sqrt{2}}$, the rest of σ_i and ρ_i can be computed for all $i \geq 3$. Given r_0 and r_1 , we then compute the value of r_i for all i , $2 \leq i \leq k$, where $r_i = \prod_{j=1}^i \sigma_j r_1$. We now incorporate the maximum transmission range R_{max} of an aggregate node into the design of the proposed algorithm. It is obvious that the maximum transmission range between two aggregate nodes in the middle ray line of a fan is the one between aggregate node v_1 and aggregate node v_2 , i.e., r_1 , since $r_{i+1} < r_i$ for all i , $1 \leq i \leq k$. Therefore, $r_1 \leq R_{max}$. In summary, we have the following theorem.

Theorem 2: Given a dense sensor network of n sensor nodes to monitor a circle area of radius L with base station at the circle center, assume that the circle is partitioned into a number of fans of angle θ , $1 \leq \theta \leq \pi/2$. Let R_{max} be the maximum transmission range of each aggregate node and D_{max} the maximum transmission delay between a sensor node and the aggregate node that covers the sensor node. Let K ($K \ll n$) be the number of powerful aggregate nodes. To maximize the lifetime of the resulting network by placing the K aggregate nodes into the dense sensor network, there is an approach of placing the K aggregate nodes into the middle ray lines of fans to form a star structure such that both constraints are met, the values of θ and r_i can be found by solving the above equations and inequalities.

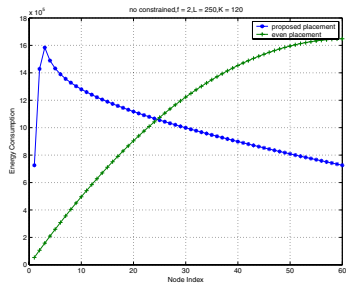
In the following a greedy algorithm is proposed, based on Theorem 2. Let f be the number of fans. We start to partition the circle into a number of fans of identical angle, and each fan is assigned with k_{min} aggregate nodes initially. It is obvious that $f = \lfloor \frac{K}{k_{min}} \rfloor$. The algorithm proceeds in iterations. Let f be the number of fans of angle θ at the current iteration. If there is a feasible solution to meet all constraints at the current iteration, we then check whether f is even. If it is, following Theorem 2 we merge every two neighboring fans into a bigger fan. As a result, there are $f/2$ fans of angle 2θ , and each fan is assigned $\lfloor \frac{2K}{f} \rfloor$ aggregate nodes at the next iteration. Otherwise, there are $\frac{f-1}{2}$ fans of angle $\frac{4\pi}{f-1}$ ($= 2\theta + \frac{2\theta}{f-1}$), and each fan is assigned $\lfloor \frac{2K}{f-1} \rfloor$ aggregate nodes at the next iteration. If there is no solution to meet the constraints at the current iteration, then, the setting value of fan angle θ at the current iteration is too large, and the fan angle θ' in the potential solution is such a maximum value between $\theta/2$ and θ that meets the constraints. Instead of working on the value of θ' directly, we use the number of aggregate nodes assigned to a fan to derive the value of θ' .

V. PERFORMANCE EVALUATION

In this section we evaluate the performance of the proposed algorithm against another commonly equal distance placement schema that has been used as the benchmark in their studies [4], [1]. In all our experiments we assume that the radius of the circle $L = 250$ meters and $n = \lceil \pi L^2 \rceil$ sensor nodes have been deployed in the circle evenly.

A. Energy consumption at individual aggregate nodes

We first examine the energy consumption at each individual aggregate node between the proposed algorithm and the uniform placement schema. Fig. V-A (a) depicts the energy consumption of each aggregate node when the circle is partitioned into two fans of angle π and each fan is deployed $k = K/2 = 120/2 = 60$ aggregate nodes. From this figure we can see that aggregate node v_k in the uniform placement schema has the maximum energy consumption, because it not only transmits the sensed data by sensor nodes in area s_k but also relays the sensed data for all the other aggregate nodes in the fan. The experimental results indicate that aggregate node v_3 in the proposed algorithm has the maximum energy

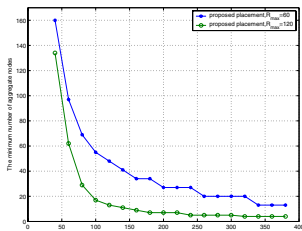


(a) The energy consumption at each aggregate node indexed from 1 to 60

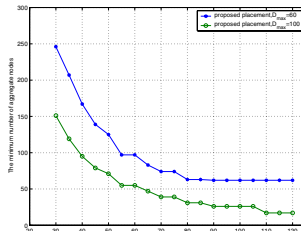
consumption, and both algorithms favor placing as many aggregate nodes in a fan as possible to reduce the maximum energy consumption among aggregate nodes.

B. The minimum number of aggregate nodes

We then investigate the minimum number of aggregate nodes needed by the proposed algorithm to meet the constraints of D_{max} and R_{max} . Fig. 4(a) shows that the minimum number of aggregate nodes needed is significantly reduced, when R_{max} is doubled from its initial value - 60 meters, whereas D_{max} is maintained from 70 meters to 250 meters. Fig. 4(b) implies that there is a big gap in terms of the number of aggregate nodes needed when D_{max} is fixed at either 60 meters or 100 meters, while R_{max} is ranged from 30 meters to 60 meters. The experimental results indicate that the proposed placement schema requires more aggregate nodes when R_{max} is small.



(b) The minimum number of aggregate nodes needed when $R_{max} = 60m, 120m$ is fixed.

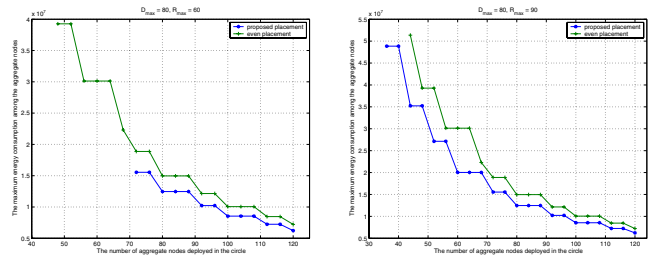


(c) The minimum number of aggregate nodes needed when $D_{max} = 60m, 100m$ is fixed.

Fig. 4. The minimum number of aggregate nodes K_{min} is needed by the proposed algorithm to meet different D_{max} and R_{max} constraints.

C. Performance evaluation with distortion constraints

We finally evaluate the performance of the proposed algorithm against the uniform placement schema, in terms of the maximum energy consumption among the aggregate nodes under different distortion constraints. Fig. 5 plots the maximum energy consumption among the aggregate nodes for both algorithms. Clearly, the maximum energy consumption among aggregate nodes by the proposed placement schema is always less than that by the uniform placement schema, if the number of aggregate nodes for the proposed algorithm is no less than $K_{min}(D_{max}, R_{max})$ that is the minimum number of aggregate nodes needed by the proposed algorithm, when D_{max} and R_{max} are given.



(a) $R_{max} = 60m$ and $D_{max} = 80m$ (b) $R_{max} = 90m$ and $D_{max} = 80m$

Fig. 5. The maximum energy consumption among the aggregate nodes for different combinations of $R_{max} = 60$ meters, $D_{max} = 80$ meters.

VI. CONCLUSIONS

In this paper we studied the aggregate node placement problem by placing a few powerful aggregate nodes into a dense sensor network such that the lifetime of the resulting network is maximized when performing data gathering. We devised a fast, scalable heuristic algorithm for it. We also conducted experiments by simulation to evaluate the performance of the proposed algorithm. The experimental results show that the proposed heuristic significantly outperforms the uniform placement schema in the prolongation of network lifetime.

ACKNOWLEDGMENT

It is acknowledged that the work by Weifa Liang is partially funded by Australian Research Council under its Discovery Schemes with Grant No:DP0449431 and the Faculty Research Grant of Australian National University. The work by Yinlong Xu is partially supported by the Natural Science Foundation of China under Grant No. 60773036.

REFERENCES

- [1] P. Cheng, C-N Chuah, and X. Liu. Energy-aware node placement in wireless sensor networks. *Proc. Globecom'04*, IEEE, 2004.
- [2] J-H Chang and L. Tassiulas. Energy conserving routing in wireless ad hoc networks. *Proc. INFOCOM'00*, IEEE, 2000.
- [3] K. Dasgupta, et al. Topology-aware placement and role assignment for energy-efficient information gathering in sensor networks. *Proc. The 8th IEEE Symposium on Computers and Communications*, IEEE, 2003.
- [4] D. Ganesan, et al. Power-efficient sensor placement and transmission structure for data gathering under distortion constraints. *ACM Trans. Sensor Networks*, Vol. 2, No. 2, pp.155-181, 2006.
- [5] A. Goel and D. Estrin. Simultaneous optimization for concave costs: single sink aggregation or single source buy-at-bulk. *Proc. SODA'03*, ACM-SIAM, 2003.
- [6] H. Gupta, et al. Connected sensor cover: self-organization of sensor networks for efficient query execution. *Proc. of MobiHoc'03*, ACM.
- [7] W. R. Heinzelman, et al. Energy-efficient communication protocol for wireless microsensor networks. *Proc. the Hawaii International Conference on System Sciences*, IEEE, 2000.
- [8] K. Kalpakis, et al. Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks. *Computer Networks*, Vol. 42, pp.697-716, 2003.
- [9] W. Liang, et al. Online data gathering for maximizing network lifetime in sensor networks. *IEEE Trans. Mobile Computing*, Vol.6, 2007.
- [10] S. Lindsey and C. S. Raghavendra. PEGASIS: Power-efficient gathering in sensor information systems. *Proc. Aerospace Conference*, IEEE, 2002.
- [11] S. Madden, et al. Supporting aggregate queries over ad hoc wireless sensor networks. *Proc. 4th IEEE Workshop on Mobile Computing and System Applications*, IEEE, 2002.
- [12] A. Singh, et al. Power-aware routing in mobile ad hoc networks. *Proc. of MobiCom*, ACM/IEEE, 1998.