Hybrid Sensor Networks: A Small World

Gaurav Sharma
Center for Wireless Systems and Applications
School of Electrical and Computer Engineering
Purdue University
gsharma@purdue.edu

Ravi Mazumdar
Dept. of Electrical and Computer Engineering
University of Waterloo
mazum@ece.uwaterloo.ca

ABSTRACT

In this paper, we investigate the use of limited infrastructure, in the form of wires, for improving the energy efficiency of a wireless sensor network. We call such a sensor network - a wireless sensor network with a limited infrastructural support - a *hybrid sensor network*. The wires act as short cuts to bring down the average hop count of the network, resulting in a reduced energy dissipation per node. Our results indicate that adding a few wires to a wireless sensor network can not only reduce the average energy expenditure per sensor node, but also the non-uniformity in the energy expenditure across the sensor nodes.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design; G.3 [**Probability and Statistics**]: Probabilistic algorithms

General Terms

Algorithms, Performance, Theory, Design

Keywords

Sensor Networks, Graph Theory, Small World Networks, System Design, and Routing.

1. INTRODUCTION

Recent advances in the MEMS-based sensor technology and wireless communications have enabled the development of relatively inexpensive and low-power wireless sensors. The common vision is of a large sensor network, consisting of thousands of these tiny devices working in collaboration with each other, to achieve some common goal. Potential applications of such networks include habitat monitoring, soil quality monitoring, detection of hazardous chemicals and forest fires, military surveillance, monitoring seismic activity, etcetera.

There are several challenges that arise in the design of sensor networks. One of the main challenges is to design energy efficient sensor networks. This has lead to significant research at various levels of the protocol stack. At the level of the physical layer, the researchers are developing low-power hardware for the sensor nodes [1, 2]. Several power aware protocols [3, 4] have been developed for the medium access control (MAC) in sensor networks. In addition to this, various energy efficient routing, clustering, and data aggregation protocols have also been developed [5, 6].

The main source of energy dissipation in a sensor node is the transmission and reception of packets over the wireless interface. In this paper, we investigate the use of a limited infrastructure, in the form of wires, for reducing the energy dissipation in a wireless sensor network. We show that the wires can be used as short cuts to reduce the average hop count of the network, resulting in a reduced energy dissipation per node. We also show that the addition of wires can significantly reduce the non-uniformity in the energy dissipation across the sensor nodes. The reduction in the per-node energy dissipation coupled with more uniform energy dissipation across the sensor nodes, significantly increases the network lifetime.

Note that in many applications like tracking enemy movements and remote surveillance, it might not be possible to augment the network with wires. For some other short term applications, adding wires to the network might be infeasible because of the cost involved in wiring. However, for applications like habitat monitoring (see [7, 8]), which require long term data collection and dissemination, it might be economically feasible to augment the network with wires.

Networked Info-mechanical Systems (NIMS), which are currently being developed at UCLA and partner universities [9, 10], are particular instances of the type of scenario we consider in this paper. NIMS infrastructure consists of a collection of steel cables, each attached to any two points - buildings, trees, or other natural structures - that serve as suspension points. The nodes which are suspended on the cables collect data about the environment through a range of sensors that can be lowered or elevated; and also move, activate and recover fixed nodes set along the cable pathway. They also have the ability to dock when necessary to recharge their energy source. In addition to these aerially suspended nodes, which can replenish their energy resources, NIMS architecture also contains untethered wireless sensor nodes. These nodes provide access to non-navigable areas. An efficient use of the infrastructural support along with the capability to harvest solar energy might possibly elim-
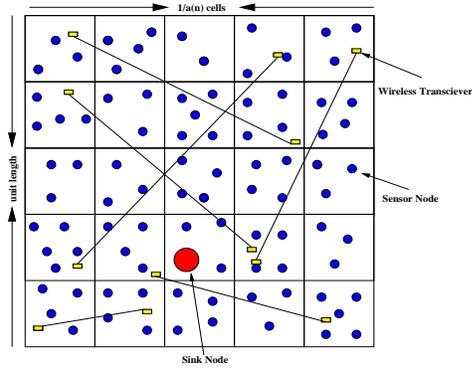
Figure 1: Complete view of the sensor network showing the wireless sensor nodes, wires with wireless transceivers attached at their ends, and the sink node.

inate the need for replacement/maintenance of the untethered wireless sensor nodes in NIMS, and other such hybrid sensor networks.

The rest of the paper is organized as follows. Our model and related work are discussed in Section 2. The average energy dissipation and the non-uniformity of the energy dissipation across the nodes in a wireless sensor network are the topics we study in Section 3. These issues are further studied in the context of a hybrid sensor network; for the static sink node case in Section 5, and for the mobile sink node case in Section 5. We summarize our main findings in Section 6, and end this paper with some concluding remarks in Section 7.

2. MODEL AND RELATED WORK

2.1 Network and Transmission Model

We consider a sensor network formed by n sensor nodes, distributed uniformly inside a square $\mathcal{A}$ of unit area (see Figure 1). The sensor nodes are equipped with wireless transceivers, for communication purposes. There is a single sink node, located arbitrarily inside $\mathcal{A}$. The sensor nodes as well as the sink node could either be static or mobile. Although some of our results are proved under static network conditions, they would also hold for mobility models under which the node distribution remains uniform at all times. The random walk mobility model, the random direction mobility model, and the random way-point mobility model (see [11] for a description of these mobility models) on a 2-d torus or a sphere are canonical examples of such mobility models.

We consider a homogeneous scenario where all sensor nodes generate data at the same rate, which is delivered to the sink node, possibly using multiple hops. For simplicity, we consider a slotted system in which each slot is of duration T seconds. All the transmissions begin and end at slot boundaries. Let τ_t denote the set of nodes which are transmitting during the time-slot t. We assume that all transmitters employ the *same power level*, P; and consider a transmission model which is similar to the physical model introduced in [12]. Under our transmission model, a node i is capable of transmitting a packet at W *bits/sec* to node j during the time-slot t if, and only if, the SINR at node j is above a threshold, say β, i.e.,

$$\frac{P\gamma_{ij}(t)}{N_o + \frac{1}{L}\sum_{k\neq i, k\in\tau_t} P\gamma_{kj}(t)} > \beta, \qquad (1)$$

where $\gamma_{ij}(t)$ is the channel gain from node i to node j during the time-slot t; N_o is the background noise power; and L is the processing gain of the system. The channel gain is assumed to be of the form

$$\gamma_{ij}(t) = \frac{1}{d_{ij}^{\alpha}(t)}, \qquad (2)$$

where $d_{ij}(t)$ is the distance between the nodes i and j during the time-slot t; and typically, $\alpha \in (2, 6)$. Since there is no power control, the per-packet energy spent in the RF amplifier at the transmitter is PT. Let E_o be the per-packet energy spent in the transmitter or the receiver circuitry, then

$$E_t = PT + E_o; \ E_r = E_o, \qquad (3)$$

where E_t and E_r are the per-packet energy spent at the transmitter and the receiver, respectively.

Note that in the above interference model we assumed that the network topology does not change during a timeslot. This is a reasonable assumption since the time-scale at which the network topology changes, either due to node mobility or due to node failure, is usually much slower than the time-scale for a packet transmission. In fact, this assumption is not really essential for the main results of this paper to hold, however, relaxing this assumption would require that we carefully look at the node mobility process; which is not the main goal of this paper.

The infrastructural support is assumed to be present in the form of wires, equipped with wireless transceivers at their ends. We assume that the wireless transceivers attached to the wire ends can replenish their energy, like in NIMS [9, 10], so that there are no energy constraints at these transceivers. The wires can either be placed in a random or deterministic manner. The wire placement schemes are considered in Sections 4 and 5.

Next, we discuss some related work in the literature.

2.2 Related Work

Our work has similarities to some aspects of the small world networks that are modeled as a connectivity graph having the three properties; namely: sparseness, clustering, and small diameter. Watts and Strogatz [13] were among the first to show that the addition of a few random connections (or rewiring of a few local connections*) to a ring lattice can make it into a small world graph. Since then, this idea has become extremely popular, and has been used in many different contexts, like, for example, in [14] it is used to do efficient resource query in wireless ad hoc networks. We note there are some significant differences between our work and the earlier work on small world networks. Some of them we list below:

- Since our goal is the optimal network design, unlike most of the previous works on small world networks, we do not restrict ourselves to randomly placed short cuts (wires), i.e., we consider both random as well as deterministic placement of short cuts (wires).

---

*Note that the rewiring can be problematic as it can result in loss of connectivity.

- For a network to be termed a small world network its diameter should be on the order of $\log n$ or at most polylogarithmic in n which, under random placement of short cuts, requires having about one long range connection (short-cut) per node. As a result, earlier works on small world networks (see, for example, [15]) have focused on a regime in which there are about the same number of short cuts as the number of nodes in the network. However, due to practical considerations, we are mainly interested in a regime in which the number of short cuts is much smaller in comparison to the number of nodes in the network.

Also similar in spirit to our work is the work in [16]. In [16], the authors investigate the use of a limited infrastructure for improving the energy efficiency of a wireless sensor network. To the best of our knowledge, they were the first to show that a few deterministically placed wired short-cuts can improve the energy efficiency of a wireless sensor network. Next, we list the main differences between the work in [16] and that of this paper, thus highlighting the main contributions of this paper:

- The objective of [16] is to analyze one particular scheme for placing the wires, whereas, the objective of this paper is to find out the best possible gains[†] that are achievable, given a cost budget; and propose schemes which can either achieve or come close to achieving those gains.

- In the scheme considered in [16], all wires originate from the sink node (which is assumed to be in the center of the network) and end at points which are equidistant from the sink node. It is easily seen that such a wire placement scheme can only achieve a constant factor of reduction in the energy expenditure for all n, whereas, some of the schemes we propose can achieve energy savings on the order of $\sqrt{n/\log n}$ (see Sections 4 and 5).

- Unlike [16], we do not restrict ourselves to the static sink node case.

Several energy efficient routing protocols for wireless sensor networks have been proposed in the literature. Most of these protocols [5, 6, 17] are based on data centric routing, and require application awareness. The energy efficiency in these protocols is achieved by means of data aggregation, caching, and elimination of redundant packet transfers by application specific naming of data (meta-data). Whenever application specific knowledge is available, these protocols can be used in a hybrid sensor network to further reduce the energy dissipation in the sensor nodes. In fact, in a hybrid sensor network the transceivers attached to the wire ends can be equipped with a small computing unit for data compression, a role normally assigned to the cluster heads in wireless sensor networks (see, for example, [5], [18], [19], [20]). This can further reduce the energy dissipation in the wireless sensor nodes. However, this is a topic for future research; we do not consider it here, in this paper.

Hybrid ad hoc networks have been studied in [21, 22]. Hybrid sensor networks differ from hybrid ad hoc networks in

that the communication scenario in a hybrid sensor network is many-to-one, rather than the many-to-many communication scenario, which is typical of hybrid ad hoc networks. More importantly, in [21, 22], the authors study the throughput capacity of hybrid ad hoc networks, whereas, we study the energy dissipation in a hybrid sensor network.

Recently, we have become aware of a related work that aims to exploit the heterogeneity in sensor networks [23]. The authors identify three common types of hardware heterogeneities: *computational heterogeneity* - some nodes having more computational power than the others, *link heterogeneity* - some nodes having a direct and highly reliable connection to the sink node, and *energy heterogeneity* - some nodes having unlimited energy supplies. They then study the impact of the energy heterogeneity and link heterogeneity on the average energy consumption, the network lifetime, and the end-to-end successful packet delivery rate. We note that by connecting the (static) sink node to the other parts of the sensor network using wires (see Section 4), in a certain way, we are able to exploit both the link heterogeneity as well as the energy heterogeneity (wireless transceivers that are attached to the wire-ends can be thought of as nodes having unlimited energy supplies and also a direct connection to the sink node). Although, we do not study the end-to-end successful packet delivery rate in this paper, it is straightforward to use our results regarding the average hop count of the network to bound the end-to-end successful packet delivery rate. The work in [23] complements our work in the sense that they consider realistic MAC and routing schemes for a specific network setting, whereas, we consider a general network setting but with ideal MAC and routing. Finally, our analytical approach allows us to provide results for networks of arbitrary (but large) sizes and arbitrary number of wires, which is not possible under a simulation based approach, as in [23].

2.3 Notation

We use the following standard notation throughout the rest of this paper. For $f(n), g(n) \geq 0$ for all n, we say that

$$f(n) = o(g(n)) \leftrightarrow \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0,$$

$$f(n) = O(g(n)) \leftrightarrow \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty,$$

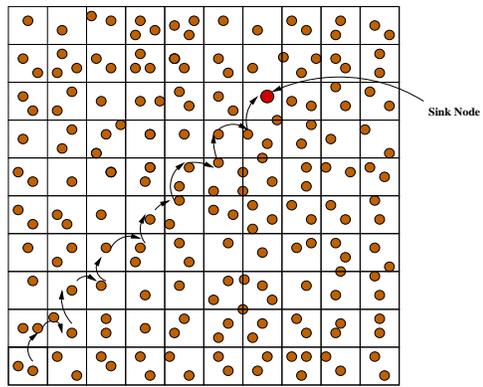$$f(n) = \omega(g(n)) \leftrightarrow g(n) = o(f(n)),$$

$$f(n) = \Theta(g(n)) \leftrightarrow f(n) = O(g(n)) \text{ and } g(n) = O(f(n)).$$

3. ENERGY DISSIPATION IN A WIRELESS SENSOR NETWORK

We now study the energy dissipation in a wireless sensor network. Consider a sensor network formed by n wireless sensor nodes and a single sink node, placed inside the square $\mathcal{A}$. The sensor nodes are assumed to be uniformly distributed within $\mathcal{A}$. The position of the sink node can be arbitrary. The network is divided into square cells of size $a(n) = \sqrt{\frac{32 \log n}{n}}$[‡]. Thus there are a total of $\mathcal{N}(n) = 1/a^2(n)$ cells in the network. Note that this choice of the cell size is not arbitrary; the cell size has been chosen keeping in mind the connectivity of the network, as will become clear shortly.

[†]The "gain" here is in terms of reducing the per node energy dissipation and the non-uniformity in the energy dissipation across the nodes.

[‡]Note that for this partitioning to be meaningful, the number of nodes must at least be on the order of a few hundred.

All the transmissions are carried out at the same power level, between *adjacent* cells; by *adjacent cells* we mean the cells which are horizontal or vertical neighbors of a given cell. We assume that each cell uses the greedy geographic routing to make the forwarding decisions, i.e., each cell forwards its packets to a neighboring cell which is "closest" to the cell containing the sink node. The distance metric used for picking the "closest" neighbor is the distance between the center of the neighboring cell and that of the cell containing the sink node. Note that each cell in the network must know the location of the cell containing the sink node. If the sink node is static, its location needs to be flooded in the network only once to enable every cell to use the greedy geographic routing. However, with a mobile sink node, the location information needs to be flooded "every now and then," depending on the speed of the sink node. A typical path taken by a packet from its source node to the sink node, using the greedy geographic routing, is shown in the Figure 2.

REMARK 1. *Note that the reason for considering the above division of the network into cells is to simplify the analysis. One would expect that our main results should hold even when the cellular structure is not present and distributed MAC and routing schemes are used. For example, GFG [24] or its variant GPRS [25] can be used for routing when the location information is available; and the scheme presented in [26] can be used to approximate greedy geographic routing when the location information is unavailable.*

Next, we provide two key results which show that the nearest neighbor communication scenario along with the greedy geographical routing can guarantee almost sure delivery of the packets to the sink node. Unless otherwise stated, we will consider the nodes to be stationary. As discussed before, the results can be generalized to the case of mobile nodes as well. We start with the following Lemma which shows that if n is large, then each cell contains $\Theta(\log n)$ nodes *almost surely*.

LEMMA 1. *Let $\mathcal{N}_i$ be the number of nodes in cell i. Then, we have*

$$\mathbb{P}\left(\liminf_{n\to\infty}\{16\log n \le \mathcal{N}_i \le 48\log n,\ 1 \le i \le \mathcal{N}(n)\}\right) = 1$$

which may alternatively be stated as: for large n,

$$\{16\log n \le \mathcal{N}_i \le 48\log n,\ 1 \le i \le \mathcal{N}(n)\}\ \textit{almost surely.}$$

PROOF. Note that the $\mathcal{N}_i$'s are all Bernoulli$(n, 1/\mathcal{N})$ random variables. For a fixed i, using the Chernoff bound (for example, see [27], page 68), we obtain

$$\mathbb{P}\left(\mathcal{N}_i < \frac{n}{2\mathcal{N}(n)}\right) \le e^{-\frac{n}{12\mathcal{N}(n)}} \qquad (4)$$

$$\mathbb{P}\left(\mathcal{N}_i > \frac{3n}{2\mathcal{N}(n)}\right) \le e^{-\frac{n}{8\mathcal{N}(n)}} \qquad (5)$$

Combining (4) and (5), and substituting the value of $\mathcal{N}$, we obtain

$$\mathbb{P}\left(|\mathcal{N}_i - 32\log n| > 16\log n\right) \le 2/n^{8/3} \qquad (6)$$

Noting that $\mathcal{N}(n) < n$ for $n > 1$, and applying the union bound, we have

$$\mathbb{P}\left(|\mathcal{N}_i - 32\log n| > 16\log n, 1 \le i \le \mathcal{N}(n)\right) < 2/n^{5/3}$$

Figure 2: A typical path taken by a packet from its source to the sink node.

Now, since

$$\sum 2/n^{5/3} < \infty,$$

the result follows from the first Borel-Cantelli Lemma and noting that

$$\left\{\limsup_{n\to\infty}\mathcal{A}_n\right\} = \left\{\liminf_{n\to\infty}\mathcal{A}_n^c\right\},$$

for a sequence of events $(\mathcal{A}_n)_{n=1}^\infty$. $\square$

The next issue we address is how to schedule the transmissions across the network in a way so that they do not interfere *too much* with each other. Obviously, if the transmissions from various cells were not co-ordinated then some of them might fail due to the interference from other simultaneous transmissions in the network. The following Lemma provides a transmission schedule in which each cell can successfully transmit at least once within a fixed number of time-slots.

PROPOSITION 1. *Under the transmission model considered in Section 2.1, there exists a transmission schedule in which each cell can successfully transmit at least once in every K^2 time-slots; with $K > 0$ being an integer, independent of n.*

PROOF. The proof is constructive; not only do we show that such a transmission schedule exists, we also construct one. Our construction is similar to the construction of a transmission schedule under the *protocol model* of interference in [28]. The idea is to construct *equivalence classes* of cells such that all cells in an equivalence class can transmit simultaneously without interfering *too much* with one another. Let $K > 0$ be an integer, to be fixed later. We say that cell u is related to cell v, denoted $u \sim v$, if u and v are a vertical and horizontal distance of exactly some multiple of K away from each other. It is easily seen that this is an equivalence relation; which divides the network into K^2 (disjoint) equivalence classes. The shaded cells in the Figure 3 belong to one such equivalence class for $K = 4$.

Now the scheduling strategy is to allow only the nodes in one particular equivalence class to transmit during a time-slot; and to schedule the equivalence classes in a round-robin fashion during successive time-slots. Clearly, each cell belongs to some equivalence class, and would therefore get a
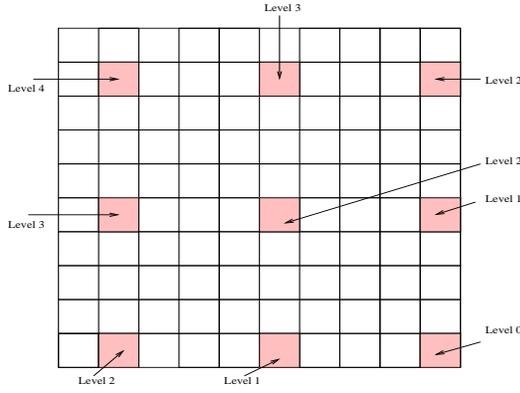
Figure 3: An equivalence class of cells for $K = 4$.

chance to transmit once every K^2 time-slots. Next, we show that under such a transmission schedule the SINR at each receiver would be at least β, provided K is large enough.

Consider an arbitrary cell in the network, say c_o. Consider a node in c_o, say node i, transmitting to a node in one of the adjacent cells, say node j. Let us call all cells at a lattice distance of lK from c_o, the level l cells. Let Φ_l denote the set of level l cells, and let $|\Phi_l|$ be its cardinality. Clearly, $|\Phi_l| \leq 4l$ for all l. Observe that a node which is inside one of the cells in Φ_l, is at least at a distance of $(Kl-2)a(n)/2$, from a node that is inside a cell adjacent to c_o (see Figure 3). Using this, we can bound the interference power at node j, denoted by $\mathcal{I}$, as follows:

$$\mathcal{I} \leq \sum_{l=1}^{\lfloor 2/Ka(n) \rfloor} 4 \cdot l \cdot \frac{P2^\alpha}{(kl-2)^\alpha a^\alpha(n)}$$

$$\leq \frac{P2^{\alpha+2}}{(K-2)^\alpha a^\alpha(n)} \sum_{l=1}^{\infty} l^{1-\alpha}$$

$$\leq \frac{P2^{\alpha+2}C}{(K-2)^\alpha a^\alpha(n)}$$

where $C = \sum_{l=1}^{\infty} l^{1-\alpha} < \infty$, since $\alpha > 2$. Now since i and j are in adjacent cells, the distance between them is less than $\sqrt{5}a(n)$, and therefore the received power at j is greater than $P/(\sqrt{5}a(n))^\alpha$. Thus, if

$$K \geq 2 + 2^{1+1/\alpha}C^{1/\alpha}\left(\frac{1}{5^{\alpha/2}\beta} - \frac{N_o a^\alpha(n)}{P}\right)^{-1/\alpha},$$

then the SINR at j would be greater than β. Now if P is large enough, say $P = \epsilon N_o a^\alpha(n) 5^{\alpha/2}\beta$ for some $\epsilon > 1$, then we see that K can be chosen to be a constant (independent of n). Since the choice of c_o, i, and j, is arbitrary, we have proved our claim. $\square$

REMARK 2. *Using the above scheduling scheme along with the greedy geographic routing, each packet would almost surely reach the sink node, thus ensuring the almost sure connectivity of the network (see [29]).*

Now we turn our attention to the main topic of this section, namely, the energy expenditure in a wireless sensor network. Let λ packets/s be the common rate of generation of packets at each node in the network. The per-node average rate of energy drainage, denoted by $\mathbb{E}[\mathcal{E}_D]$, can then be calculated as:

$$\mathbb{E}[\mathcal{E}_D] = \lambda \mathbb{E}[\mathcal{H}](E_t + E_r),$$

where $\mathbb{E}[\mathcal{H}]$ is the average number of hops, with the average being taken across all the nodes and across all possible realizations of the random distribution of the nodes. We now estimate $\mathbb{E}[\mathcal{E}_D]$.

PROPOSITION 2. $\mathbb{E}[\mathcal{H}] = \Theta\left(\sqrt{n/\log n}\right)$. *Thus,* $\mathbb{E}[\mathcal{E}_D] = \lambda(E_t + E_r)\Theta\left(\sqrt{n/\log n}\right)$.

PROOF. The upper bound follows by noting that a packet generated at any arbitrary node in the network can reach the sink node in at most $2\sqrt{\mathcal{N}(n)} - 2 = \Theta(\sqrt{n/\log n})$ hops. In order to prove the lower bound, observe that for any arbitrary position of the sink node there would at least be $\mathcal{N}(n)/4$ cells at a distance greater than $1/4$ from the cell containing the sink node. This follows easily by a simple geometrical argument (see Figure 1). Now, a packet generated at one of these cells would take at least $1/4a(n) = \sqrt{\mathcal{N}(n)}/4$ hops to reach the destination node. Using Lemma 1, it follows that *almost surely* each of these cells would contain more than $16 \log n$ nodes, for large n. Thus, we obtain

$$\mathbb{E}[\mathcal{H}] \geq \frac{\mathcal{N}(n)}{4} \cdot 16 \log n \cdot \frac{\sqrt{\mathcal{N}(n)}}{4n}$$

Noting that $\mathcal{N}(n) = 1/a^2(n) = \Theta(n/\log n)$, the result follows. $\square$

The above Lemma gives the rate of energy drainage averaged across all the nodes in the network. Note that the rate of energy drainage may not be uniform across different nodes in the network. Under static conditions, it is clear that the nodes which are closer to the sink node, will relay more packets than the other nodes, and correspondingly, will drain their energies much quicker than the other nodes. In order to characterize this disparity in the energy dissipation, we define:

DEFINITION 1 (ENERGY DISSIPATION SKEW). *Let E_i be the average rate at which the node i drains its energy due to the transmission and reception of packets. Then, the energy dissipation skew or the EDS is given by:*

$$EDS = \max_{i=1,\dots,n} E_i / \min_{i=1,\dots,n} E_i \qquad (7)$$

The EDS is a random variable whose value depends on how the nodes are distributed as well as on the position of the sink node. Clearly, we would like the EDS to be close to 1, so as to ensure a uniform rate of energy dissipation across the sensor nodes. However as we show now, the EDS is usually very high in case of a wireless sensor network, implying a highly non-uniform rate of energy dissipation across the sensor nodes.

Consider a wireless sensor network using the adjacent cell communication scheme discussed above. Observe that in order to reach the sink node, each packet must pass through one of the four cells that are adjacent to the cell containing the sink node, and each of which contains $\Theta(\log n)$ nodes almost surely. Note also that the nodes located in one of the four corner cells are not involved in any relaying. Therefore, the EDS for a wireless sensor network is $\Theta(n/\log n)$ almost surely.

In the above analysis, we ignored the node mobility; the node mobility can possibly result in an improvement of the EDS over long periods of time. But such an improvement, if any, can only be estimated if complete knowledge of the motion of each node is available. Since we do not assume any knowledge of the node mobility process, we will not consider this issue any further. The EDS we consider should therefore be interpreted as the worst case or the short term EDS.

In the next section, we will show that both the average hop count and the (worst case) EDS can be significantly reduced with the introduction of a small number of wired short cuts in the network.

4. HYBRID SENSOR NETWORKS: STATIC SINK

We now study the energy dissipation in a hybrid sensor network with a static sink node. The sensor nodes can either be static or mobile. The problem we consider is the following: given a budget of $l(n)$ wires per cell, how should the wires be placed so as to minimize the average hop count of the network, and to reduce the EDS as much as possible. Note that our problem formulation is motivated by the cost constraints that one might have in placing the wires, i.e., the value of $l(n)$ would typically be dictated by the cost constraints. It turns out that in order to have non-negligible reductions in the average hop count or the EDS, $l(n)$ must be $\omega(\log n/n)$; that is, the total number of wires, $\mathcal{N}(n)l(n)$, should be an unbounded function of n. Also, due to practical considerations, we will only be interested in the case where $l(n) = O(1)$. In particular, we will assume that $l(n) < 1$ for all n, i.e., less than one wire per cell on an average.

As far as minimizing the average hop count is concerned, it is intuitive to believe that an optimal placement of the wires should have all the wires originating from the cell containing the sink node. This is indeed the case, as we now show.

LEMMA 2. *A wire placement scheme which minimizes the average hop count of the network should have all the wires originating from the cell containing the sink node.*

PROOF. We will give a proof by contradiction. Let Υ be a scheme that minimizes the average hop count of the network. Let c_s be the cell containing the sink node (assumed static). Suppose Υ places a wire between cells $u, v \neq c_s$. Let $d_{\Upsilon}(s, t)$ be the distance in hops between cells s and t under Υ. Note that a "hop" could either be a *wired hop* or a *wireless hop*. Without loss of generality, let us suppose $d_{\Upsilon}(u, c_s) \leq d_{\Upsilon}(v, c_s)$. In this case, it is clear that no node will use the link between u and v, denoted by l_{uv}, in the direction from u to v for routing its packets to the sink node (as this would increase the number of hops). Let Φ denote the set of cells which use l_{uv}, and let $|\Phi|$ denote its cardinality. Now, consider scheme Υ' which is obtained from Υ by replacing l_{uv} with a link between v and c_s. Also, consider Υ'' obtained from Υ by replacing l_{uv} with a link between cells o and c_s, where o is a cell which is not connected with c_s under Υ (such a cell always exists since $l(n) < 1$). Let

$$\Upsilon''' = \begin{cases} \Upsilon', & \text{if } |\Phi| > 0 \\ \Upsilon'', & \text{otherwise.} \end{cases}$$

We claim that the average hop count under Υ''' is strictly smaller than the average hop count under Υ. Note that once

we prove this, we will be done as this would contradict our initial assumption that Υ minimizes the average hop count of the network. Let us first consider $|\Phi| > 0$. Observe that no cell uses l_{uv} in the direction from u to v, for routing its packet to c_s. And for a cell, say t, which uses l_{uv} in the direction from v to u, we have:

$$\begin{aligned} d_{\Upsilon'}(t, c_s) &= d_{\Upsilon}(t, v) + 1 \\ &< d_{\Upsilon}(t, v) + 1 + d_{\Upsilon}(u, c_s) \ (u \neq c_s) \\ &= d_{\Upsilon}(t, c_s) \end{aligned}$$

Since $|\Phi| > 0$, it follows that the average hop count under Υ' is strictly smaller than the average hop count under Υ. Now when $|\Phi| = 0$, Υ'' cannot increase the increase the distance between a cell and c_s. Clearly $d_{\Upsilon''}(o, c_s) < d_{\Upsilon}(o, c_s)$, and therefore, the average hop count under Υ'' is strictly smaller than the average hop count under Υ. ☐

We know now that in order to minimize the average hop count of the network, all the wires should originate from the cell containing the sink node. Next, we determine how much can one possibly reduce the average hop count and the EDS, when placing at most $l(n)$ wires per cell on an average.

PROPOSITION 3. *Under any wiring and routing scheme which uses at most $\mathcal{N}(n)l(n)$ wires, the average hop count must be $\Omega(1/\sqrt{l(n)})$. Furthermore, if the average hop count under such a scheme is $o(\sqrt{n/\log n})$, then the EDS must be $\Omega(1/l(n))$.*

PROOF. First, let us look at the average hop count. In view of Lemma 2, we may assume that all wires originate from the sink node. Observe that the above result is trivial if $l(n) = \Theta(1)$, so assume that $l(n) = o(1)$. In this case, without loss of generality we can assume that $1/4\sqrt{l(n)}$ is an integer; if not, the proof can be carried out with $\lfloor 1/4\sqrt{l(n)} \rfloor$ in place of $1/4\sqrt{l(n)}$. Now consider clusters of cells with $1/16l(n)$ cells each, similar to what is shown in Figure 4. We call such clusters *level1* clusters. Observe that for large n, there would be close to $16\mathcal{N}(n)l(n)$ *level1* clusters. Consider groups of *level1* clusters, each consisting of 9 *level1* clusters. We call such a group a *level2* cluster. Note that a *level2* cluster consists of a *level1* cluster, along with its 8 surrounding *level1* clusters. For large n, the number of *level2* clusters would be close to $16\mathcal{N}(n)l(n)/9$. Now consider one particular *level2* cluster, which does not contain the sink node. Observe that if none of the cells in this *level2* cluster has a wired connection with the cell containing the sink node, then all cells which are inside the *level1* cluster situated in the middle of this *level2* cluster, would have to route their packets over more than $1/4\sqrt{l(n)}$ hops in order to reach the sink node. Since there are a total of $\mathcal{N}(n)l(n)$ wires, there would be close to $7\mathcal{N}(n)l(n)/9 > 2\mathcal{N}(n)l(n)/3$ *level2* clusters which do not have a wired connection. Noting that each *level2* cluster consists of $9/16l(n)$ cells, it follows that the average hop count must at least be

$$\frac{1}{4\sqrt{l(n)}} \cdot \frac{2\mathcal{N}(n)l(n)}{3} \cdot \frac{9}{16l(n)} \cdot \frac{1}{\mathcal{N}(n)} = \frac{3}{32\sqrt{l(n)}}.$$

This proves our claim regarding the average hop count. Now, let us look at the EDS. Suppose the average hop count under the given scheme is $o(\sqrt{n/\log n})$. Observe that for any arbitrary position of the sink node, there would be at least $7\mathcal{N}(n)/8$ cells at a lattice distance of more than
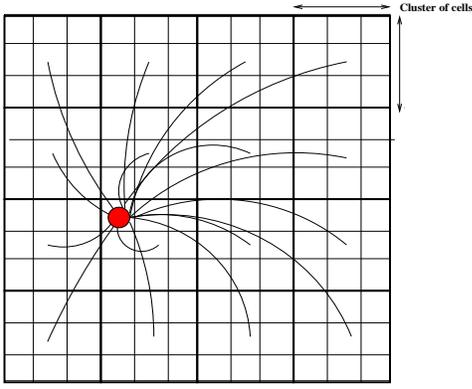
Figure 4: The clustering and wiring of cells for $l(n) = 1/9$.

$1/4a(n)$ from the cell containing the sink node. Now since the total number of wires is $\mathcal{N}(n)l(n)$, and $l(n) = o(1)$, more than $3\mathcal{N}(n)/4$ of these cells will have no wired connections. Recalling that the average hop count of the network is $o(\sqrt{n/\log n})$, it follows that at least $\mathcal{N}(n)/2$ (actually, close to $3\mathcal{N}(n)/4$) of these, must be served by a cell having a wired connection. For if not, then the average hop count would be more than

$$\frac{\mathcal{N}(n)}{4} \cdot \frac{1}{4a(n)} \cdot \frac{1}{\mathcal{N}(n)} = \frac{1}{16a(n)} = \Theta(\sqrt{n/\log n}).$$

Since there are at most $\mathcal{N}(n)l(n)$ cells having a wired connection, the EDS must at least be $1/2l(n)$. $\square$

Having established the lower bounds on the average hop count and the EDS, we now provide a simple deterministic wiring scheme that achieves those bounds.

4.1 Scheme 1

In this scheme, the wires are placed in a deterministic manner. For simplicity of notation, assume that $1/\sqrt{l(n)}$ is an odd integer and $\sqrt{\mathcal{N}(n)/l(n)}$ is an integer. The scheme is the following:

- *Wire Placement*: Group the cells into clusters consisting of $1/l(n)$ cells each, as shown in the Figure 4. For each cluster of cells, connect the cell in the middle the cluster with the cell containing the sink node.

- *Routing*: Each cell, first routes its packet to the nearest cell having a wired connection (using wireless transmissions), and from there onward, the packet is delivered to the sink node using the wired connection.

Clearly, the average hop count under the above scheme is $\Theta(1/\sqrt{l(n)})$. Now let us look at the EDS. Observe that all $(1/l(n))$ cells inside a cluster route their packets through the cell situated in the middle of the cluster, having a wired connection. Note also that the cells at the boundary of a cluster are not involved in any forwarding. Thus, the EDS under the above scheme is $1/l(n)$. These results can easily be extended to the general case, where $1/\sqrt{l(n)}$ is not necessarily an odd integer and $\sqrt{\mathcal{N}(n)/l(n)}$ not necessarily an integer. In the general case, however, the EDS might not be exactly $1/l(n)$; but it will still be $\Theta(1/l(n))$. In view of Proposition 3, this is the best achievable (in order terms)

EDS and average hop count, under any wiring and routing scheme which uses at most $l(n)$ wires per cell.

We have so far assumed that each cell (every node inside each cell) knows the location of the nearest cell having a wired connection. We now consider a simple protocol which can be used for maintaining this information in the presence of node mobility and node failures.

4.1.1 Route Maintenance Protocol 1 (RMP1)

In this protocol, the transceivers attached to the wire ends periodically send out an advertisement (ADV) packet. The ADV packet is received by the nodes which are in the same cell as one of the transceivers. An ADV packet has a *hop count* field, which is initially set to 0. Upon receiving an ADV packet, one of the nodes will broadcast this packet to the adjacent cells, with the hop count field incremented by 1. We have already seen that the nodes which are present in one of the adjacent cells can successfully receive the packet (see the proof of Proposition 1). Each node keeps track of the neighboring cell (or the *next hop*) from which an ADV packet was received; this enables the cell to forward its packets to the "appropriate" cell when trying to reach the sink node. This process is continued until the maximum limit on the *hop count* is reached. The maximum hop count limit can be set close to $1/\sqrt{l(n)}$, so as to limit the number of retransmissions and at the same time ensure that every cell receives the ADV packet from at least one wire. Observe that with the above choice of the hop count each ADV packet is received by $\Theta(1/l(n))$ cells, and since there are $\mathcal{N}(n)l(n)$ wires in the network, the total number of transmissions of ADV packets in the network is $\Theta(\mathcal{N}(n)) = \Theta(n/\log n)$ per update; and the total number of receptions of ADV packets is $\Theta(\mathcal{N}(n)\log n) = \Theta(n)$ per update (note $\Theta(\log n)$ nodes receive an ADV packet in one transmission). Thus, the overall energy dissipation in one route maintenance epoch is $\Theta(nE_r + nE_t/\log n) = \Theta(n)^\S$ per update. Note that instead of every node maintaining the next hop information, the location of the cell containing a wire-end can be included in the ADV packet. With this location information the nodes can use geographic routing, which is generally more robust to node failures, to reach the nearest wire-end.

How frequently the ADV packets should be sent would, in general, depend on the node mobility process as well as on the lifetime of the sensor nodes. If the sensor nodes are static and there are no node failures, then it would suffice to send the ADV packets only once. However, in dynamic network conditions the overhead of maintaining the routing information can be significant. There are two possible ways of eliminating this periodic route maintenance overhead:

1. *One-time flooding:* All wires flood their location information in the entire network, and each sensor node keeps track of this information. Since there are a total of $\mathcal{N}(n)l(n)$ wires, the cost of this one-time flooding would be $\Theta(n^2l(n)/\log n)$. One-time flooding can totally eliminate the route maintenance overhead. However, note that the cost of one-time flooding can be significant when there are large of wires in the net-

---

§Note that in a practical setting, an ADV packet would be much smaller in comparison to a DATA packet, and therefore, the energy spent in transmission/reception of an ADV packet would be much smaller than the energy spent in transmission/reception of a DATA packet.
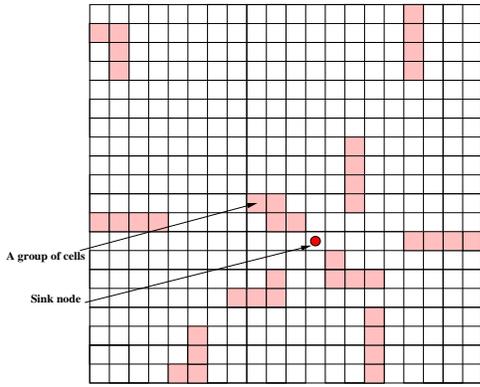
Figure 5: Disjoint groups of cells for $l(n) = 1/16$. Each group consists of a cell, along with the first $1/4l(n) - 1$ nodes on a shortest path from the cell to the sink node.

work, and moreover each node must keep track of the location of all the wires at all times.

2. *Greedy geographic routing:* Each node directly routes its packets to the sink node using the greedy geographic routing. In this case, the nodes only need to know the location of the sink node. Since the sink node is static, it is enough to flood the location information only once, with an overhead of $\Theta(n)$.

Clearly, the routing overhead under greedy geographical routing is much smaller than RMP1 or one-time flooding, and more over it does not require each node to keep track of the location of all the wires, as in one-time flooding. Next, we look at the performance of greedy geographic routing in terms of the average hop count and the EDS.

PROPOSITION 4. *If each node uses greedy geographic routing to reach the sink node, then under all wiring schemes which use at most $\mathcal{N}(n)l(n)$ wires, for $l(n) = \Omega(\sqrt{\log n/n})$, the average hop count must be $\Omega(1/l(n))$.*

PROOF. We only provide a proof for $l(n) = \omega(\sqrt{\log n/n})$. The proof for the case when $l(n) = \Theta(\sqrt{n/\log n})$ can be carried out in a similar fashion, and is omitted. Clearly, if $l(n) = \Theta(1)$, then we have nothing to prove since the average hop count cannot be any smaller than $\Theta(1)$. So suppose $l(n) = o(1)$. Consider disjoint groups of cells; each group consisting of a cell at a lattice distance $> 1/4l(n)$ from the cell containing the sink node, along with the first $\frac{1}{4l(n)} - 1$[¶] cells which are on a shortest path[‖] from the cell to the sink node (see Figure 5). For large n, there are close to $4\mathcal{N}(n)l(n) > 3\mathcal{N}(n)l(n)$ such groups. Since there are less than $\mathcal{N}(n)l(n)$ wires in the network, at least in $2\mathcal{N}(n)l(n)$ groups there are no cells having a wired connection. Observe that the average number of hops traversed by a packet

---

[¶]We are assuming that $1/4l(n)$ is an integer; if not, the proof can be carried out with $\lfloor 1/4l(n) \rfloor$ in place of $1/4l(n)$.

[‖]By shortest path we mean a path which minimizes the hop count when there are no wires in the network. Note that there could be more than one shortest path from a cell to the sink node.

originating in one of these groups is more than $1/8l(n)$, and therefore the average hop count must at least be

$$2\mathcal{N}(n)l(n) \cdot \frac{1}{4l(n)} \cdot \frac{1}{8l(n)} \cdot \frac{1}{\mathcal{N}(n)} = \frac{1}{16l(n)},$$

proving the claim. $\square$

The above result shows that there is a price to be paid for reducing the routing overhead in that the average hop count becomes $\Omega(1/l(n))$ instead of $\Omega(1/\sqrt{l(n)})$. It is easy to see that if the wires are placed as in Scheme 1, and greedy geographic routing is used by all the nodes, then the average hop count as well as EDS are of same order as in a wireless sensor network. Thus, there is almost no benefit of adding the wires in this case. Intuitively, this happens because the packets originating at "most" of the cells end up following the same path as in the case when there were no wires in the network. Next, we consider a simple probabilistic wire placement scheme which works well with the greedy geographic routing.

4.2 Scheme 2

We now consider a simple probabilistic wiring scheme that achieves the lower bound on the average hop count, given in Proposition 4. For technical convenience, we will consider a slightly different constraint on the total number of wires. In particular, the constraint would be to have at most $l(n) < 1$ wires per cell on an average, with the average being taken across all possible realizations of the wiring scheme. The scheme is the following:

1. *Wire Placement*: For each cell in the network, connect it with the cell containing the sink node with a probability $l(n)$[**], independent of the other cells.

2. *Routing:* Each node uses greedy geographic routing to deliver its packets to the sink node.

We will now determine what the average hop count and the EDS are under the above scheme.

PROPOSITION 5. *For any $l(n) = \omega(\sqrt{\log n/n})$, the average hop count and the EDS under Scheme 2 are $\Theta(1/l(n))$ and $\Theta(1/l^2(n))$, respectively.*

PROOF. The above results are trivial if $l(n) = \Theta(1)$; so suppose $l(n) = o(1)$. Let us first look at the average hop count. Consider an arbitrary node in the network, say node o. Let c_o be the cell containing the node o, and let c_s be the cell containing the sink node. Let N_o be the random variable representing the number of hops required by a packet to reach the sink node starting from the node o. Also, let $(\Omega_o, \mathcal{F}, \mathbb{P})$ be the probability space on which the events associated with random distribution of the wires and the nodes are defined, and let $\mathbb{E}$ denote the expectation under $\mathbb{P}$. Observe that $\mathbb{P}(N_o > i)$ is smaller than the probability that the first i nodes encountered by a packet do not have a wired connection with the sink node, that is, $P(N_o > i) \leq (1 - l(n))^i$. Therefore, we have

$$\mathbb{E}[N_o] = \sum_{i=0}^{\infty} \mathbb{P}(N_o > i) \leq \sum_{i=0}^{\infty} (1 - l(n))^i = 1/l(n).$$

[**]Note that by choosing this connection probability to be $cl(n)$, for some appropriate $c < 1$, it is possible to ensure that the number of wires placed be *almost surely* smaller than $\mathcal{N}(n)l(n)$, for large enough n.

Since the choice of node o was arbitrary, we have proved the upper bound on the average hop count. Now for the lower bound, observe that for any arbitrary position of the sink node, there are at least $\mathcal{N}(n)/2$ (actually, close to $\mathcal{N}(n)$) cells at a lattice distance more than $1/l(n)$ from c_s. Further, a packet originating in one such cell, takes at least $1/l(n)$ hops to reach the sink node, provided the first $1/l(n)$ cells on the path followed by this packet to-wards the sink node do not have a wired connection with c_s; an event of probability $(1 - l(n))^{1/l(n)}$. Thus, the average hop count must at least be

$$\frac{1}{2} \cdot \frac{1}{l(n)}(1 - l(n))^{1/l(n)} \geq \frac{1}{8l(n)},$$

for large enough n. Now, let us look at the EDS. We will first establish the upper bound. Consider any arbitrary cell in the network, say cell $c_o \neq c_s$. Note that the probability that a packet generated inside a cell which is at a lattice distance of i from c_o goes via c_o in order to reach the sink node is less than $(1 - l(n))^i$. Since there are at most $4i$ cells at a lattice distance of i from c_o, the average number of cells served by c_o is no more than

$$\sum_{i=0}^{\infty} 4i \cdot (1 - l(n))^i \leq \frac{4}{l^2(n)}$$

Thus, the EDS is no more than $4/l^2(n)$. For the lower bound, observe that a packet originating inside a cell which is at a lattice distance of i from c_s must pass through one of the cells adjacent to c_o, provided the intermediate cells do not have a wired connection with c_s; an event of probability $(1 - l(n))^{i-1}$. Since for any arbitrary position of the sink node, there are at least i cells at a lattice distance of i from c_o, for $i \leq \lceil 1/l(n) \rceil$[††], the average number of cells served by the cells adjacent to c_s is no less than

$$\sum_{i=0}^{\lceil 1/l(n)\rceil} i \cdot (1 - l(n))^{i-1} \geq \sum_{i=\lceil 1/2l(n)\rceil}^{\lceil 1/l(n)\rceil} \frac{1}{2l(n)} \cdot \frac{1}{4} \geq \frac{1}{16l^2(n)},$$

for large n. Noting that there are at most four cells adjacent to c_s, it follows that at least one cell in the network must be serving no less than $\frac{1}{64l^2(n)}$ cells. Also, since the nodes inside c_s are not involved in any forwarding, it follows that the EDS must be $\Omega(1/l^2(n))$. $\square$

REMARK 3. *We note that there is a gap between the lower bound on the EDS given in Proposition 3 ($\Theta(1/l(n))$), and the EDS under the above scheme ($\Theta(1/l^2(n))$). We are unable to close this gap so far. We suspect that a lower bound of $\Theta(1/l^2(n))$ on the EDS might hold under the greedy geographic routing.*

5. HYBRID SENSOR NETWORKS: MOBILE SINK

We now study the energy dissipation in a hybrid sensor network, with a mobile sink node. The setting is almost the same as in the earlier section, with the difference being that the sink node is now mobile. Once again, we have a budget of $l(n) < 1$ wires per cell, with $l(n) = \omega(\log n/n)$, and the objective is to reduce the average hop count and the EDS as much as possible. However, a little amount of

---

[††]Here, we are using $l(n) = \omega(\sqrt{\log n/n})$.

reflection shows that with a mobile sink node, we can no longer control the (short-term) EDS. To be more precise, the EDS would now depend on the motion of the sink node. The reason for this is the following: since we have a budget of $l(n) < 1$ wires per cell, at least $1 - l(n)$ fraction of cells have no wired connections. Now if the sink node happens to be in one of these cells, then the entire traffic must pass through the four neighboring cells, which contain a total of $\Theta(\log n)$ nodes *almost surely*. Since the nodes, which are in the same cell as the sink node are not involved in any forwarding, the EDS would be $\Theta(n/\log n)$ *almost surely*. Note that this is the same as in the case of a wireless sensor network. However, it is possible that the motion of the sink node might result in some sort of equalization in the energy dissipation of the sensor nodes over time. Since we do not assume any knowledge of the motion of the sink node, we will not consider the EDS any further in this section. Our focus, therefore, would be on minimizing the average hop count and the routing overhead as much as possible.

In view of Proposition 3, we know that we cannot make the average hop count any smaller than $\Theta(1/\sqrt{l(n)})$. We now consider a wiring and routing scheme, which is a simple modification of the Scheme 1 in the previous Section, and achieves this lower bound on the average hop count.

5.1 Scheme 3

In this scheme, the wires are placed in a deterministic manner. The scheme is the following:

- *Wire Placement*: Group the cells into clusters consisting of $1/l(n)$ cells each, as in the Scheme 1 (see Figure 4). Choose any arbitrary cluster in the network. Let c_o be the cell in the middle of this cluster. Connect the middle cell of each cluster with cell c_o.

- *Routing*: All cells first route their packet to the nearest cell having a wire-end (using wireless transmissions), and from thereon the packet is forwarded to the cell c_o, using the wired connection. From the cell c_o, the packet is forwarded to the cluster containing the sink node (using the wired connection), and finally, the packet is delivered to the sink node using wireless transmissions.

Clearly, the average hop count under the above scheme is $\Theta(1/\sqrt{l(n)})$. Note that the above routing scheme requires that:

- All the cells know how to reach the closest wire-end.

- The transceivers attached to the wire that connects c_o with the cluster containing the sink node know how to reach the sink node.

We now discuss one possible way of maintaining this information.

5.1.1 Route Maintenance Protocol 2 (RMP2)

In this protocol, in addition to the wires sending out ADV packets periodically, the sink node also sends a HELLO packet periodically. In addition to the *hop limit* (which is set to $2/\sqrt{l(n)}$, in this case), the HELLO packet also contains a parameter called the LOC. This parameter is initially set to some arbitrary value by the sink node, when transmitting within its own cell. One of the nodes who receives this packet

(in the same cell as the sink node), broadcasts this packet to its neighbors with the LOC parameter set to the location (or the identity) of the wire-end it communicates with. Note that in order for the nodes to know this, we also need a LOC field in the ADV packets, which can be set appropriately by the wire-end. Any node which receives a HELLO packet, would broadcast the packet to its nearest neighbors if, and only if, the hop limit has not been reached, and the LOC parameter in the HELLO packet matches with the LOC parameter of the wire-end it communicates with. This will not only ensure that there are no redundant broadcasts of a HELLO packet, but also ensure that only the wire-end which is in the same cluster as the sink node receives the HELLO packet. This wire-end can simply send the information to the other wire-end. Note that in this route maintenance scheme, in addition to the $\Theta(n)$ amount of overhead for forwarding of the ADV packets, there is also an additional $\Theta(\log n/l(n))$ amount of overhead involved in forwarding of the HELLO packet. However, since $l(n) = \omega(\log n/n)$, the overall overhead is still $\Theta(n)$.

Note that only one-time sending of the ADV packets would suffice if the sensor nodes are immobile, reducing the periodic route maintenance overhead significantly. However, in highly dynamic network conditions, where the sensor nodes are mobile and fail frequently, the overhead involved in the above route maintenance scheme can be substantial. We will now consider a simple wiring and routing scheme, which incurs an overhead that is dependent only on the motion of the sink node, and not on the motion of the sensor nodes or the failure rate of the sensor nodes.

5.2 Scheme 4

In this section, we consider a simple probabilistic wiring scheme which is inspired by the work in [15]. For technical convenience, we will consider a slightly different constraint on the total number of wires. In particular, as in Scheme 2, the constraint would be to have $l(n)$ wires per cell on an average, with the average being taken across all possible realizations of the wiring scheme. Also, for the analysis, we assume the wires to be simplex, that is, the information can only flow in one direction inside a wire. Thus, the statement: *cell a has a wired connection with cell b*, would mean that *cell a can send data to cell b through this wired connection*, but not vice-versa. Note that this assumption will not change the order of the average hop count, which is what we are interested in. We are now ready to describe our scheme.

1. *Wire Placement*: Consider an arbitrary pair of cells, say cells u and v. Connect the cell u to cell v using a wire, with a probability

$$\mathbb{P}(u,v) = l(n)\frac{d(u,v)^{-\kappa}}{\sum_{v \neq u} d(u,v)^{-\kappa}}, \qquad (8)$$

where $d(u,v)$ is the Euclidean distance between the centers of the cells u and v, and κ is a constant, to be determined later.

2. *Routing:* Each node uses greedy geographic routing to deliver its packets to the sink node. Note that the frequency with which the location updates need to be sent, depend only on the motion of the sink node; and not on the motion of the sensor nodes or the failure rate of the sensor nodes. Since the sink node is the one initiating these location updates, the frequency of these updates can be controlled precisely by the sink node.

Since the location of the sink node needs to be flooded across the entire network, the overhead per location update is $\Theta(n)$, which is the same as in RMP2. However, in situations, where the sink node moves on a much slower pace than the sensor nodes or if the sensor nodes fail very frequently, one would expect that the overhead under this routing scheme to be much smaller than in the case of RMP2.

We now study the average hop count under the above scheme for different values of κ. We directly state the result in this paper; the proof is available in the longer version of this paper [30].

PROPOSITION 6. *Let $\mathbb{E}[\mathcal{H}]$ denote the average hop count under Scheme 4, and let*

$$\eta = \inf\left\{x : \lim_{n \to \infty}\frac{l(n)}{n^x} = 0\right\},$$

then

1. *for $\kappa = 2$: $\mathbb{E}[\mathcal{H}]$ is $O(\min\{(\log^2 n)/l(n), \sqrt{n/\log n}\})$ and $\Omega(\min\{1/l(n), \sqrt{n/\log n}\})$;*

2. *for $\kappa < 2$: $\mathbb{E}[\mathcal{H}] = \Omega(n^\delta)$, for all $\delta < (2 - \kappa - 2\eta)/(6 - 2\kappa)$;*

3. *and, for $\kappa > 2$: $\mathbb{E}[\mathcal{H}] = \Omega(n^\delta)$, for all $\delta < (\kappa - 2 - 2\eta)/(2\kappa - 2)$.*

REMARK 4. *We note that Proposition 6 extends the results in [15], where similar results were derived for $l(n) = 1$. Also with $l(n) = 1$, the lower bound stated in Proposition 6 for $\kappa < 2$ is tighter than the corresponding lower bound in [15]. More precisely, arguing as in [15], one obtains a lower bound of $\Theta\left((n/\log n)^{\frac{2-\kappa}{6}}\right)$ on the average hop count, which is weaker than the bound we obtain for all κ between 0 and 2. The reason for this is that we use a more refined argument than the one used in [15].*

It is easy to see that the average hop count is minimized when $\kappa = 2$. To see this, let us consider $l(n)$ of the form n^{-c}, for some $0 < c < 1/2$. Then, for $\kappa = 2$ the average hop count is $O(n^c \log^2 n)$. Now for $\kappa < 2$, we have $(2 - \kappa - 2\eta)/(6 - 2\kappa) > c$, and therefore, we can choose a δ lying between c and $(2 - \kappa - 2\eta)/(6 - 2\kappa)$, and the average hop count would be $\Omega(n^\delta)$. In particular, it would be $\omega(n^c \log^2 n)$. Now, let us consider $\kappa > 2$. In this case, $(\kappa - 2 - 2\eta)/(2\kappa - 2) > c$, and a similar argument as above shows that the average hop count under this case is $\omega(n^c \log^2 n)$ as well. Thus $\kappa = 2$ is indeed the best case.

To find out how much "penalty" we might incur on choosing a "bad" value of κ, let us consider $\kappa = 0$. Note that, in this case any two nodes are connected with a probability that does not depend on the distance between the nodes. Using Proposition 6, we see that the average hop count for $\kappa = 0$ is close to $\Theta(n^{c+1}/3)$. Thus even with $c = 0$, the average hop count under $\kappa = 0$ is close to $\Theta(n^{1/3})$; whereas, with $\kappa = 2$ it is $O(\log^2 n)$!

The main intuition behind the above results is that when κ is small, majority of short cuts connect "faraway" nodes, and therefore, most of the packets end up traveling a large

number of hops toward the end of their delivery to the sink node; when the parameter κ is large, majority of the short cuts connect "nearby" nodes, and therefore, most of the packets end up using a large number of hops in order to reach the sink node; and for $\kappa = 2$, it so happens that there is a nice balance between the number of short cuts connecting "nearby" and "faraway" nodes, resulting in the smallest average hop count.

Similar results have been shown to hold under a continuum setting, where the number of nodes as well as the number of connections (links) per node are unbounded (see [31]).

REMARK 5. *We note that it is possible to come up with a deterministic scheme that can achieve an average hop count of $\Theta(1/l(n))$ under the greedy geographic routing. The purpose of presenting the above scheme is not just to show that the average hop count close to $\Theta(1/l(n))$ can be achieved under the greedy geographic routing, but also to provide some insights into what the average hop count would likely be when the wires are placed in some random fashion. Of particular importance is the case when one simply throws the wires without any planning, thereby connecting a pair of cells with a probability that is independent of the distance between them. This case corresponds to $\kappa = 0$ and, as Proposition 6 shows, it results in a much higher average hop count for the same number of wires, compared to the case when $\kappa = 2$. Thus we see that the wire placement is crucial when the greedy geographic routing is being used.*

6. DISCUSSION

In this section, we provide a summary of our results. The following are our main findings:

- If the sink node is static, one can significantly reduce both, the average hop count and the EDS, with the addition of a few wires. In particular, with the addition of $\Theta(nl(n)/\log n)$ wires, one can reduce the average hop count to $\Theta(1/\sqrt{l(n)})$, and the EDS to $\Theta(1/l(n))$. This, however, is at the cost a periodic route maintenance overhead of $\Theta(n)$ per epoch. The route maintenance overhead can be eliminated, if one is ready to sacrifice a bit on the average hop count and the EDS. In particular, we have shown that when the greedy geographic routing is used, the average hop count is at least $\Theta(1/l(n))$. We suspect that a lower bound of $\Theta(1/l^2(n))$ on EDS might hold in this case.

- If the sink node is mobile, it is not possible to guarantee any reduction in the (short-term) EDS when the budget only allows one to place less than one wire per cell on average. (Having at least one wire per cell, does not appear to be an economically feasible option.) The average hop count can be reduced to $\Theta(1/\sqrt{l(n)})$ with the addition of $\Theta(nl(n)/\log n)$ wires, even in this case. The periodic route maintenance overhead cannot be completely eliminated in this case. The reason being that even with the greedy geographic routing scheme, any changes in the location of the sink node must be conveyed to all nodes in the network. Nevertheless, the greedy geographic routing can help reduce the overhead significantly under dynamic network conditions. Once again, the price one has to pay for reducing the overhead is that the average hop count becomes $\Omega(1/l(n))$ instead of $\Omega(1/\sqrt{l(n)})$.

- If greedy geographical routing is being used along with a random placement of wires, then it is best to place the wires in a way so that the probability that a pair of cells is connected with a wire is inversely proportional to the square of the distance between the cells.

- If the budget allows placing less than $\sqrt{n/\log n}$ wires, that is, $l(n) = O(\sqrt{\log n/n})$, then it is impossible to reduce the average hop count when the greedy geographic routing is being used. Thus, in order for the schemes that use the greedy geographical routing to be useful, the investment in the infrastructure should be high enough.

Since both our probabilistic wire placement schemes achieve an average hop count of $\Omega(1/l(n))$, the reader might think that no probabilistic wire placement scheme can achieve an average hop count close to $\Theta(1/\sqrt{l(n)})$. This is, however, not the case. We used the probabilistic wire placement schemes along with the greedy geographical routing and, we believe, it is the greedy geographical routing which has resulted in the average hop count being $\Omega(1/l(n))$ under both our probabilistic wire placement schemes. It is definitely possible to achieve an average hop count close to $\Theta(1/\sqrt{l(n)})$ even with a probabilistic wire placement scheme. For example, with a completely random placement of wires, that is, with $\kappa = 0$, one can achieve an average hop count of $\Theta(\log n/\sqrt{l(n)})^{\ddagger\ddagger}$, provided the packets are routed optimally as in Schemes 1 and 3. We have not presented these results here due to space constraints.

7. CONCLUDING REMARKS

We studied the average energy dissipation and the non-uniformity in the energy dissipation across the sensor nodes in wireless, as well as hybrid, sensor networks. We defined a meaningful notion of EDS to study the non-uniformity in the energy dissipation across the nodes.

In the static sink node case, our results show that a few wires when placed in an appropriate fashion, along with the optimal routing of packets, can significantly reduce both, the energy dissipation in the network as well as the EDS. In the mobile sink node case, we showed that the advantage of placing the wired short cuts is limited to reducing the average hop count of the network, and there is not much that one can do about the EDS in this case. We also showed that the greedy geographic routing, although capable of reducing the routing overhead significantly, provides smaller reductions in the average hop count and (most likely) the EDS, in comparison to the optimal routing. Thus, there is a price to be paid for reducing the routing overhead.

We believe that there is a need to understand the fundamental trade-off between the infrastructure cost and the network lifetime in hybrid sensor networks, and we see this paper as the first step in that direction. Indeed, one would expect the notions of the average hop count and EDS that we study in this paper, to play an important role in determining the network lifetime. We plan to address these issues in our future work.

There are several avenues for future research in this area. We considered a homogeneous sensor network with no data aggregation and clustering; it will be interesting to study

---

$\ddagger\ddagger$See [32], for a related result on the diameter of the small world networks.

other scenarios where the nodes are heterogeneous, possibly arranged into multiple hierarchies, and data aggregation is used to reduce the communication overhead. We believe that our results can provide useful guidelines for placing the wires even in such scenarios.

ACKNOWLEDGMENTS

8. REFERENCES

[1] G. Asada et al. Low power wireless communication and signal processing circuits for distributed microsensors. In *Proceedings of IEEE International Symposium on Circuits and Systems: Circuits and Systems in the Information*, 1997.

[2] A. P. Chandrakasan et al. Power aware wireless microsensor systems. Keynote Paper ESSCIRC, Florence, Italy, Sep 2002.

[3] W. Ye, J. Heidemann, and D. Estrin. An energy-efficient mac protocol for wireless sensor networks. In *IEEE INFOCOM*, June 2002.

[4] S. Singh and C. S. Raghavendra. Pamas - power aware multi-access protocol with signalling for ad hoc networks. In *ACM Computer Communications Review*, 1999.

[5] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocols for wireless microsensor networks. In *Proc. Hawaaian Int'l Conf. on Systems Science*, Jan 2000.

[6] W. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. In *Proc. 5th ACM/IEEE Mobicom Conference, Seattle, WA*, Aug 1999.

[7] A. Mainwaring et. al. Wireless sensor networks for habitat monitoring. In *First ACM Workshop on Wireless Sensor Networks and Applications, Atlanta, GA, USA*, Sep 2002.

[8] A. Cerpa et. al. Habitat monitoring: Application driver for wireless communications technology. In *ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean, Costa Rica*, April 2001.

[9] A. Kansal et al. Sensing uncertainty reduction using low complexity actuation. Technical Report, Center for Embedded Networked Sensing (CENS), UCLA, Feb 2004.

[10] M. Rahimi et. al. Adaptive sampling for environmental robotics. Technical Report, Center for Embedded Networked Sensing (CENS), UCLA, Nov 2003.

[11] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. In *Wireless Communication and Mobile Computing: Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2002.

[12] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Trans. on Information Theory*, IT-46(2):388–404, March 2000.

[13] D. J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, pages 440–442, 1998.

[14] A. Helmy. Small worlds in wireless networks. *IEEE Communications Letters*, 7(10):490–492, Oct 2003.

[15] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proc. 32nd ACM Symposium on Theory of Computing*, 2000.

[16] R. Chitradurga and A. Helmy. Analysis of wired short cuts in wireless sensor networks. In *IEEE/ACS International Conference on Pervasive Services*, July 2004.

[17] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *ACM MOBICOM*, Aug 2000.

[18] S. Bandyopadhyay and E. J. Coyle. Minimizing communication costs in hierarchically-clustered networks of wireless sensors. *Computer Networks*, 44(1):1–16, 2004.

[19] V. Mhatre, C. Rosenberg, D. Kofman, R. Mazumdar, and N. B. Shroff. A minimum cost heterogeneous sensor network with a lifetime constraint. *IEEE Trans. Mob. Comput.*, 4(1):4–15, 2005.

[20] V. Mhatre and C. Rosenberg. Design guidelines for wireless sensor networks: Communication, clustering, and aggregation. *Ad Hoc Networks Journal, Elsevier Science*, 2:45–63, 2004.

[21] B. Liu, Z. Liu, and D. Towsley. On the capacity of hybrid wireless networks. In *IEEE INFOCOM*, 2003.

[22] S.R. Kulkarni and P. Viswanath. Throughput scaling in heterogeneous networks. In *IEEE ISIT*, 2003.

[23] M. Yarvis et al. Exploiting heterogeneity in sensor networks. In *IEEE INFOCOM*, March 2005.

[24] P. Bose et al. Routing with guaranteed delivery in ad hoc wireless networks. In *3rd ACM Int. Workshop on Discrete Algorithms and Methods for Mobile Comp. and Comm.*, Aug 1999.

[25] B. Karp and H.T. Kung. Gpsr: Greedy perimeter stateless routing for wireless networks. In *ACM MOBICOM*, Aug 2000.

[26] A. Rao et al. Geographic routing without location information. In *ACM MOBICOM, San Diego, California, USA*, Sep 2003.

[27] R. Motwani and P. Raghavan. *Randomized algorithms.* Cambridge Univ. Press, 1995.

[28] S.R. Kulkarni and P. Viswanath. A deterministic approach to throughput scaling in wireless networks. *IEEE Trans. on Information Theory*, 50(6):1041–1049, June 2004.

[29] P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity. In *Proc. IEEE 37th Conf. on Decision and Control*, 1998.

[30] G. Sharma and R. Mazumdar. A case for hybrid sensor networks. *Available at: http://ece.purdue.edu/~gsharma.*

[31] M. Franceschetti and R. Meester. Navigation in small world networks, a scale-free continuum model. Preprint, Dept. of ECE, University of California, San Diego, 2004.

[32] A.D. Barbour and G. Reinert. Small worlds. *Random Structures and algorithms*, 19:54–74, 2001.