

A Minimum Cost Heterogeneous Sensor Network with a Lifetime Constraint

Vivek P. Mhatre, *Student Member, IEEE*, Catherine Rosenberg, *Senior Member, IEEE*, Daniel Kofman, *Member, IEEE*, Ravi Mazumdar, *Senior Member, IEEE*, and Ness Shroff, *Senior Member, IEEE*

Abstract—We consider a heterogeneous sensor network in which nodes are to be deployed over a unit area for the purpose of surveillance. An aircraft visits the area periodically and gathers data about the activity in the area from the sensor nodes. There are two types of nodes that are distributed over the area using two-dimensional homogeneous Poisson point processes; type 0 nodes with intensity (average number per unit area) λ_0 and battery energy E_0 ; and type 1 nodes with intensity λ_1 and battery energy E_1 . Type 0 nodes do the sensing while type 1 nodes act as the cluster heads besides doing the sensing. Nodes use multihopping to communicate with their closest cluster heads. We determine the optimum node intensities (λ_0, λ_1) and node energies (E_0, E_1) that guarantee a *lifetime* of at least T units, while ensuring connectivity and coverage of the surveillance area with a high probability. We minimize the overall cost of the network under these constraints. Lifetime is defined as the number of successful data gathering trips (or cycles) that are possible until connectivity and/or coverage are lost. Conditions for a sharp cutoff are also taken into account, i.e., we ensure that almost all the nodes run out of energy at about the same time so that there is very little energy waste due to residual energy. We compare the results for random deployment with those of a grid deployment in which nodes are placed deterministically along grid points. We observe that in both cases λ_1 scales approximately as $\sqrt{\lambda_0}$. Our results can be directly extended to take into account unreliable nodes.

Index Terms—Sensor networks, energy, lifetime, stochastic geometry, Voronoi cells.

1 INTRODUCTION

SENSOR networks are dense networks of low cost, wireless nodes that sense certain phenomena in the area of interest and report their observations to a central base station for further analysis. Sensor networks have been predicted to have a wide range of applications in both civilian as well as military domains [2]. An important application of sensor networks is surveillance of a battlefield or sensitive borders of countries. A simple way to monitor such areas is to deploy sensors. Deployment could either be deterministic, i.e., placing the nodes along grid points (may not be practical in general), or the nodes could be deployed randomly (e.g., from an aircraft). In this paper, we study a scenario in which an aircraft (possibly unmanned) or a LEO satellite passes over these areas periodically and collects updates from the deployed nodes. Thus, in the above scenario, the aircraft acts as the (mobile) base station. We assume that the nodes are organized as clusters and the cluster heads perform data aggregation. Data aggregation is used because it reduces the amount of data that is sent from the sensor nodes to the base station and, thereby, improves energy efficiency of the network [3], [4].

We consider a heterogeneous network with two types of nodes, type 0 nodes deployed with intensity λ_0 and battery energy E_0 , and type 1 nodes deployed with intensity λ_1 and battery energy E_1 . Type 0 nodes do the basic sensing as well as the relaying of packets since multihop communication is used within each cluster. Type 1 nodes are the cluster heads. They do data fusion within each cluster and directly transmit the aggregated data to the aircraft. Type 1 nodes also participate in sensing. Since type 0 nodes communicate over short range, their hardware requirements are simple. On the other hand, type 1 nodes perform long range transmissions to the aircraft, perform data aggregation, and coordinate MAC and routing within the clusters. Hence, the type 1 nodes have more complex hardware than the type 0 nodes. We first assume that all the nodes are reliable. We show later that our analysis can be easily extended to the case of unreliable nodes.

Each visit of the aircraft triggers a sensing and data gathering cycle on the ground during which every node sends a packet to its cluster head. This model of periodic monitoring is different from an event detection sensor network where nodes send data only when an event occurs. The total battery energy of each node is finite and, so, there is a limit on the number of successful data gathering trips that the aircraft can make. Our objective is to determine the optimum node deployment parameters that will ensure a certain minimum number of successful data gathering cycles before the sensor system becomes unusable, i.e., connectivity and/or coverage can no longer be ensured. Each type of node has a cost function associated with it that takes into account its hardware and battery cost. We

• V.P. Mhatre, C. Rosenberg, R. Mazumdar and N. Shroff are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907.
E-mail: {mhatre, cath, mazum, shroff}@ecn.purdue.edu.

• D. Kofman is with ENST, 46, rue Barrault 75634 Paris Cedex 13 France.
E-mail: daniel.kofman@enst.fr.

Manuscript received 4 Mar. 2003; revised 25 Sept. 2003; accepted 10 Jan. 2004; published online 1 Dec. 2004.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-0032-0303.

formulate an optimization problem with the above mentioned constraints and find a solution that minimizes the overall cost of the network. We dimension the network (in terms of cluster head intensity and battery energies of the nodes) so that both types of nodes expire at about the same time in order to reduce waste of residual energy. We study two deployment scenarios; namely, grid and random deployment and obtain results for λ_0 , E_0 , λ_1 , and E_1 for both scenarios. We observe that, in both scenarios, λ_1 , i.e., the cluster head intensity scales approximately as $\sqrt{\lambda_0}$.

This paper is organized as follows: In Section 2, we discuss some of the related work. In Section 3, we first discuss the models that we use for communication, connectivity and coverage, lifetime, etc., and then formulate the design problem. In Section 4, we provide an exact solution to the problem and then obtain an approximate solution for some typical radio parameter settings. Section 5 contains some numerical results that validate the approximations that we make in our analysis and give insights on the nature of solutions that we get. We conclude in Section 6. The proofs of some of the results that we use in Sections 3 and 4 are provided in Appendices A and B.

2 RELATED WORK

In [12], Heinzelman et al. consider a homogeneous clustered network in which each cluster head collects data from its one hop neighbors, aggregates the gathered data, and transmits it directly to the remote base station. The cluster heads are periodically rotated for efficient load balancing. All the nodes in the network are identical, and there is no multihop communication. In this scenario, the authors provide results supporting their idea that, even though the cluster heads have the highest energy drainage rate, periodic rotation ensures good load balancing and, hence, a higher lifetime. It has also been shown in [12] that, in order to minimize the total energy spent in the network, the required number of cluster heads has to scale as the square root of the total number of sensor nodes; a result similar to what we obtain.

The scenario that we study consists of a multihop network with heterogeneous nodes. In [12], cluster head rotation requires that all the nodes be capable of performing data aggregation as well as long range transmissions to the remote base station. This results in extra hardware complexity in all the nodes. In contrast, in our approach, the above complexity is embedded in only a few nodes (type 1 nodes). As a result, the vast majority of the nodes (type 0 nodes) have less complex hardware. Also, in our system, since the cluster heads are predetermined and fixed, there is no need for a cluster head election protocol.

In [11], Bandyopadhyay and Coyle consider a homogeneous sensor network in which nodes are uniformly deployed using a two-dimensional Poisson point process over a unit area. Cluster heads are chosen from among these nodes randomly with a probability p . The nodes send their data to their closest cluster head node. This leads to the formation of Voronoi cells wherein the cluster heads are the nuclei of the cells. The nodes use multihopping to communicate with the cluster heads, and the cluster heads use multihopping to communicate with the base station.

The authors find an expression for the expected value of the total energy that is spent in the network during each data gathering cycle. The authors minimize this total energy to determine an optimum value of p .

In [7], Bhardwaj et al. provide loose bounds on the lifetime of a sensor network. In [20], the authors provide upper bounds on the lifetime of a sensor network by taking into account all the possible collaborative data gathering strategies over all the possible network routes. However, Bhardwaj and Chandrakasan [20] do not provide any practical ways to achieve these bounds. In [10], Kalpakis et al. provide bounds on the lifetime of a sensor network by assuming a communication model in which any node can communicate with any other node, as well as with the base station; clearly, a limited model. In [8], Chiasserini et al. determine an optimum strategy to allocate sensor nodes to different clusters so as to improve the system lifetime. In [22], Meguerdichian et al. study the coverage problem in a wireless sensor network by using Voronoi diagram tools. The authors propose an algorithm to determine a maximal breach path for a given sensor network topology. In [19], Zussman and Segall study energy-efficient routing in emergency sensor networks by using the principles of network flows. In [21], Cerpa and Estrin propose using redundant nodes to improve the connectivity of the network. Nodes use a distributed protocol to determine if they need to stay awake to improve the network connectivity. In [18], Gupta and Younis study a heterogeneous sensor network (with two types of nodes) and study the impact on clustering of node failure at the higher level of the node hierarchy.

In our approach, we observe that, in a sensor network, energy drainage is not uniform over the entire network. The cluster heads have the highest energy burden due to the long-range transmissions to the base station. Also, the nodes close to the cluster heads have a high energy burden due to relaying of packets. Hence, these nodes are likely to expire before other nodes. To determine a bound on the lifetime of the network, these observations have to be taken into account. Besides, it is also important to ensure a sharp cutoff effect, i.e., to ensure that almost all the nodes expire at about the same time when the network becomes unusable. This guarantees that there is very little residual energy left behind. The above related papers mainly focus on homogeneous networks and do not take all the above factors into account. We also note that in a heterogeneous network the right objective function to minimize is not the overall battery energy, but the overall network cost. This cost takes into account the cost of the battery as well as the node hardware. We also take into account the conditions for connectivity and coverage of the area by extending the results in [13]. In our approach, we model all the above factors as constraints of an optimization problem. Our results can be directly extended to the case of unreliable nodes.

3 PROBLEM FORMULATION AND MODELS USED

A circular disk of unit area is to be covered with sensor nodes. If the region to be covered has an area A , all the distances are normalized by dividing them by \sqrt{A} . There are two types of nodes; nodes of energy level E_0 deployed

with intensity λ_0 and nodes of energy level E_1 deployed with intensity λ_1 . A surveillance aircraft flying at an altitude of H sweeps the area periodically and triggers a data sensing *cycle* during which all the type 0 nodes in each cluster send their data to the closest type 1 node using multihop communication. The type 1 nodes aggregate the received data and then send it to the aircraft using a direct transmission. Thus, the aircraft serves as the remote base station. In this paper, we assume that the base station receives updates from every cluster. However, if the base station is interested in receiving updates from only a few clusters (extra sensitive regions), then our analysis can be modified as follows: Deploy more nodes over the regions of frequent updates and take these nodes into account in the overall network cost. The redundant nodes stay inactive while the battery energy of other nodes lasts and join the cluster when other nodes start to expire. A similar problem in which redundant nodes are used to improve network connectivity is studied in [21]. Thus, it is easy to modify our approach when frequent updates are desired from certain clusters. For the sake of simplicity, in the rest of this paper, we assume that all the clusters send information to the base station during every data gathering cycle.

There are two ways to deploy the nodes. In the first scenario (which is more realistic), nodes are thrown from an aircraft and this can be modeled using a two-dimensional homogeneous Poisson point process for each type of nodes. In the second scenario, nodes are deterministically placed along grid points. In the case of random deployment, λ_0 and λ_1 are the intensities of the corresponding (independent) Poisson point processes. In the case of random deployment, clustering leads to the formation of Voronoi cells with type 1 nodes being the nuclei of these cells. In the case of a grid, λ_0 and λ_1 are simply the number of type 0 and type 1 nodes that are used. The topology will consist of λ_1 equispaced type 1 nodes and λ_0 equispaced type 0 nodes placed along the grid points.

We now describe the various models that we use in our analysis, and we then formulate the design problem.

3.1 Cost Model

Let C_0 and C_1 be the cost per node for each type of node. Then, a simple model for a cost function is,

$$C_i = \alpha_i + \beta E_i,$$

where α_i and β are some constants that depend on the manufacturing process. The constant α_i is the cost of the hardware of a type i node (excluding the battery cost), while β is the proportionality constant for the battery cost. The overall cost of the network as a function of the vector $[\bar{\lambda}, \bar{E}] = [\lambda_0, \lambda_1, E_0, E_1]$ is

$$C(\bar{\lambda}, \bar{E}) = \alpha_0 \lambda_0 + \alpha_1 \lambda_1 + \beta(\lambda_0 E_0 + \lambda_1 E_1). \quad (1)$$

3.2 Connectivity and Coverage Model

For the sensor network to provide sensing coverage of the region and for the nodes to successfully use multihop communication, it is necessary to ensure that the conditions for node connectivity and area coverage be met. Results for connectivity and coverage of an area when unreliable nodes are placed along grid points have been obtained in [13].

Using a similar approach, we can study the case in which nodes are deployed over a unit area with a two-dimensional homogeneous Poisson point process. We assume that the sensing radius of each node is r . In reality, the model for node sensing coverage is highly application dependent. For example, there are scenarios in which the sensor nodes can only sense in a certain direction as in the case of certain acoustic sensors. However, we restrict ourselves to omnidirectional sensing in order to keep the model generic. For simplicity, we assume that r is also the critical distance between any two nodes for successful transmission. This distance r depends on the allowable signal to noise ratio for successful packet reception, modulation scheme, propagation loss exponent, etc. If the communication and sensing radii are different, we can use a similar approach to determine conditions for connectivity and coverage with some modifications (see [13]).

Let λ be the intensity of the Poisson process and p be the reliability probability of each node. The probability of connectedness of nodes and coverage of area is given as follows (see Appendix A).

$$\begin{aligned} &P(\text{network is connected and region is covered}) \\ &\geq 1 - \left(\frac{1}{\gamma r}\right)^2 e^{-\pi\theta^2 p r^2 \lambda} \quad (2) \\ &\forall \gamma, \theta > 0 \text{ such that } \gamma + 2\theta = 1. \end{aligned}$$

The result in (2) is similar in form to the result that the authors obtained in [13]. In [13], node placement is deterministic but randomness is induced due to node failure, while, in our scenario, randomness is due to the deployment process. Note that the result stated above is valid for $r \ll 1$, i.e., $\lambda \gg 1$. So, all the results are applicable only when the dimensions of the area are much larger than r . In practical applications where an area is to be covered by sensors, $r \ll 1$ indeed holds. This is because we are interested in deploying a large number of nodes, each having a coverage area that is much smaller than the total area of the region.

For simplicity, we assume that all the nodes are reliable, i.e., $p = 1$. The case of unreliable nodes is not fundamentally different since the functional form of the above relation is not affected. With unreliable nodes, we simply have to replace λ by λp (the resultant thinned process is also a Poisson process). In our scenario, there are two types of nodes with intensity λ_0 and λ_1 and both deployment processes are independent homogeneous Poisson point processes. Assuming that both types of nodes do the sensing and have the same sensing radius r ,

$$\begin{aligned} &P(\text{network is connected and region is covered}) \\ &\geq 1 - \left(\frac{1}{\gamma r}\right)^2 e^{-\pi\theta^2 p r^2 (\lambda_0 + \lambda_1)}. \quad (3) \end{aligned}$$

$\forall \gamma, \theta > 0$ such that $\gamma + 2\theta = 1$. In a network dimensioning problem, the designers provide a parameter ϵ such that the probability of connectivity and coverage be at least $1 - \epsilon$. Therefore, we require

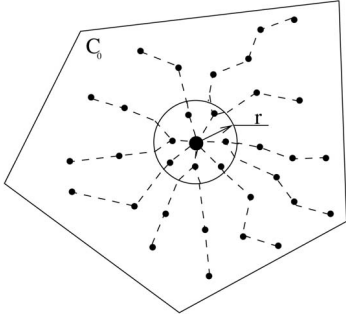


Fig. 1. A typical Voronoi cell.

$$\lambda_0 + \lambda_1 \geq \frac{1}{(\pi\theta^2 pr^2)} \log \left(\frac{1}{\epsilon(\gamma r)^2} \right). \quad (4)$$

$\forall \gamma, \theta > 0$ such that $\gamma + 2\theta = 1$. We assume that $\epsilon r^2 < 1$. Otherwise, (4) holds trivially for any positive value of $\lambda_0 + \lambda_1$. This scenario corresponds to the case when ϵ is large which is not of much interest to us. When $\epsilon r^2 < 1$, the right-hand side of (4) can be minimized as a function of θ under the constraint of $\gamma + 2\theta = 1$. This is because the right-hand side can be rewritten as a function of θ only (by eliminating γ)

$$u(\theta) = \frac{1}{(\pi\theta^2 pr^2)} \log \left(\frac{1}{\epsilon((1-2\theta)r)^2} \right). \quad (5)$$

Note that $u(\theta)$ approaches ∞ as θ approaches 0 as well as $1/2$. Hence, there is a point in between where $u(\theta)$ is minimized since $u(\cdot)$ is continuous and lower bounded by 0. In fact, we can show that there is a unique point where the function takes this minimum value. The proof is technical and, therefore, has been omitted. If θ_0 is the point where this minimum is attained (and correspondingly $\gamma_0 + 2\theta_0 = 1$), then ensuring $\lambda_0 + \lambda_1 \geq u(\theta_0)$ is sufficient for (4) to be true. Note that this θ_0 depends only on ϵ , p , and r which are constants.

Hence, the constraint in (4) reduces to a constraint of the form

$$\lambda_0 + \lambda_1 \geq u(\theta_0) = a, \quad (6)$$

where a is completely determined by ϵ , p , and r .

Note that, in the grid case, the required number of nodes is exactly $\lambda_0 + \lambda_1$ (all the nodes are assumed to be reliable). Hence, the connectivity-coverage requirement for a unit area takes the simple form:

$$\lambda_0 + \lambda_1 = 1/r^2. \quad (7)$$

3.3 Lifetime Constraint Model

We call those type 0 nodes which are within a distance r from a cluster head as the *critical nodes*. Since the transmission radius of each type 0 node is r , we observe that every transmission of a type 0 node to its cluster head has to go through one of these critical nodes. This is because the critical nodes are the last hop nodes for all the paths (see Fig. 1). Hence, among all the type 0 nodes in a cluster, the critical nodes have the highest burden of relaying data. As a result, these nodes are likely to exhaust their battery energy before other type 0 nodes. When critical nodes in a cluster

expire, connectivity is lost. We also observe that a cluster head has a high energy burden due to the long range transmissions to communicate with the aircraft, as well as due to the computations that it has to perform during data aggregation. Hence, the energy drainage rates of the critical nodes and the cluster heads determine the lifetime of the system. In light of these facts, it is natural to define the *lifetime* of the system to be the number of cycles until which all the cluster heads as well as all the critical nodes are active. In order to have a sharp cutoff effect, we also require that almost all the nodes in the network expire at about the same time. This ensures that very little residual energy (waste) is left behind when the system becomes unusable, i.e., when coverage and/or connectivity are lost. We cannot ensure a sharp cutoff for all the nodes in the network due to the inherent nonuniform nature of energy drainage in a cluster. For example, the type 0 nodes that are near the periphery of a cluster have very little relaying to do and, so, whenever the critical nodes or the cluster heads expire, the residual energy in the peripheral nodes is wasted. However, this is inevitable. The best we can do is to ensure that the two types of nodes which determine the lifetime of the system, i.e., the cluster heads and the critical nodes, expire at about the same time.

Let P_0 be the average energy spent by a typical critical node during each cycle, and let P_1 be the average energy spent by a typical cluster head during each cycle. Then, E_0/P_0 (respectively, E_1/P_1) is the average number of cycles that the critical nodes (respectively, the cluster heads) can sustain. For a sharp cutoff effect, we require:

$$\frac{E_1}{P_1} = \frac{E_0}{P_0}.$$

To ensure a lifetime of at least T cycles, we require

$$\frac{E_1}{P_1} = \frac{E_0}{P_0} \geq T. \quad (8)$$

3.4 Energy Model

Let P_0 be the amount of energy spent by a critical node during one cycle. This consists of energy spent on relaying packets of other nodes that are in the same cell (P_0^r), and transmitting one's own data (E_0^t per packet). We assume that type 0 nodes do not perform data fusion. We also assume that during each data gathering cycle a type 0 node sends one packet of its own to its cluster head. Let P_1 denote the amount of energy spent by a type 1 node during a single data gathering cycle. This consists of energy spent on receiving data from other nodes in the cell (E_0^r per packet), processing and compressing the received data (E_f per packet), and transmitting the compressed data to the aircraft (E_1^t per packet). We assume the radio model used in [12] wherein the energy required to transmit a packet over distance x is $l + \mu x^k$. l is the constant per packet energy spent in the transmitter electronics circuitry while μx^k is the energy spent in the RF amplifier to counter the propagation loss. The energy required to receive a packet is just l . Hence,

$$E_1^t = l + \mu H^k, \quad E_0^t = l + \mu r^k, \quad E_0^r = l. \quad (9)$$

We also assume that the cluster heads coordinate MAC and routing in their respective clusters so that no energy is wasted on packet collisions or idle listening (ideal MAC assumption). A cluster head performs fusion of the data packets that it receives from all the sensors in its cluster, and transmits a single packet to the aircraft during each cycle.

$$P_0 = E_0^t + P_0^r, \quad (10)$$

$$P_1 = \mathbf{E}[N_v](E_0^r + E_f) + E_1^t, \quad (11)$$

where E_f is the processing energy spent on data fusion. We denote by $\mathbf{E}[N_v]$ the expected number of type 0 nodes in a typical cluster. Note that the type 0 nodes use multihopping to communicate with the cluster head, but the cluster head still has to receive as many as $\mathbf{E}[N_v]$ packets during each cycle since no aggregation is performed at the type 0 nodes.

3.5 Problem Statement

We would like to determine $\bar{x} = [\lambda_0, \lambda_1, E_0, E_1]$, the parameters of the minimum cost network so that a lifetime of at least T cycles is guaranteed while ensuring connectivity and coverage with a probability of at least $1 - \epsilon$. Hence, we have the following optimization problem for the random deployment scenario:

$$\begin{aligned} & \text{minimize} && \alpha_0 \lambda_0 + \alpha_1 \lambda_1 + \beta(\lambda_0 E_0 + \lambda_1 E_1) \\ & \text{subject to} && \lambda_0 + \lambda_1 \geq a \\ & && \frac{E_1}{P_1} = \frac{E_0}{P_0} \geq T. \end{aligned}$$

For a grid deployment, the problem formulation is along similar lines with (7) as the connectivity-coverage constraint. Note that P_0 and P_1 are functions of λ_0 and λ_1 and are yet to be determined.

4 SOLUTIONS FOR THE TWO DEPLOYMENT SCENARIOS

4.1 Random Deployment

We begin the solution by first determining an expression for P_0^r , i.e., the energy spent by a critical type 0 node to relay packets. For that, we need some results from stochastic geometry [14].

When the deployment is random, each cluster is a Voronoi cell because sensor nodes choose the closest type 1 node as their cluster head. We first find the expected number of critical nodes in a typical Voronoi cell. This number is simply the expected number of type 0 nodes in a circle of radius r around a type 1 node. We then find the expected number of type 0 nodes in that Voronoi cell that lie outside this circle of radius r . From this, we can find the average relaying load on a critical node.

We use the approach used in [14] to determine the expected number of type 0 nodes in a cluster, as well as the expected number of critical type 0 nodes. Let $\sigma(\Pi_1)$ denote the sigma algebra generated by the point process corresponding to the type 1 nodes. Since type 0 nodes as well as type 1 nodes are deployed using a homogeneous Poisson point process, we can shift the origin to one of the type 1 points and use Campbell's theorem and Slivnyak's theorem [16] to compute the expected number of type 0 nodes in a typical Voronoi cell. Let \mathcal{C}_0 denote the Voronoi cell of a

typical type 1 node located at the origin, and $\{x_i \in \Pi_0\}$ denote the set of all the type 0 points. Then, $1_{\{x_i \in \mathcal{C}_0\}}$ is the indicator function which is one when a type 0 node i lies in cell \mathcal{C}_0 . Let $\mathbf{E}[N_v]$ be the expected number of type 0 nodes in cell \mathcal{C}_0 .

$$\begin{aligned} \mathbf{E}[N_v] &= \mathbf{E}\left[\sum_{x_i \in \Pi_0} 1_{\{x_i \in \mathcal{C}_0\}}\right] \\ &= \mathbf{E}\left[\mathbf{E}\left[\sum_{x_i \in \Pi_0} 1_{\{x_i \in \mathcal{C}_0\}} \mid \sigma(\Pi_1)\right]\right] \\ &= \int_0^{2\pi} \int_0^\infty e^{-\lambda_1 \pi x^2} \lambda_0 x dx d\theta. \end{aligned}$$

The event that a type 0 point located at (x, θ) belongs to the Voronoi cell \mathcal{C}_0 is equivalent to the event that there is a point of type 0 in a small area $x dx d\theta$ located at (x, θ) , and there is no other point of type 1 in a circle of radius x around that type 0 point. The latter condition ensures that this type 0 node belongs to the Voronoi cell \mathcal{C}_0 . From this, we get

$$\mathbf{E}[N_v] = \frac{\lambda_0}{\lambda_1}. \quad (12)$$

This is a well-known result. Using a similar argument, we can find the expected number of type 0 nodes located within a distance of r from a type 1 node as follows:

$$\begin{aligned} \mathbf{E}[N_v(r)] &= \mathbf{E}\left[\sum_{x_i \in \Pi_0} 1_{\{x_i \in \mathcal{C}_0, |x_i| < r\}}\right] \\ &= \mathbf{E}\left[\mathbf{E}\left[\sum_{x_i \in \Pi_0} 1_{\{x_i \in \mathcal{C}_0, |x_i| < r\}} \mid \sigma(\Pi_1)\right]\right] \\ &= \int_0^{2\pi} \int_0^r e^{-\lambda_1 \pi x^2} \lambda_0 x dx d\theta, \end{aligned}$$

which gives

$$\mathbf{E}[N_v(r)] = \frac{\lambda_0}{\lambda_1} (1 - e^{-\lambda_1 \pi r^2}). \quad (13)$$

These $\mathbf{E}[N_v(r)]$ number of critical nodes relay the data of the $\mathbf{E}[N_v] - \mathbf{E}[N_v(r)]$ type 0 nodes that are located outside the circle of radius r in the same Voronoi cell (see Fig. 1). Hence, the average relaying load on a typical critical node (P_0^r) is

$$\begin{aligned} P_0^r &= (E_0^r + E_0^t) \left(\frac{\mathbf{E}[N_v] - \mathbf{E}[N_v(r)]}{\mathbf{E}[N_v(r)]} \right) \\ &= (E_0^r + E_0^t) \left(\frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \right). \end{aligned}$$

Substituting the above in (10) and the value of $\mathbf{E}[N_v]$ from (12) in (11), we have the following:

$$\begin{aligned} P_0 &= E_0^t + (E_0^r + E_0^t) \left(\frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \right) \\ &= c_0 + c_1 \left(\frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \right). \end{aligned} \quad (14)$$

$$\begin{aligned}
P_1 &= (E_0^r + E_f) \left(\frac{\lambda_0}{\lambda_1} \right) + E_1^t \\
&= c_2 + c_3 \left(\frac{\lambda_0}{\lambda_1} \right).
\end{aligned} \tag{15}$$

We have introduced the constants c_0 , c_1 , c_2 , and c_3 for ease of notation. Combining the above equations with the equality constraint of (8), we have the following:

$$E_1 P_0 - E_0 P_1 = 0,$$

which gives

$$\begin{aligned}
c_1 E_1 \left(\frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \right) + c_0 E_1 - c_2 E_0 \\
- c_3 E_0 \left(\frac{\lambda_0}{\lambda_1} \right) = h(\bar{x}) = 0,
\end{aligned} \tag{16}$$

where, $\bar{x} = [\lambda_0, E_0, \lambda_1, E_1]$. We can further rewrite the coverage constraint from (6) as follows:

$$a - (\lambda_0 + \lambda_1) = g_1(\bar{x}) \leq 0. \tag{17}$$

Using the inequality constraint of (8),

$$E_1 \geq T P_1$$

with (15) gives:

$$c_2 T + c_3 T \left(\frac{\lambda_0}{\lambda_1} \right) - E_1 = g_2(\bar{x}) \leq 0. \tag{18}$$

Also, from (1), we have:

$$\begin{aligned}
C(\lambda_0, E_0, \lambda_1, E_1) &= f(\bar{x}) \\
&= \alpha_0 \lambda_0 + \alpha_1 \lambda_1 + \beta(\lambda_0 E_0 + \lambda_1 E_1).
\end{aligned} \tag{19}$$

We now have all the constraint equations in (16) to (18) and the cost function in (19) to formulate our optimization problem as follows:

$$\begin{aligned}
&\text{minimize} && f(\bar{x}) \\
&\text{subject to} && h(\bar{x}) = 0 \\
&&& g_1(\bar{x}) \leq 0 \\
&&& g_2(\bar{x}) \leq 0.
\end{aligned}$$

This is a standard optimization problem with equality and inequality constraints and can be solved using the Karush-Kuhn-Tucker (KKT) theorem [17]. An exact solution can be obtained by numerically solving the following equation for λ_1 (see Appendix B).

$$\begin{aligned}
\frac{c_2 - c_0 - c_3}{c_1} + \frac{(\alpha_1 - \alpha_0)}{\beta T c_1} \\
= \frac{e^{-\lambda_1 \pi r^2}}{(1 - e^{-\lambda_1 \pi r^2})} \left(1 + \frac{(a - \lambda_1) \pi r^2}{(1 - e^{-\lambda_1 \pi r^2})} \right).
\end{aligned} \tag{20}$$

In order to obtain a closed form expression for λ_1 , we note that in general $H \gg r$ and $\lambda_0 \gg \lambda_1$ since we expect the number of type 0 nodes in each cluster (λ_0/λ_1) to be much larger than one. With these approximations, we obtain the following closed form expression for the required cluster head intensity (see Appendix B),

$$\lambda_1 = \left(\frac{1}{\pi r^2} \right) \log \left(1 + \frac{\lambda_0 \pi r^2}{c} + \sqrt{\left(1 + \frac{\lambda_0 \pi r^2}{c} \right)^2 - 1} \right), \tag{21}$$

where c is given by,

$$c = \frac{c_2 - c_0 - c_3}{c_1} + \frac{(\alpha_1 - \alpha_0)}{\beta T c_1}. \tag{22}$$

We can make further simplifications to (21) for typical transceiver radio parameters. It has been noted in [7] and [12] that, for very short range communication (up to 10m), l is often comparable to μr^k . For such cases, it is not possible to simplify the expression in (21) any further. However, when $\mu r^k \gg l$, we can further simplify the above expression (see (58) in Appendix B),

$$\lambda_1 = \frac{r^{\frac{k-2}{2}} \sqrt{\lambda_0}}{\sqrt{\frac{\pi}{4} \left(H^k + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right)}}.$$

When the propagation loss index k equals 2 (line of sight communication), we obtain the following simple relationship,

$$\lambda_1 = \frac{\sqrt{\lambda_0}}{\sqrt{\frac{\pi}{4} \left(H^2 + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right)}}. \tag{23}$$

Since H , α_0 , α_1 , μ , β , and T are all constants, we obtain that λ_1 scales approximately as $\sqrt{\lambda_0}$. An exact solution for λ_1 can be obtained by numerically solving (20) (see Appendix B). In [12], the authors find the optimum number of cluster heads (k_{opt}) required when N sensor nodes are uniformly distributed over a unit area and nodes use single hop communication to reach the cluster heads. Cluster heads are rotated randomly and periodically for efficient load balancing. For this scenario, minimizing the total energy spent in the network gives:

$$k_{opt} = \frac{\sqrt{N}}{\sqrt{2\pi}} \sqrt{\frac{\epsilon_1}{\epsilon_2}} \frac{1}{H^2}.$$

The authors assume a propagation loss model of $\epsilon_1 x^2$ for communication between a node and its cluster head, and $\epsilon_2 x^4$ for communication between the cluster heads and the base station. To modify this argument to fit our scenario, the latter model needs to be replaced by $\epsilon_1 x^2$. This is because the communication between the cluster heads and the base station (in our case, the aircraft) is line of sight and not multipath. With this modification to their analysis, the above equation takes the following form.

$$k_{opt} = \frac{\sqrt{N}}{\sqrt{2\pi}} \frac{1}{H}. \tag{24}$$

If we interpret k_{opt} as λ_1 and N as λ_0 , then from (23) and (24), we can clearly see that λ_1 has a similar form as a function of λ_0 in both cases. However, since typically $\alpha_1 \gg \alpha_0$ (hardware complexity of ordinary sensor nodes is much less than the hardware complexity of the cluster heads), the cluster head intensity λ_1 that we obtain will be smaller than k_{opt} .

In summary, so far we have (20) which can be solved numerically to obtain λ_1 . We also have an approximate solution for λ_1 in (23). An approximate solution for λ_0 is a , and is given by (25) (see Appendix B), while the exact solution can be determined as $a - \lambda_1$. Once λ_0 and λ_1 have been determined, we can determine E_1 using (26) and E_0 using (27) (see Appendix B), and the problem is completely solved. Note that, due to our model of lifetime (number of data gathering cycles), both E_0 and E_1 are proportional to T .

$$a = \frac{1}{(\pi\theta_0^2 pr^2)} \log \left(\frac{1}{\epsilon(\gamma_0 r)^2} \right), \quad (25)$$

$$E_1 = c_2 T + c_3 T \frac{\lambda_0}{\lambda_1}, \quad (26)$$

$$E_0 = E_1 \left(\frac{c_0 + c_1 \left(\frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \right)}{c_2 + c_3 \left(\frac{\lambda_0}{\lambda_1} \right)} \right). \quad (27)$$

So far, we have assumed that it is possible to equip type 0 nodes with battery energy E_0 and type 1 nodes with battery energy E_1 . However, after determining E_0 and E_1 , it might not be possible to provide those battery energies to the nodes due to physical and/or manufacturing limitations. In such cases, it is possible to use redundant nodes as proposed in [21] so as to make up for the battery energy shortage. However, for the sake of simplicity, we assume in this paper that it is possible to equip nodes with as much battery energy as required for T data gathering cycles. Analysis with redundant nodes is along similar lines.

Since the intensity of the cluster heads, i.e., λ_1 varies approximately as $\sqrt{\lambda_0}$, for large λ_0 we have $\lambda_0 \gg \lambda_1$. Hence, it does not matter whether the cluster heads do the sensing in addition to fusion since their intensity is very small as compared to λ_0 . Consequently, the cluster heads could serve purely as fusion centers.

4.2 Grid Deployment

Now, we consider a simple grid of nodes in which nodes are placed along grid points with distance r between them. If all the nodes are reliable, this grid will trivially provide connectivity as well as coverage. We would like to compare the results that we get for the random node deployment scenario with those for the grid deployment scenario. We use the same notation as in the previous subsection, but λ_0 and λ_1 are to be interpreted, respectively, as the number of type 0 nodes and the number of type 1 nodes placed along the grid.

The cost function is the same as before. However, the connectivity and coverage condition takes the following form:

$$a' = (\lambda_0 + \lambda_1) = \frac{1}{r^2}, \quad (28)$$

where a' is the total number of nodes required and is analogous to a from (46).

The only expression that is different from the random network is P_0^r which we need to compute. P_0^r can be evaluated easily by noting that, in a grid, there can be only

four single hop neighbors for each cluster head which share the burden of last hop communication, i.e., there are exactly four critical nodes. Hence,

$$\begin{aligned} P_0^r &= (E_0^r + E_0^t) \frac{(\lambda_0/\lambda_1 - 4)}{4} \\ &= c_1 \left(\frac{a'}{4\lambda_1} - \frac{5}{4} \right). \end{aligned}$$

We can again prove that the original problem in the KKT theorem with inequality constraints reduces to a simple minimization problem with equality constraints as in the random deployment case. The proof is along similar lines as before and is therefore skipped. We finally obtain:

$$\begin{aligned} f(\lambda_1) &= a'(\alpha_0 + \beta T(c_0 + c_3)) \\ &\quad + (\beta T(c_2 - c_0 - c_3) + (\alpha_1 - \alpha_0))\lambda_1 \\ &\quad + \beta T c_1 (a' - \lambda_1) \left(\frac{a'}{4\lambda_1} - \frac{5}{4} \right). \end{aligned} \quad (29)$$

For local minimum,

$$\frac{d}{d\lambda_1} (f(\lambda_1)) = 0 \Rightarrow \lambda_1 = \frac{a'}{\sqrt{4c + 5}}. \quad (30)$$

Note that, even in this case, c is given by (22). As the energy spent in countering the propagation loss to communicate with the aircraft (μH^k) is much larger than the other energy terms (μr^k , l , E_f), c_2 dominates over other c_i (see (14) and (15)). Hence, we obtain that $c \gg 1$ and (30) takes the form,

$$\lambda_1 \approx \frac{a'}{2\sqrt{c}}.$$

We assume $\lambda_0 \gg \lambda_1$ and $\mu r^k \gg l$, as we did in obtaining (23). Further, if we assume a propagation constant of $k = 2$, then substituting for c from (22), and after some simplifications we obtain,

$$\lambda_1 = \frac{\sqrt{\lambda_0}}{\sqrt{2 \left(H^2 + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right)}}. \quad (31)$$

A comparison of (23) and (31) reveals a striking similarity in the overall form of λ_1 for both random deployment as well as grid deployment. In both cases, λ_1 scales as the square root of λ_0 , although one model has random deployment while the other has deterministic deployment. Note that λ_0 (and, hence, λ_1) is much larger in case of a random network (46) as compared to λ_0 for a grid (28). This is due to the randomness involved in deployment. However, a reliable grid is a highly idealized scenario. Usually, nodes are unreliable and from the results of Shakkottai et al. in [13], it is clear that if the nodes are unreliable with a certain probability $(1 - p)$, then even for a grid deployment the number of nodes n required for coverage and connectivity is such that n^2 scales as $\frac{\log n}{p}$. If the nodes are unreliable and we use the random deployment model, all our results are still valid with minor modifications because even in this case, the coverage-connectivity constraint retains the same $\log \lambda / \lambda$ form (lemma in Appendix A). This indicates that whenever there is randomness (either due to node failure or due to deployment, or both), we get the $\log \lambda / \lambda$ form for

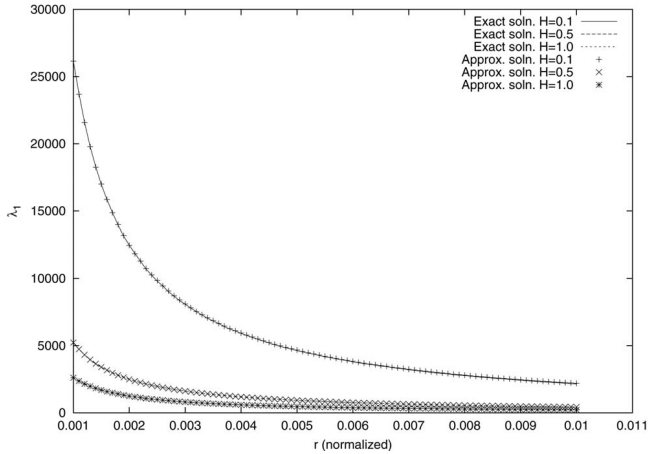


Fig. 2. Approximate and exact solutions for λ_1 in random deployment overlap.

coverage-connectivity constraint). Hence, our work can be easily generalized to the case of unreliable nodes.

5 NUMERICAL RESULTS

In this section, we provide justifications for the approximations that we made in obtaining (23) by using some typical transceiver radio parameters. Consider an area of $A = 10\text{km} \times 10\text{km}$ to be covered by sensor nodes. If the total area A is normalized to 1, all the distances should be divided by a corresponding normalizing factor \sqrt{A} . Hence, for the above settings, all the distances should be divided by 10km. The sensing radius r of the nodes is varied from 10m to 100m (normalized r from 0.001 to 0.01) and the distance of the nodes from the aircraft H is varied from 1km to 10km (normalized H varying from 0.1 to 1). These are typical values for practical surveillance networks. For simplicity, we assume $\alpha_1 = \alpha_0$, $k = 2$, and $E_f = l = 0$. Note that, if $\alpha_1 > \alpha_0$, our main approximation $c \gg 1$ for c given by (54) works even better. We solve (20) numerically to find an exact solution and compare the result with the solution obtained through approximations in (23). From Fig. 2, it is clear that the approximation works quite well for the settings of practical interest since the curves corresponding to the exact and the approximate solutions overlap. Note that λ_0 is determined by the connectivity-coverage constraint and, hence, does not depend on H . As H decreases, the communication between the cluster heads and the aircraft becomes less expensive. Hence, it is worthwhile having more type 1 nodes since that leads to a reduction in the average cluster size which in turn lowers the relaying burden on the critical nodes. Hence, we see from Fig. 2 that, for smaller values of H , λ_1 is higher. Also, from Fig. 3, it can be seen that, for the above settings, $\lambda_0 \gg \lambda_1$ which validates our earlier approximations.

6 CONCLUSIONS

We consider two types of hierarchical sensor networks: one that uses random uniform deployment and the other that uses grid deployment. Our approach involves using two types of nodes: type 0 nodes which do the sensing and

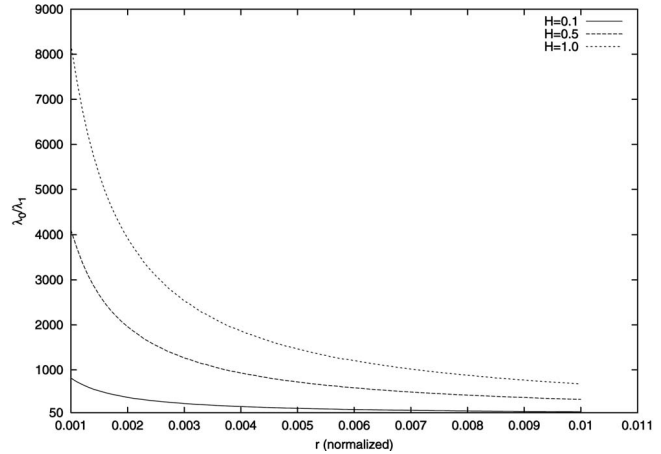


Fig. 3. λ_0/λ_1 computed from the exact solution.

relaying on the ground, and type 1 nodes (cluster heads) which do data aggregation and communication with the aircraft. Type 0 nodes have energy level E_0 and are deployed with intensity λ_0 , while type 1 nodes have energy level E_1 and are deployed with intensity λ_1 . We provide results that guarantee a minimum lifetime (i.e., at least T successful data gathering trips) of the sensor network. We also ensure conditions for connectivity and coverage of the area during the lifetime of the network. We take into account constraints for a sharp cutoff effect so that when the network becomes unusable due to the loss of connectivity and/or coverage, very little residual energy is left behind in the nodes. For this, we note that the cluster heads as well as the nodes within one hop of the cluster heads, i.e., the critical nodes have the maximum relaying burden and, therefore, these nodes are likely to run out of battery before other nodes. We minimize the overall network cost while satisfying the above constraints. We compare the results for the random deployment with those of the grid deployment. We observe that in both deployment scenarios, the required cluster head intensity λ_1 scales as $\sqrt{\lambda_0}$. Our analysis can be easily extended to scenarios in which unreliable nodes are deployed randomly or along grid points.

APPENDIX A

Lemma: Connectivity and coverage for Poisson deployment. *If unreliable nodes are distributed in a plane with a homogeneous two dimensional Poisson point process with intensity λ , then the probability of connectivity and coverage is*

$$\begin{aligned} & P(\text{network is connected and region is covered}) \\ & \geq 1 - \left(\frac{1}{\gamma r(\lambda)} \right)^2 e^{-\lambda p(\lambda) \pi \theta^2 r^2(\lambda)} \\ & \forall \gamma, \theta > 0 \text{ such that } \gamma + 2\theta = 1, \end{aligned}$$

where $r(\lambda)$ is the communication (or the sensing) radius of the nodes and $p(\lambda)$ is the probability of a reliable node.

Proof. We use a similar bin-packing argument as used in [13]. In [13], the authors cover a square of unit area with circles of radii $\theta r(\lambda)$ which are shifted by $\gamma r(\lambda)$ where $\gamma + 2\theta = 1$. Then, a sufficient condition for connectivity and coverage is that there should be at least one active

node in each circle. Using the same notation as used in [13], let:

$$P_s(\lambda) = \Pr\{\text{at least one active node in each circle}\}$$

$$A_i = \{\text{at least one active node in circle } i\},$$

$$P_s(\lambda) = \Pr\left(\bigcap_i A_i\right) = 1 - \Pr\left(\bigcup_i \bar{A}_i\right)$$

$$\geq 1 - \sum_i \Pr(\bar{A}_i) = 1 - \left(\frac{1}{\gamma r(\lambda)}\right)^2 \Pr(\bar{A}_1).$$

The probability of \bar{A}_1 , i.e., the probability that there are no active nodes in a circle of area $x = \pi\theta^2 r^2(\lambda)$ is the probability that k nodes fall in the area and all k of them fail. Assuming that nodes fail independently of each other,

$$\Pr(\bar{A}_1) = \sum_{k=0}^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} (1 - p(\lambda))^k$$

$$= e^{-\lambda p(\lambda)x} = e^{-\lambda p(\lambda)\pi\theta^2 r^2(\lambda)}.$$

Hence,

$$P_s(\lambda) \geq 1 - \left(\frac{1}{\gamma r(\lambda)}\right)^2 e^{-\lambda p(\lambda)\pi\theta^2 r^2(\lambda)}; \quad \gamma + 2\theta = 1,$$

which is similar in form to the expression that the authors in [13] get in terms of n , when n nodes are deployed along grid points,

$$P_s(n) \geq 1 - \left(\frac{1}{\gamma r(n)}\right)^2 e^{-np(n)\pi\theta^2 r(n)^2}; \quad \gamma + 2\theta = 1.$$

It can also be shown that even the asymptotic results for coverage and connectivity follow exactly along the same lines as [13]. \square

APPENDIX B

In this section, we solve the following minimization problem where the corresponding functions are given by (16) to (19).

$$\begin{aligned} &\text{minimize } f(\bar{x}) \\ &\text{subject to } h(\bar{x}) = 0 \\ &\quad g_1(\bar{x}) \leq 0 \\ &\quad g_2(\bar{x}) \leq 0. \end{aligned}$$

We find the solution to the above problem by solving:

$$\nabla f(\bar{x}) + \mu_0 \nabla h(\bar{x}) + \mu_1 \nabla g_1(\bar{x}) + \mu_2 \nabla g_2(\bar{x}) = \mathbf{0}, \quad (32)$$

$$\text{with } \mu_1 \geq 0, \mu_2 \geq 0, \mu_1 g_1(\bar{x}) + \mu_2 g_2(\bar{x}) = 0, \quad (33)$$

where μ_0, μ_1 , and μ_2 are the constants of the KKT theorem. From (19):

$$\nabla f(\bar{x}) = \begin{bmatrix} \alpha_0 + \beta E_0 \\ \alpha_1 + \beta E_1 \\ \beta \lambda_0 \\ \beta \lambda_1 \end{bmatrix}. \quad (34)$$

From (16),

$$\nabla h(\bar{x}) = \begin{bmatrix} \frac{-c_3 E_0}{\lambda_1} \\ \frac{-c_1 E_1 \pi r^2 e^{-\lambda_1 \pi r^2}}{(1 - e^{-\lambda_1 \pi r^2})^2} + \frac{c_3 E_0 \lambda_0}{\lambda_1^2} \\ -c_2 - \frac{c_3 \lambda_0}{\lambda_1} \\ c_0 + c_1 \frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \end{bmatrix}. \quad (35)$$

From (18):

$$\nabla g_1(\bar{x}) = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}. \quad (36)$$

From (17):

$$\nabla g_2(\bar{x}) = \begin{bmatrix} \frac{c_3 T}{\lambda_1} \\ -\frac{c_3 T \lambda_0}{\lambda_1^2} \\ 0 \\ -1 \end{bmatrix}. \quad (37)$$

Substituting (34) to (37) in (32) we have the following four equations:

$$0 = (\alpha_0 + \beta E_0) + \mu_0 \left(\frac{-c_3 E_0}{\lambda_1}\right) + \mu_1(-1)$$

$$+ \mu_2 \left(\frac{c_3 T}{\lambda_1}\right), \quad (38)$$

$$0 = (\alpha_1 + \beta E_1) + \mu_0 \left(\frac{-c_1 E_1 \pi r^2 e^{-\lambda_1 \pi r^2}}{(1 - e^{-\lambda_1 \pi r^2})^2} + \frac{c_3 E_0 \lambda_0}{\lambda_1^2}\right)$$

$$+ \mu_1(-1) + \mu_2 \left(-\frac{c_3 T \lambda_0}{\lambda_1^2}\right), \quad (39)$$

$$0 = (\beta \lambda_0) + \mu_0 \left(-c_2 - \frac{c_3 \lambda_0}{\lambda_1}\right) + \mu_1(0) + \mu_2(0), \quad (40)$$

$$0 = (\beta \lambda_1) + \mu_0 \left(c_0 + c_1 \frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}}\right) + \mu_1(0) + \mu_2(-1). \quad (41)$$

Assuming that a feasible solution exists, i.e., $\lambda_0, \lambda_1, E_0, E_1 > 0$, from (40), we get

$$\mu_0 = \frac{\beta \lambda_0 \lambda_1}{c_2 \lambda_1 + c_3 \lambda_0} > 0. \quad (42)$$

Hence, from (41), we get

$$\mu_2 = (\beta \lambda_1) + \mu_0 \left(c_0 + c_1 \frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}}\right) > 0. \quad (43)$$

But, from the KKT theorem, this implies that the corresponding inequality (18) becomes an equality, i.e.,

$$c_2 T + c_3 T \left(\frac{\lambda_0}{\lambda_1}\right) - E_1 = g_2(\bar{x}) = 0. \quad (44)$$

Using the value of μ_0 from (42) in (38),

$$\mu_1 = \alpha_0 + E_0 \left(\frac{\beta c_2 \lambda_1}{c_2 \lambda_1 + c_3 \lambda_0}\right) + \mu_2 \left(\frac{c_3 T}{\lambda_1}\right) > 0.$$

Again using the KKT theorem, this implies that the inequality of (17) becomes an equality, i.e.,

$$a - (\lambda_0 + \lambda_1) = g_1(\bar{x}) = 0. \quad (45)$$

So, now we can reformulate our problem as follows:

$$\begin{aligned} & \text{minimize} && f(\bar{x}) \\ & \text{subject to} && h(\bar{x}) = 0 \\ & && g_1(\bar{x}) = 0 \\ & && g_2(\bar{x}) = 0. \end{aligned}$$

We have three equations and four unknowns. So, we eliminate the three variables λ_0 , E_1 , and E_0 from the cost function. Thus, we obtain $f(\bar{x})$ as a function of a single variable λ_1 .

From (4) and (6)

$$a = \frac{1}{(\pi\theta_0^2 pr^2)} \log \left(\frac{1}{\epsilon(\gamma_0 r)^2} \right). \quad (46)$$

Since $g_1(\bar{x}) = 0$ using (45),

$$\lambda_0 = a - \lambda_1. \quad (47)$$

Since $g_2(\bar{x}) = 0$ using (44),

$$E_1 = c_2 T + c_3 T \frac{\lambda_0}{\lambda_1}. \quad (48)$$

Since $h(\bar{x}) = 0$ using (16),

$$E_0 = E_1 \left(\frac{c_0 + c_1 \left(\frac{e^{-\lambda_1 \pi r^2}}{1 - e^{-\lambda_1 \pi r^2}} \right)}{c_2 + c_3 \left(\frac{\lambda_0}{\lambda_1} \right)} \right), \quad (49)$$

$$\begin{aligned} f(\lambda_1) = & a(\alpha_0 + \beta T(c_0 + c_3)) \\ & + (\beta T(c_2 - c_0 - c_3) + (\alpha_1 - \alpha_0))\lambda_1 \\ & + \beta T c_1 (a - \lambda_1) \frac{e^{-\lambda_1 \pi r^2}}{(1 - e^{-\lambda_1 \pi r^2})}. \end{aligned} \quad (50)$$

The local extremum (minimum or maximum) of $f(\cdot)$ is attained when $df/d\lambda_1 = 0$.

$$\begin{aligned} & \frac{c_2 - c_0 - c_3}{c_1} + \frac{(\alpha_1 - \alpha_0)}{\beta T c_1} \\ & = \frac{e^{-\lambda_1 \pi r^2}}{(1 - e^{-\lambda_1 \pi r^2})} \left(1 + \frac{(a - \lambda_1) \pi r^2}{(1 - e^{-\lambda_1 \pi r^2})} \right). \end{aligned} \quad (51)$$

An exact solution for λ_1 can be obtained by solving the above equation numerically. To confirm that the solution of (51) is indeed a local minimum of (50), we verify that:

$$\begin{aligned} \frac{d^2}{d\lambda_1^2}(f(\lambda_1)) = & \frac{2\beta T c_1 \pi r^2 e^{-\lambda_1 \pi r^2}}{(1 - e^{-\lambda_1 \pi r^2})^2} + \\ & \frac{\beta T c_1 (a - \lambda_1) (\pi r^2)^2 e^{-\lambda_1 \pi r^2} (1 + e^{-\lambda_1 \pi r^2})}{(1 - e^{-\lambda_1 \pi r^2})^3} \\ & > 0 \text{ for } \lambda_1 \in (0, a]. \end{aligned}$$

The above also implies that $f(\lambda_1)$ is convex on $(0, a]$ and hence the local minimum is also the global minimum. Note that λ_1 is constrained to $(0, a]$ because of (46).

However, for typical practical scenarios, it is possible to make certain useful approximations and simplify this equation. Note that, from (46), $a = \lambda_0 + \lambda_1$. But, in typical practical scenarios, we would expect the number of type 0 nodes in each cluster (λ_0/λ_1) to be much larger than 1. In other words, $\lambda_0 \gg \lambda_1$ and, hence,

$$\begin{aligned} a - \lambda_1 & \approx a \\ \lambda_0 & \approx a. \end{aligned} \quad (52)$$

With this approximation (51) can be simplified to a quadratic in $t = e^{-\lambda_1 \pi r^2}$.

$$\begin{aligned} \frac{c_2 - c_0 - c_3}{c_1} + \frac{(\alpha_1 - \alpha_0)}{\beta T c_1} = c = \frac{t}{(1-t)} \left(1 + \frac{a\pi r^2}{(1-t)} \right) \\ \Rightarrow (c+1)t^2 - (2c+1+a\pi r^2)t + c = 0, \end{aligned} \quad (53)$$

where c is given by:

$$c = \frac{c_2 - c_0 - c_3}{c_1} + \frac{(\alpha_1 - \alpha_0)}{\beta T c_1}.$$

Using (9),

$$c = \frac{\mu H^k - E_f - l - \mu r^k}{2l + \mu r^k} + \frac{(\alpha_1 - \alpha_0)}{\beta T (l + \mu r^k + l)}.$$

Since the distance of the nodes from the aircraft H is much larger than the distance r , we ignore the r^k terms in the numerator. Over large distances, the propagation loss term also dominates over l . We also use the fact that the energy spent on long range transmissions (E_l^t) is much higher than the energy spent on computations (E_f) as noted in [15]. Hence, we obtain,

$$\begin{aligned} c & \approx \frac{H^k}{2l/\mu + r^k} + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu (2l/\mu + r^k)} \\ & = \frac{1}{2l/\mu + r^k} \left(H^k + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right). \end{aligned} \quad (54)$$

Finally, we note that since $\mu H^k \gg l + \mu r^k$, we obtain $c \gg 1$. Using these facts, (53) becomes

$$t^2 - \left(2 + \frac{a\pi r^2}{c} \right) t + 1 = 0. \quad (55)$$

The only feasible ($t < 1$) solution to the above equation is:

$$t = \left[\left(1 + \frac{a\pi r^2}{c} \right) + \sqrt{\left(1 + \frac{a\pi r^2}{c} \right)^2 - 1} \right]^{-1}$$

Since $t = e^{-\lambda_1 \pi r^2}$,

$$\lambda_1 = \left(\frac{1}{\pi r^2} \right) \log \left(1 + \frac{a\pi r^2}{c} + \sqrt{\left(1 + \frac{a\pi r^2}{c} \right)^2 - 1} \right) \quad (56)$$

$$= \left(\frac{1}{\pi r^2} \right) \log \left(1 + \frac{\lambda_0 \pi r^2}{c} + \sqrt{\left(1 + \frac{\lambda_0 \pi r^2}{c} \right)^2 - 1} \right). \quad (57)$$

Furthermore, if $\mu r^k \gg l$, i.e., if the energy spent in the RF power amplifiers dominates over the energy spent in the transmitter electronics circuitry over a range of r using (54),

$$c \approx \frac{1}{r^k} \left(H^k + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right) > \frac{H^k}{r^k}.$$

Also, since a satisfies (46), we eliminate a to obtain

$$\left(\frac{a\pi r^2}{c} \right) < \left(\frac{1}{H^k \theta_0^2 p} \right) r^k \log \left(\frac{1}{(\epsilon \gamma_0^2) r^2} \right).$$

As already noted, we are interested in the scenarios for which $r \ll 1$. Note that

$$\lim_{r \rightarrow 0} r^k \log \frac{1}{r^2} = 0.$$

Hence, for sufficiently small r

$$\left(\frac{a\pi r^2}{c} \right) \ll 1.$$

Note that H is fixed and we are looking at the scenario in which the nodes have very small sensing/coverage radius r . Then, a first order approximation for λ_1 in (56) can be obtained as follows:

$$\begin{aligned} \lambda_1 &\approx \left(\frac{1}{\pi r^2} \right) \log \left(\left(1 + \frac{a\pi r^2}{c} \right) + \sqrt{\frac{2a\pi r^2}{c}} \right) \\ &\approx \left(\frac{1}{\pi r^2} \right) \log \left(1 + \sqrt{\frac{2a\pi r^2}{c}} \right) \\ &\approx \left(\frac{1}{\pi r^2} \right) \sqrt{\frac{2a\pi r^2}{c}} \\ &= \frac{r^{\frac{k-2}{2}} \sqrt{a}}{\sqrt{\frac{\pi}{4} \left(H^k + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right)}} \\ &= \frac{r^{\frac{k-2}{2}} \sqrt{\lambda_0}}{\sqrt{\frac{\pi}{4} \left(H^k + \frac{(\alpha_1 - \alpha_0)}{\beta T \mu} \right)}}. \end{aligned} \quad (58)$$

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions that greatly helped us improve the final quality of this paper. This work was supported in part by a DARPA grant (contract no. MDA 972-02-1-0032), a US National Science Foundation grant (contract no. 0087266) and a grant from Purdue Research Foundation. Some of these results were presented at the First European Workshop on Wireless Sensor Networks, Berlin, Germany, January 19-21 2004, see [1].

REFERENCES

[1] V. Mhatre, C. Rosenberg, D. Kofman, R. Mazumdar, and N. Shroff, "Design of Surveillance Sensor Grids with a Lifetime Constraint," *Proc. First European Workshop Wireless Sensor Networks (EWSN)*, Jan. 2004.
 [2] I.F. Akyildiz, W. Su, Y. Sankarsubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer Networks*, vol. 38, pp. 393-422, Mar. 2002.

[3] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, "Impact of Network Density on Data Aggregation in Wireless Sensor Networks," *Proc. Int'l Conf. Distributed Computing Systems (ICDCS)*, July 2002.
 [4] B. Krishnamachari, D. Estrin, and S. Wicker, "The Impact of Data Aggregation in Wireless Sensor Networks," *Proc. Int'l Workshop Distributed Event Based Systems (DEBS)*, July 2002.
 [5] A.D. Amis, R. Prakash, T.H.P. Vuong, and D.T. Huynh, "Max-Min D-Cluster Formation in Wireless Ad Hoc Networks," *Proc. IEEE INFOCOM Conf.*, Mar. 2000.
 [6] C.R. Lin and M. Gerla, "Adaptive Clustering for Mobile Wireless Networks," *J. Selected Areas in Comm.*, pp. 1265-1275, Sept. 1997.
 [7] M. Bhardwaj, T. Garnett, and A.P. Chandrakasan, "Upper Bounds on Lifetime of Sensor Networks," *Proc. IEEE Int'l Conf. Comm. (ICC '01)*, June 2001.
 [8] C.F. Chiasserini, I. Chlamtac, P. Monti, and A. Nucci, "Energy Efficient Design of Wireless Ad Hoc Networks," *Proc. Networking 2002 Conf.*, E. Gregori, M. Conti, A.T. Campbell, G. Omidyar, M. Zukerman, eds., May 2002.
 [9] A. Amis and R. Prakash, "Load-Balancing Clusters in Wireless Ad Hoc Networks," *Proc. ASSET 2000 Conf.*, Mar. 2000.
 [10] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Maximum Lifetime Data Gathering and Aggregation in Wireless Sensor Networks," *Proc. IEEE Int'l Conf. Networking (NETWORKS '02)*, pp. 685-696, Aug. 2002.
 [11] S. Bandyopadhyay and E. Coyle, "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks," *Proc. IEEE INFOCOM'03 Conf.*, Apr. 2003.
 [12] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Trans. Wireless Comm.*, vol. 1, no. 4, Oct. 2002.
 [13] S. Shakkottai, R. Srikant, and N. Shroff, "Unreliable Sensor Grids: Coverage, Connectivity and Diameter," *Proc. IEEE INFOCOM'03 Conf.*, Apr. 2003.
 [14] S.G. Foss and S.A. Zuyev, "On a Voronoi Aggregative Process Related to a Bivariate Poisson Process," *Advances in Applied Probability*, vol. 28, no. 4, pp. 965-981, 1996.
 [15] G.J. Pottie and W.J. Kaiser, "Wireless Integrated Network Sensors," *Comm. ACM*, vol. 43, no. 5, pp. 51-58, 2000.
 [16] F. Baccelli and P. Bremaud, *Elements of Queueing Theory*. Springer Verlag, 1994.
 [17] E. Chong and S. Zak, *An Introduction to Optimization*, second ed. Wiley and Sons, Inc., 2001.
 [18] G. Gupta and M. Younis, "Fault-Tolerant Clustering of Wireless Sensor Networks," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC 2003)*, Mar. 2003.
 [19] G. Zussman and A. Segall, "Energy Efficient Routing in Ad Hoc Disaster Recovery Networks," *Proc. IEEE INFOCOM 2003 Conf.*, Apr. 2003.
 [20] M. Bhardwaj and A. Chandrakasan, "Bounding the Lifetime of Sensor Networks Via Optimal Role Assignments," *Proc. IEEE INFOCOM 2002 Conf.*, June 2002.
 [21] A. Cerpa and D. Estrin, "ASCENT: Adaptive Self-Configuring Sensor Networks Topologies," *Proc. IEEE INFOCOM 2002 Conf.*, June 2002.
 [22] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. Srivastava, "Coverage Problems in Wireless Ad-Hoc Sensor Networks," *Proc. IEEE INFOCOM 2001 Conf.*, vol. 3, pp. 1380-1387, Apr. 2001.



Vivek P. Mhatre graduated with the BTech degree in electrical engineering from the Indian Institute of Technology (IIT) Bombay, India in August 2000. He is currently working towards the PhD degree at the School of Electrical and Computer Engineering at Purdue University, Indiana. His research interests include wireless sensor networks and ad-hoc networks. He is a student member of the IEEE.



Ravi Mazumdar received the Btech degree in electrical engineering from the Indian Institute of Technology, Bombay, India in 1977, the MSc degree DIC in control systems from Imperial College, London, UK in 1978, and the PhD degree in systems science from the University of California, Los Angeles, USA in 1983. He is a professor of electrical and computer engineering at Purdue University. Prior to joining Purdue, he was a professor of mathematics at the University of Essex, Colchester, UK from 1996-1999. From 1988-1996, he was a professor at INRS-Télécommunications—a graduate research institute of the Université du Québec and an invited professor at McGill University, Montreal, Canada. From 1985-1988, he was an assistant professor of electrical engineering at Columbia University, New York. He has held visiting positions at UCLA, University of Twente (The Netherlands), Indian Institute of Science, and the Ecole Nationale Supérieure des Télécommunications (Paris). He is a Fellow of the Royal Statistical Society. His research interests are in game theory, applied probability, and stochastic analysis focusing on applications in wireless communication networks, statistical signal processing, and mathematical finance. He is a senior member of the IEEE.



Catherine Rosenberg was educated at Ecole Nationale Supérieure des Télécommunications de Bretagne, Diplôme d'Ingénieur in electrical engineering in 1983 and received Doctorat en Sciences in computer science at the University of Paris, Orsay, in 1986 and the MS degree in computer science from the University of California at Los Angeles in 1984. She has worked in several countries including USA, UK, Canada, France, and India. In particular, she worked for Nortel Networks in the UK, AT&T Bell Laboratories in the USA, Alcatel in France, and taught at Ecole Polytechnique of Montreal (Canada). She is currently a professor in the School of Electrical and Computer Engineering at Purdue University, Indiana. She is also the director of the university-wide Center for Wireless Systems and Applications at Purdue University. Her research has been supported by agencies and industries including NSERC (The Canadian National Science Foundation), FCAR (The Quebec counterpart of NSERC), CRC (Canadian Ministry of Communications), EEC (European Commission), ESA (European Space Agency), France-Telecom, CISCO, and Nortel Networks. Her research interests are in broadband networks (IP and ATM), wireless networking, broadband satellite networks, network security, peer-to-peer networks, and traffic engineering (QoS, Charging, Dynamic Provisioning, Network Design, and Routing). She has authored more than 70 papers on ATM, satellite broadband networking, wireless networking, and traffic engineering and has filed several patents in the UK and the USA. She is a senior member of the IEEE.



Ness Shroff received the PhD degree from Columbia University, New York, in 1994. He is currently an associate professor in the School of Electrical and Computer Engineering at Purdue University, Indiana. His research interests span the areas of wireless and wireline communication networks. He is especially interested in fundamental problems in the design, performance, scheduling, capacity, pricing, and control of these networks. His research is funded by various companies such as Intel, Hewlett Packard, Nortel, AT&T, BAE systems, and L. G. Electronics; and government agencies such as the US National Science Foundation (NSF), Defense Advanced Research Projects Agency (DARPA), Indiana Department of Transportation, and the Indiana 21st Century fund. Dr. Shroff is an editor for *IEEE/ACM Transactions on Networking* and the *Computer Networks Journal*, and past editor of *IEEE Communications Letters*. He was the technical program cochair for IEEE INFOCOM'03 conference, the panel cochair for ACM Mobicom'02 conference, the program cochair for the symposium on high-speed networks, Globecom 2001, and the conference chair for the 14th Annual IEEE Computer Communications Workshop in 1999. He received the US NSF CAREER award in 1996. He is a senior member of the IEEE.



Daniel Kofman is an associate professor at the Ecole Nationale Supérieure des Télécommunications GET, Paris, France. He is an expert in new networking technologies and architectures in particular in the evolution towards high-speed multiservice-convergent networks and in the optimization of IP networks. He directs research of a number of research projects, in collaboration with main industrial and academic organizations. He is the scientific and technical coordinator of a network of excellence granted by the European Commission on the "Design and Engineering of the Next Generation Internet" that assembles the most prominent European experts of 58 research laboratories. He is member of the scientific Committee of the French Parliament and an expert for the CSTI (Conseil Stratégique des Technologies de l'Information), a committee reporting directly to the French Prime Minister, for which he issued a report on the "Synthesis of Telecommunication Network Evolution." He is a member of the French Commission 3 of the RNRT on "Architecture des réseaux et systèmes de télécommunication." This commission is in charge of overseeing the French research projects on networking issues. He has also been invited as an expert by the European Commission in the context of the FP6. His teaching at the ENST is on "IP Networking" and "New Networking Technologies." he has published two books and a number of papers on related topics. He is a reviewer for numerous scientific revues. He is also a member of the scientific or program committees for major international scientific conferences in the telecommunication domain, such as the IEEE Infocom, ITC, IFIP, and others. He is one of the founders of a start-up (Net-Tone) that has developed a solution to help the telecommunication operators and corporations to migrate towards convergent service architectures. He is a member of the IEEE.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.