Initializing Sensor Networks of Non-uniform Density in the Weak Sensor Model

Martín Farach-Colton and Miguel A. Mosteiro

Department of Computer Science, Rutgers University, Piscataway, NJ 08854, USA
{farach,mosteiro}@cs.rutgers.edu

Abstract. Assumptions about node density in the Sensor Networks literature are frequently too strong or too weak. Neither absolutely arbitrary nor uniform deployment seem feasible in most of the intended applications of sensor nodes. We present a Weak Sensor Model-compatible distributed protocol for hop-optimal network initialization, under the assumption that the maximum density of nodes is some value Δ known by all of the nodes. In order to prove lower bounds, we observe that all nodes must communicate with some other node in order to join the network, and we call the problem of achieving such a communication the *Group Therapy Problem*. We show lower bounds for the Group Therapy Problem in Radio Networks of maximum density Δ, regardless of the use of randomization, and a stronger lower bound for the important class of randomized fair protocols. We also show that even when nodes are distributed uniformly, the same lower bound holds, even in expectation and even for the simpler problem of Clear Transmission.

1 Introduction

Although some papers analyze problems in Radio Networks under the assumption of an arbitrary distribution of nodes, in most applications the layout of nodes is not the result of an uncontrolled random experiment in which the probability of some highly undesirable outcome is positive. On the other hand, a uniform distribution of nodes in the plane, as is customary to assume in the Sensor Networks literature, may be difficult or impossible to achieve in settings where the environment is hostile or remote. Furthermore, for any reasonable model of non-uniform distribution of nodes chosen, a minimum density of nodes has to be ensured in order to guarantee connectivity, and a non-trivial maximum density of nodes can indeed be guaranteed. An example of a feasible model for the distribution of nodes that reflects the random nature of the deployment, yet excludes highly unlikely pathological cases is a multiple bivariate normal distribution.

In this paper, we do not limit ourselves to any particular distribution but we define bounds on the density of nodes for any reasonable model. More specifically, we define a distribution of nodes as a *Smooth Distribution* if the maximum density of nodes in any one-hop neighborhood is some value $\Delta \leq n$, and for any constant $\alpha > 0$, in any disc of radius αr there exists a constant $\beta > 0$ such that the number of nodes is at least $\beta \log n$[1]. The rationale behind the choice of the lower bound is that, when the nodes are deployed uniformly at random, with enough density to achieve connectivity, a logarithmic density is guaranteed w.h.p.[2] [5]. Given that deterministic deployment is not possible in hostile or remote environments, we assume the random deployment to be a process with the goal of achieving uniform distribution but where nodes are dropped in excess in some areas.

The Problem. We study the network initialization problem with smooth distribution of nodes under the restrictions of the Weak Sensor Model [5], a harsh and comprehensive model that summarizes the literature on sensor node restrictions. The initialization of a Sensor Network is the problem of self-organizing from scratch as a radio-communication network called a sensor network. Even though communication among sensor nodes is through radio broadcast, it is useful to set up explicit links between nodes in order to establish routing paths and prevent flooding. In this model, collision detection is not available and a transmission in a given time slot is successful if only if exactly one node transmits in the one-hop neigborhood of the receiver. Therefore, achieving non-colliding transmissions fast without knowledge of the topology is not trivial. The problem of achieving non-colliding transmissions has been well studied within other problems. For some problems in multihop networks, such as Sensor Network initialization, the messages transmitted not necessarily must reach all nodes in the network. If in a given time slot a node transmits and all nodes in a two-hop neighborhood do not transmit, we say that a *Clear Transmission* has been produced and all nodes in the one-hop neighborhood of the transmitter have received the Clear Transmission. If all nodes in the network have to either produce or receive a Clear Transmission the problem is called the *Clear Transmission Problem* [6]. For settings where all nodes in the network have to receive all the messages to be transmitted, the various problems studied differ in the number of nodes that have messages. When some arbitrary number k of nodes have a message the problem is known as *k-selection* [10]. If $k = 1$ the problem is called *Broadcast* [2,12], and if $k = n$, it is called *Gossiping* [4,13]. For the purpose of proving lower bounds for Sensor Network initialization, we take as a lower bound the problem of transmitting so that at least one transmission of each node has to be received by some other node. Given that to solve the problem all the participants have to be *heard*, we term this problem the *Group Therapy Problem*.[3]

Previous work. The literature on Sensor Networks is vast and includes both theoretical and empirical research work. Many of the solutions proposed do not sufficiently handle all the aspects of the problem. The protocol in [21] builds a flat topology, but it is assumed that there are enough channels to accomodate

---

[1] Througout this paper, log means $\log_2$ unless otherwise stated.

[2] Define *with high probability* to mean with probability at least $1 - O(n^{-\Omega(1)})$.

[3] In fact, were it not for the extensive literature on gossiping, we would reverse these terms. After all, it is hardly the point of gossiping to tell everyone your news. Conversely, in group therapy one expects to be heard by all. Nonetheless, we use the current notation for consistency.

each link among neighbors which is not possible under the Weak Sensor Model. The protocol in [3] builds a network where every node has at most k neighboring nodes. However, the number of available radii of transmission is a function of n, and the protocol relies in distance estimation hardware. The protocol in [22] is an energy efficient topology control scheme. Unfortunately, global synchronization is necessary and an underlying contention resolution mechanism is assumed. In all these protocols the memory size is assumed to be in $\omega(1)$. Recently, a $O(\log^2 n)$ protocol that builds a network under the Weak Sensor Model was presented [5]. This protocol builds a hop-optimal network in settings where the nodes are deployed uniformly at random. This was the first protocol for network formation that is implementable in sensor nodes even theoretically. More general information about sensor networks can be obtained from the surveys [1,9,19,20,23].

Regarding lower bounds, Kushilevitz and Mansour [12] proved the first lower bound of $\Omega(\log n)$ on the expectation of the running time of any randomized algorithm for clear transmissions in radio networks. For *uniform* protocols, i.e., protocols where all nodes use the same sequence of probabilities of transmission, a lower bound of $\Omega(\log n \log(1/\epsilon)/(\log\log n + \log\log(1/\epsilon)))$ for achieving a Clear Transmission with probability $1-\epsilon$ in a one-hop globally-synchronized Radio Network was proved in [8]. Recently, this lower bound was improved to $\Omega(\log n \log(1/\epsilon))$ in [6] closing the gap for this problem for uniform protocols. A simple application of this result gives a lower bound of $\Omega(\log \Delta \log(1/\epsilon))$ steps in order to solve the Clear Transmission problem with probability at least $1-\epsilon$ in Sensor Networks where the maximum density is Δ.

Related work. An important building block of the initialization protocol presented in this paper, which dominates the overall running time, is an algorithm that guarantees that in a network of size n, each node produces a Clear Transmission within $O(\Delta \log n)$ steps w.h.p. A $O(h + \log n \log\log n)$ adaptive algorithm to solve the problem of realizing arbitrary *h-relations* w.h.p. was presented in [7]. In an h-relation, each processor is the source as well as the destination of at most h messages. Also, for the k-selection problem, Martel's [14] randomized adaptive protocol operates in expected time $O(k + \log n)$. These algorithms could seemingly be used as our building block, thus speeding up the overall running time. However, they rely on collision detection and global synchronization.

Our results. We use two different problems to prove our lower bounds, both necessary conditions to solve the Sensor Network initialization problem. We observe that in order to initialize the network the Group Therapy problem must be solved. $\Omega(\Delta)$ steps are required to solve this problem regardless of the use of randomization. Combining this observation and our previous results [6] for Clear Transmission, a lower bound of $\Omega(\Delta + \log \Delta \log(1/\epsilon))$ for solving the problem with probability $1-\epsilon$ is obtained for uniform protocols. Restricting ourselves to the important class of *fair* protocols, i.e., protocols where the probability of transmission of every node is the same in the same time step, we show here a lower bound of $\Omega(\Delta(\log(1/\epsilon) + \log \Delta))$ to solve the Group Therapy problem, which we also show to be tight. The analysis for fair protocols is relevant

given that, to the best of our knowledge, asymptotically faster adaptive protocols that work under this model for the Group Therapy problem are not known. Finally, we also show that even if the nodes are distributed uniformly, and thus $\Delta = \Theta(\log n)$, a seemingly simpler problem such as Clear Transmission, takes $\Omega(\log^2 n)$ *expected* time using fair protocols. As for upper bounds, we present a distributed protocol to self-organize a Sensor Network where the maximum density in any one hop neighborhood is an arbitrary value Δ. More specifically, we show that, upon waking up, a node joins the network w.h.p. within $O(\Delta \log n)$ time steps. A bottleneck of this algorithm is a local-spanner construction where *each* node produces a Clear Transmission within $O(\Delta \log n)$ steps w.h.p. using a fair protocol. Given that if *every* node produces a Clear Transmission, the simpler Group Therapy problem is also solved, this protocol matches the lower bound showed here. We leave open the question of how to use adaptive algorithms that work under these conditions to speed up the overall running time.
Roadmap. In Section 2 we define the models used througout the paper and we prove our lower bounds in Section 3. Upper bounds are showed in Sections 4 and 5. We include some acknowledgements in Section 6.

2 The Model

In Sensor Networks nodes are expected to be deployed at random in large quantities over an area of interest. Therefore, an abstraction frequently used in Sensor Networks to model the reachability of nodes is the *Random Geometric Graph Model* (RGG), where the nodes are assumed to be distributed uniformly at random in the plane. In this paper, we relax the assumption about the distribution and we use the *Geometric Graph Model* where the nodes are deployed in $\mathbb{R}^2$, and a pair of nodes is connected if and only if they are at an Euclidean distance of at most a parameter r.

Regarding models of sensor node constraints, Bar-Yehuda, Goldreich and Itai [2] detailed a formal model of a radio network that specifies many of the important restrictions on sensor nodes, including limits on contention resolution, but they make no mention of computational limits such as small memory. After this model was introduced, some papers [11, 18] have added more restrictions, although often such restrictions are implicit in the text or algorithms rather than fully specified. In this paper, we use the Weak Sensor Model [5], a harsh and comprehensive model that summarizes the literature on sensor node restrictions taken the most restrictive choices when possible. In this model, the communication among neighboring nodes is through broadcast on a *shared channel*. A node receives a message if and only if exactly one of its neighbors transmits. There is *no collision detection* mechanism available and the channel is assumed to have only two states: single transmission and silence/collision. Sensors nodes have *non-simultaneous reception and transmission*. Time is assumed to be slotted and all nodes have the same clock frequency, but no global synchronizing mechanism is assumed. Nodes are woken up by an adversary perhaps at different times. Sensor nodes may store only a constant number of $O(\log n)$ bit words.

We assume that sensor nodes can adjust their power of transmission to only a *constant* number of levels. Other restrictions include: limited life cycle, short transmission range, only one shared channel of communication, lack of position information and unreliability.

3 Lower Bounds

To solve any problem in a communication network at least one successful transmission is necessary. However, in multihop networks, one non-colliding transmission is not enough to solve most of the problems. For the Radio Network initialization problem we observe that, in order to join the network, for every node at least one transmission has to be received by at least one neighboring node. More precisely, let V be the set of nodes in the network and $N(v) \subseteq V$ denote the set of nodes adjacent to $v \in V$. Then, for all $v \in V$ there is at least one time slot in which there exists a node $u \in N(v)$ such that exactly one node in $N(u)$ transmits and this node is v. We term this problem the *Group Therapy Problem*. In order to provide stronger lower bounds, we relax the Weak Sensor Model to a minimum set of restrictions, namely, only one channel of communication is available, there is no collision-detection mechanism, a node can not receive and transmit in the same time slot, no global synchronism and adversarial node wake-up schedule. We refer to this model as the Radio Network model. We begin observing that a lower bound for the Group Therapy problem, regardless of the use of randomization, is $\Omega(\Delta)$, a claim that we formalize in the following theorem.

Theorem 1. *In order to solve the Group Therapy problem in a multihop Radio Network where the maximum density in any one-hop neighborhood is Δ, any algorithm requires $\Omega(\Delta)$ time-slots.*

Proof. Exploiting the assumption of an adversarial wake-up schedule, let us assume the existence of an adversary that, at a given time, wakes up only a subset of Δ neighboring nodes, i.e., a set of Δ nodes whose connectivity graph is a clique. We call them *active* nodes. Such a subset of nodes exists since the maximum density is Δ. Upon waking up, the active nodes start the execution of the protocol. All the other nodes remain non-active and do not participate in the protocol. In this setting, in order to solve the problem, every node has to achieve a non-colliding transmission in a different time slot, therefore the claim follows.

Combining Theorem 1 with our lower bound [6] of $\Omega(\log n \log(1/\epsilon))$ time steps to solve the Clear Transmission problem with probability $1-\epsilon$ in a one-hop Radio Network of n nodes using uniform protocols, the following lower bound for the Group Therapy problem in Radio Networks is obtained.

Corollary 1. *In order to solve the Group Therapy problem with probability $1-\epsilon$ in a multihop Radio Network where the maximum density in any one-hop neighborhood is Δ, any randomized uniform protocol requires $\Omega(\Delta + \log \Delta \log(1/\epsilon))$ time-slots.*

We consider now lower bounds for *fair* protocols, i.e., protocols where the probability of transmission of every node in the same time step is the same. The analysis for fair protocols is relevant given that, to the best of our knowledge, asymptotically faster adaptive protocols for the Group Therapy problem that work under this model are not known. We prove this lower bound under the assumption of the existence of a weak adversary that, at a given time, wakes up some subset of neighboring nodes of size $\{2^i | 0 \leq i \leq \log \Delta\}$. We call them *active* nodes. We know that such a subset exists given that there is a clique of size Δ. Upon waking up, the active nodes start the execution of the protocol. All the other nodes remain non-active and do not participate in the protocol.

We define a *randomized fair protocol* to be a sequence $p_1, p_2, \ldots$ where each node transmits with probability p_ℓ in the ℓ^{th} time step after waking up. Given our adversary, this means that all active nodes transmit with the same probability as each other in each time slot. We further assume that all $p_\ell \in \{2^{-j} | 1 \leq j \leq \log \Delta\}$. If this assumption is not true of a particular algorithm A, we can always produce an algorithm A' from A by replacing one attempt in A by a constant number of attempts in A' where the probabilities of transmission in A' have been rounded off to the closest power of 1/2.

Let p_{ij} denote the probability that a given node fails to achieve a non-colliding transmission when 2^i active nodes transmit with probability 2^{-j}. Then, we know that $p_{ij} = 1 - (1/2^j)(1-1/2^j)^{2^i-1}$. Let t_j be the number of time-slots that nodes are transmitting with probability 2^{-j}. Then, the total probability of failure for any number of active nodes 2^i, needs to be bounded by $2^i \prod_j p_{ij}^{t_j} \leq \epsilon$, or taking logarithms $\sum_j t_j \ln p_{ij} \leq \ln(\epsilon) - \ln 2^i$.

A lower bound can be obtained by minimizing the total number of time-slots needed to satisfy the previous constraints. Here, we reuse our proof technique from [6], i.e., we formulate the problem as a linear program and use a feasible solution of the dual formulation as a lower bound. However, due to the differences between the problems to be solved, the slack variables of the dual need to be more carefully defined, which we do below. The function to be minimized together with the constraints can be formulated as the following *primal* linear program which yields the corresponding *dual*.

Minimize $\mathbf{1}^T\mathbf{t}$,	Maximize $\boldsymbol{\epsilon}^T\mathbf{u}$,	where:
subject to:	subject to:	$\mathbf{t} \triangleq [t_j]$,
$\mathbf{P}\mathbf{t} \geq \boldsymbol{\epsilon}$	$\mathbf{P}^T\mathbf{u} \leq \mathbf{1}$	$\boldsymbol{\epsilon} \triangleq [-\ln(\epsilon) + \ln 2^i]$,
$\mathbf{t} \geq \mathbf{0}$,	$\mathbf{u} \geq \mathbf{0}$,	$\mathbf{P} \triangleq [-\ln(p_{ij})]$.

The primal linear program has a finite minimum solution, and hence its dual has a finite maximum solution. The value of the objective function for every feasible solution of the dual is a lower bound on the minimum value of the objective function for the primal. Thus any feasible solution for the dual will give the lower bound sought. We first define the slack variables as $u_i = 2^i(1-1/\sqrt{e})^2$, and show that these values satisfy the constraints of the dual.

Lemma 1. *For any* $1 \leq j \leq \log \Delta$, $\sum_{i=0}^{\log \Delta}(-\ln p_{ij})u_i \leq 1$.

Proof. We want to prove that

$$\sum_{i=0}^{\log \Delta} \left(-\ln \left(1 - \frac{1}{2^j} \left(1 - \frac{1}{2^j} \right)^{2^i - 1} \right) \right) 2^i \left(1 - \frac{1}{\sqrt{e}} \right)^2 \leq 1$$

Using that for $0 < x < 1$, $e^{-x/(1-x)} \leq 1-x \leq e^{-x}$ [15, §2.68] and maximizing for $j = 1$ it is enough to prove $\sqrt{e}(1 - 1/\sqrt{e})^2 \sum_i (2^i/\sqrt{e}^{2^i}) \leq 1$. Diferentiating the arithmetic-geometric series and replacing, the claim follows.

Now, we use the value of the objective function for this feasible solution to show our lower bound.

Theorem 2. *In order to solve the Group Therapy problem with probability* $1-\epsilon$ *in a multihop Radio Network where the maximum density in any one-hop neighborhood is* Δ*, any fair randomized algorithm requires* $\Omega(\Delta(\log(1/\epsilon) + \log \Delta))$ *time-slots.*

Proof. From lemma 1, we know that $u_i = 2^i(1 - 1/\sqrt{e})^2$ satisfies the constraints of the dual LP, replacing

$$\begin{aligned} \boldsymbol{\epsilon}^T \mathbf{u} &= \sum_{i=0}^{\log \Delta} \left(\ln \frac{1}{\epsilon} + \ln 2^i \right) 2^i \left(1 - \frac{1}{\sqrt{e}} \right)^2 \\ &\in \Omega(\Delta(\log(1/\epsilon) + \log \Delta)). \end{aligned}$$

As proved in [5], when the nodes are distributed uniformly, the density of nodes in any disc of radius $\Theta(r)$ is in $\Theta(\log n)$. Therefore, a simple application of Theorem 2 gives a lower bound of $\Omega(\log n(\log(1/\epsilon) + \log \log n))$ for the Group Therapy problem within uniform density settings. However, it can be proved that to solve even a seemingly simpler problem such as Clear Transmission in a Radio Network with uniformly distributed nodes, it takes $\Omega(\log^2 n)$ *expected* time, which we do as follows.

The topology of active nodes chosen by the adversary for this proof consists of a set of disjoint pairs of cliques connected by a single node. One clique of the pair has node density in $\Theta(1)$, the other in $\Theta(\log n)$ and the intermediate node connects to all nodes in both cliques . We call this construction a *clique-pair*. In order to be disjoint, nodes are woken up so the resulting clique-pairs are separated by a distance of r, the maximum range of transmission of any node.

We first give the intuition of why this structure gives a good lower bound on the number of time steps needed to solve the Clear Transmission problem. Recall that in a multi-hop setting a transmission is a Clear Transmission if no node within two hops of the transmitter transmits in the same time slot. To solve the Clear Transmission problem every node has to receive or produce a Clear Transmission. Hence, in order to solve this problem, a necessary condition is that each node in the low-density clique either receives or produces a Clear

Transmission. That means that there must exist at least one time slot in which exactly one node in the clique pair transmits, that node being the intermediate node or a node in the low-density clique.

Given the different densities and that the protocol is fair, when the sum of probabilities of transmission in the low density clique reaches a constant, and therefore the probability of having a succesful transmission in that clique is constant, the sum of probabilities of transmission in the 2-hop neighboring high density clique is asimptotically more than a constant and the probability of silence is low. On the other hand, when the sum of probabilities of transmission in the whole clique-pair reaches a constant, and the probability of having a non-colliding transmission is high, the probability that the transmitting node is in the low-density clique or it is the intermediate node is low. Then, the probability that nodes in the low density clique produce or receive a Clear Transmission fast is low.

Lemma 2. *Given a Radio Network with nodes deployed as a connected RGG, the total number of clique-pairs activated by the adversary is in* $\Theta(n/\log n)$ *w.h.p.*

Proof. It follows from the $\Theta(\log n)$ density bound in any disk of radius $\Theta(r)$ proved in [5].

Theorem 3. *Every fair randomized algorithm takes* $\Omega(\log^2 n)$ *expected time in order to solve the Clear Transmission problem in a multi-hop Radio Network where nodes are deployed uniformly at random.*

Proof. The proof is based on minimizing the probability of failing to achieve a Clear Transmission in a low density clique. The details are omitted in this extended abstract for brevity.

4 An Optimal Upper Bound for the Group Therapy Problem

A common observation in the literature is that a fair protocol, i.e., a protocol where all nodes are assumed to use the same probability of transmission in the same time slot, has a higher probability of achieving a non-colliding transmission when the probability and the inverse of the number of active nodes agree up to a constant factor and this probability is lower otherwise. Therefore, a main challenge for any protocol is to estimate the density accurately and fast. However, as we show in this section, if all nodes have to achieve successful transmissions by means of a fair protocol it is enough to know the maximum density to achieve a running time of $O(\Delta \log n)$ w.h.p. In achieving a Clear Transmission for all nodes, the Group Therapy problem is also solved. Thus, given the lower bound of Theorem 2, it is optimal for the latter problem. We leave open the question of whether it can be done faster or not using adaptive algorithms. The algorithm is simple to describe, for a network where the maximum density of nodes in any disk of radius r is Δ, every node repeatedly transmits with probability $1/\Delta$.

Theorem 4. *Given a multihop Radio Network where the maximum density in any one-hop neighborhood is Δ, using the protocol described above every node achieves a Clear Transmission within $O(\Delta \log n)$ time steps w.h.p.*

Proof. For a given node, consider a circle of radius $2r$ centered on it. This circle can be completely covered by a constant number, say β_1, of circles of radius r. In each of these circles there are at most Δ nodes, since Δ is the maximum number of nodes in any one-hop neighborhood and all nodes within a circle of radius r are connected. Therefore, $\beta_1 \Delta$ is an upper bound of the number of nodes in the 2-hop neighborhood of any node. Hence, the probability of *some* node not achieving a Clear Transmission after $\beta_2 \Delta \log n$ steps, where β_2 is a constant is

$$Pr(fail) \leq n \left(1 - \frac{1}{\Delta}\left(1 - \frac{1}{\Delta}\right)^{\beta_1 \Delta}\right)^{\beta_2 \Delta \log n}$$

$$\in O(n^{-\gamma}), \text{ for some constants } \beta_2, \gamma > 0.$$

Where we used that for all $n \geq 1$ and $|x| \leq n$, $e^x(1-x^2/n) \leq (1+x/n)^n \leq e^x$ [16].

5 Non-uniform Density Network Initialization

As proved in [5], under the Weak Sensor Model, an optimal network should have low hop-stretch while maintaining links to a constant number of neighbors due to memory constraints. The hop-stretch is the maximum, among all pairs of nodes, of the ratio between the minimum number of hops in a path connecting two nodes, and the optimal number of hops given by the Euclidean distance and the maximum range. In the same paper, it was presented a distributed protocol that builds from scratch a network with such a topology, under the assumption that nodes are deployed uniformly at random sufficiently densely to ensure connectivity w.h.p. We show here that even if the density of nodes is not uniform, as long as it is *Smooth* as defined in Section 1, a network with such a topology can be obtained fast using the same general technique adequately implemented for this setting.

To model the reachability of nodes we use the *Geometric Graph Model* or $\mathcal{G}_{n,r,\ell}$, where n nodes are deployed in a space of size $[0, \ell]^2$, and a pair of nodes is connected if and only if they are at an Euclidean distance of at most r. An instance of $\mathcal{G}_{n,r,\ell}$ is called a *Geometric Graph* (GG) and noted $G(n, r, \ell)$. Given that the network we aim to obtain has to have low hop-stretch and constant number of neighbors, the graph that models its topology has to have constant degree and asymptotically optimal path length in terms of number of edges. We call such a graph a *Constant-degree Hop-optimal Spanning Graph* (CHSG).

In [5] was proved the existence of a CHSG subgraph of any connected RGG by means of a dissection technique called bin-covering [17]. Further, in the same paper was given a *Disk Covering Scheme* that produces such a subgraph. Given the smooth distribution assumed here, the minimum density of nodes in any disc of radius $\Theta(r)$ is $\Omega(\log n)$. Therefore, the same results apply to our setting, i.e., given a GG with smooth distribution of nodes, the Disk Covering Scheme produces a CHSG. The Disk Covering Scheme has four phases, namely, small disk layout, bridge interconnection, disks expansion and local spanner construction. The first three phases of the Disk Covering Scheme can be implemented as detailed in the journal version of [5]. Given that in our setting the maximum density is Δ, the probabilities of transmission and counters used need to be changed appropriately. We omit the details in this extended abstract for brevity. The last phase of spanner construction for smooth distributions is detailed in the following section.

5.1 Spanner Construction

The last phase of the Disk Covering Scheme is the construction of a constant-degree spanner within each expanded disk. As shown in [5], the diameter of the spanner must be logarithmic in order to achieve asymptotically optimal hop-stretch. In this paper, we consider Sensor Networks where the maximum density of nodes in any one-hop neighborhood is an arbitrary value Δ bounded only by n. Thus, a straightforward solution such as a linked list can not be used. Instead, we simply use a balanced binary tree. To build such a tree, we locally rank the nodes according with their unique ID and the ID of the bridge node that covers them. Once unique consecutive labels given by the rank within the disk are assigned to all nodes, each node can easily compute to which nodes is connected within the tree.

We give here a description of the distributed algorithm and we omit the details in this extended abstract for brevity. The spanner construction algorithm consists of three phases. First, every node broadcasts its ID keeping track of the ID of its predecessor among the nodes covered by the same bridge for $\Theta(\Delta \log n)$ steps. As we prove in Theorem 4, at this point all nodes have achieved a Clear Transmission w.h.p. so, all nodes have received a transmission from their local predecessor. To obtain their local rank, nodes enumerates themselves one by one in a second phase as follows. Upon receiving the rank i of its predecessor, a node defines its rank as $i+1$ and broadcasts it with constant probability for $\Theta(\log n)$ steps. As shown in lemma 3, there will be at least one transmission without collision w.h.p. The first node in this ordering does not have any predecessor so it starts this phase of the algorithm with rank 1. At this point, all nodes know their local rank and it only remains to connect them as a balanced tree. A final phase broadcasting the rank where node i connects to nodes $\lfloor i/2 \rfloor$, $2i$ and $2i+1$, if they exist, achieves this. The root of such a tree is therefore the node with the smallest local rank which connects to the bridge.

In order to avoid conflicts with nodes waking up while building the spanner, the Sensor Network initialization protocol has to include an initial waiting phase of $\beta\Delta \log n$ time steps, for some constant β. Nodes can be covered by more than one bridge but, given the geometric restrictions, every node is covered by a constant number of them. Messages and bookkeeping must be replicated for each covering bridge as needed. Nodes running other phases may introduce interference but as long as the sum of their probabilities of transmission is a constant, the analysis can be done as if each phase runs in a different channel in the presence of a source of noise of constant probability of transmission, which we fold into the constants included in the analysis. The details follow.

Lemma 3. *Any node running the second phase of the spanner construction algorithm described above, achieves a transmission without collision within* $O(\log n)$ *steps w.h.p.*

Proof. Let $\beta_1 \in O(1)$ denote the maximum number of interfering neighbors also running the second phase of the spanner construction algorithm. Let $1/\beta_3 \in O(1)$ be the probability of transmission used by a node running such phase. Let $Pr[\text{fail}]$ denote the probability that any node fails to transmit without collision after $\beta_2 \log n$ steps for some constant β_2. Using the union bound and for some constants $\beta_1, \beta_2, \beta_3$

$$\begin{aligned} Pr[\text{fail}] &\leq n\left(1 - \frac{1}{\beta_3}\left(1 - \frac{1}{\beta_3}\right)^{\beta_1}\right)^{\beta_2 \log n} \\ &\in O(n^{-\gamma}), \text{ for some constant } \gamma > 0 \end{aligned}$$

Theorem 5. *Any node running the spanner algorithm joins the spanner within* $O(\Delta \log n)$ *steps w.h.p.*

Proof. The first and third phase take $O(\Delta \log n)$ time by definition of the algorithm. In the second phase, each of the at most Δ nodes in turn transmit for $O(\log n)$ steps. Hence, the overall running time of the algorithm is $O(\Delta \log n)$. As shown in Theorem 4, every node achieves at least one non-colliding transmission within $O(\Delta \log n)$ steps w.h.p. therefore the claim follows.

6 Acknowledgements

We would like to thank Rohan Fernandes for helpful discussions. This research was supported in part by DIMACS, Center for Discrete Mathematics & Theoretical Computer Science, grant numbered NSF CCR 00-87022.

References

1. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cyirci. Wireless sensor networks: A survey. *Computer Networks*, 38(4):393–422, 2002.
2. R. Bar-Yehuda, O. Goldreich, and A. Itai. On the time-complexity of broadcast in multi-hop radio networks: An exponential gap between determinism and randomization. *Journal of Computer and System Sciences*, 45:104–126, 1992.
3. D. M. Blough, M. Leoncini, G. Resta, and P. Santi. The k-neigh protocol for symmetric topology control in ad hoc networks. In *Proc. of the 4th ACM international symposium on Mobile ad hoc networking and computing*, pages 141–152, 2003.
4. B. Chlebus, L. Gąsieniec, A. Lingas, and A. Pagourtzis. Oblivious gossiping in ad-hoc radio networks. In *Proc. of 5th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 44–51, 2001.

5. M. Farach-Colton, R. J. Fernandes, and M. A. Mosteiro. Bootstrapping a hop-optimal network in the weak sensor model. In *Proc. of the 13th Annual European Symposium on Algorithms*, volume 3669 of *Lecture Notes in Computer Science*, pages 827–838. Springer-Verlag, Berlin, 2005.
6. M. Farach-Colton, R. J. Fernandes, and M. A. Mosteiro. Lower bounds for clear transmissions in radio networks. In *Proceedings of the 7th Latin American Theoretical Informatics Symposium*, volume 3887 of *Lecture Notes in Computer Science*, pages 447–454. Springer-Verlag, Berlin, 2006.
7. M. Gerèb-Graus and T. Tsantilas. Efficient optical communication in parallel computers. In *4th Annual ACM Symposium on Parallel Algorithms and Architectures*, pages 41–48, 1992.
8. T. Jurdziński and G. Stachowiak. Probabilistic algorithms for the wakeup problem in single-hop radio networks. *Theory of Computing Systems*, 38(3):347–367, 2005.
9. H. Karl and A. Willig. A short survey of wireless networks. Technical Report TKN-03-018, Technical University Berlin, October 2003.
10. D. R. Kowalski. On selection problem in radio networks. In *Proceedings 24th Annual ACM Symposium on Principles of Distributed Computing*, pages 158–166, 2005.
11. V. S. A. Kumar, M. V. Marathe, S. Parthasarathy, and A. Srinivasan. End-to-end packet-scheduling in wireless ad-hoc networks. In *Proc. of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1021–1030, 2004.
12. E. Kushilevitz and Y. Mansour. An $\Omega(D \log(N/D))$ lower bound for broadcast in radio networks. *SIAM Journal on Computing*, 27(3):702–712, 1998.
13. D. Liu and M. Prabhakaran. On randomized broadcasting and gossiping in radio networks. *Lecture Notes in Computer Science*, 2387:340–349, 2002.
14. C. U. Martel. Maximum finding on a multiple access broadcast network. *Inf. Process. Lett.*, 52:7–13, 1994.
15. D. S. Mitrinović. *Elementary Inequalities.* P. Noordhoff Ltd. - Groningen, 1964.
16. R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1995.
17. S. Muthukrishnan and G. Pandurangan. The bin-covering technique for thresholding random geometric graph properties. In *Proc. of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 989–998, 2005.
18. K. Nakano and S. Olariu. Energy-efficient initialization protocols for radio networks with no collision detection. In *ICPP*, pages 263–270, 2000.
19. V. A. S. Ponduru and A. Bharathidasan. Sensor networks: An overview. Technical report, University of California, Davis, 2003.
20. P. Rentala, R. Musumuri, U. Saxena, and S. Gandham. Survey on sensor networks. http://citeseer.nj.nec.com/479874.html.
21. K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie. Protocols for self-organization of a wireless sensor network. *Personal Communications, IEEE*, 7(5):16–27, 2000.
22. W. Song, Y. Wang, X. Li, and O. Frieder. Localized algorithms for energy efficient topology in wireless ad hoc networks. In *Proc. of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pages 98–108, 2004.
23. O. Younis, M. Krunz, and S. Ramasubramanian. Node clustering in wireless sensor networks: Recent developments and deployment challenges. *IEEE Network Magazine*, 20:20–25, 2006.