

2003

Energy-Efficient Routing Schemes for Sensor Networks

Maleq Khan

Gopal Pandurangan

Bharat Bhargava

Purdue University, bb@cs.purdue.edu

Report Number:

03-013

Khan, Maleq; Pandurangan, Gopal; and Bhargava, Bharat, "Energy-Efficient Routing Schemes for Sensor Networks" (2003). *Computer Science Technical Reports*. Paper 1562.

<http://docs.lib.purdue.edu/cstech/1562>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

**ENERGY-EFFICIENT ROUTING SCHEMES
FOR SENSOR NETWORKS**

**Maleq Khan
Gopal Pandurangan
Bharat Bhargava**

**Department of Computer Sciences
Purdue University
West Lafayette, IN 47907**

**CSD TR #03-013
May 2003**

Energy-Efficient Routing Schemes for Sensor Networks*

Maleq Khan, Gopal Pandurangan, and Bharat Bhargava

Department of Computer Science,
Purdue University, West Lafayette, IN 47907
{mmkhan, gopal, bb}@cs.purdue.edu

Abstract

Microsensors operate under severe energy constraints. The sensor nodes need to be deployed in large numbers without any pre-configuration. The nodes are thrown in the area of interest randomly and exact position of each node cannot be pre-specified. However, in some cases, manually placing the nodes in pre-specified position can be possible. In this paper, we analyze energy consumptions, optimal network configuration, and data-centric routing schemes to minimize energy consumption for both random and manual placement of the nodes. We show that in a linear network, energy consumption is minimal when nodes are equally spaced. For a two dimensional network, energy consumptions for various manual uniform arrangements such as triangular, square, and hexagonal array of sensors are analyzed and compared. We show that minimal spanning tree is the optimal data aggregation tree when nodes are randomly distributed. To minimize energy consumption in network configuration phase, an algorithm, called nearest neighbor tree, that approximates minimal spanning tree, is developed. The simulation results show that it is a 2-approximation algorithm.

Keywords. Sensor Networks, Data-Centric Routing, Energy-Efficient Routing, Self-Configuration.

* This research was supported by CERIAS, and NSF Grants CCR-0001788 and EIA-0103676.

1 Introduction

Advances in integrated circuit technology have enabled mass production of tiny, cost-effective, and energy-efficient wireless sensor devices with on-board processing capabilities. The emergence of mobile and pervasive computing has created new applications for them. Sensor-based applications span a wide range of areas, including remote monitoring of seismic activities, environmental factors (e.g., air, water, soil, wind, chemicals), condition-based maintenance, smart spaces, military surveillance, precision agriculture, transportation, factory instrumentation, and inventory tracking [13, 14].

A *microsensor* is a device which is equipped with a sensor module (e.g., an acoustic, a seismic, or an image sensor) capable of sensing some entity in the environment, a digital unit for processing the signals from the sensors and performing network protocol functions, a radio module for communication, and a battery to provide energy for its operation [14]. Currently, microsensors typically consist of 8-bit 4-MHz processors (80% of all microprocessors shipped in 2000 were 8-bit [15]), with slow 10-Kbps communication, an 8-Kbyte read-only program memory, and a 512-byte RAM [16]. These parameters ensure limited weight, size, and cost. We use the term *sensor* to refer to a microsensor.

There are some similarities between wireless sensor networks and wireless ad-hoc networks. One of the similar characteristics for both of them is multi-hop communications. Some of the routing protocols proposed for wireless ad-hoc networks can be examined in the context of wireless sensor networks. Especially the power-aware routing protocols proposed in [1, 2, 3, 4] can be explored since sensor networks have stringent energy constraint. However, these protocols may not be efficient, effective or feasible, in sensor networks. The nature of applications and routing requirements for the two are significantly different in several aspects [5]. First, the typical mode of communication in a sensor network is from multiple data sources to a data recipient/sink rather than communication between any pair of nodes. Second, since the data being collected by multiple sensors is based on common phenomena, there is likely to be some redundancy in the data being communicated by the various sources in sensor networks. Third, in most envisioned scenarios the sensors are not mobile, so the nature of the dynamics in the two networks is different. Finally, the single major resource constraint in sensor networks is that of energy. The situation is much worse than in traditional wireless networks, where the communicating devices handled by human users can be replaced or recharged relatively often. The scale of sensor networks and the necessity of unattended operation for months at a time mean that energy resources have to be managed even more carefully. This, in turn, precludes really high data rate communication and demands energy-efficient routing protocols.

Aggregating data enroute, known as data-centric routing, provides a way to reduce volume of data that saves transmission energy. In this paper, we derive optimal and suboptimal data-centric routing protocols by analyzing energy consumption for various network configurations. The rest of the paper is organized as follow. Previous works related to routing in sensor networks are briefly discussed in Section 2. Analyses of the energy and node placement for a simple linear network are given in Section 3. Energy analysis and optimal routing schemes in a two dimensional sensor network are presented in Section 4. A self-configuration algorithm for building nearest neighbor tree, which is an approximation of optimal routing scheme, minimal spanning tree, is given in Section 5. Experimental results are presented in Section 6 and the conclusions are in Section 7.

2 Related Works

Data aggregation has been put forward as a particularly useful paradigm for wireless routing in sensor networks [6, 7]. The idea is to combine the data coming from different sources enroute – eliminating redundancy, minimizing the number of transmissions and thus saving energy. This paradigm shifts the focus from the traditional *address-centric* approaches to a more *data-centric* approach [5]. In [5], address-centric and data-centric protocols are defined as follows:

Address-centric Protocol (AC): Each source independently sends data along the shortest path to sink based on the route that the queries took (“end-to-end routing”).

Data-centric Protocol (DC): The sources send data to the sink, but routing nodes enroute look at the content of the data and perform some form of aggregation/consolidation function on the data originating at multiple sources.

The authors theoretically bound the number of transmissions required in DC protocol and showed that DC protocol need fewer transmission that that of AC protocol.

Heinzelman et al [9] presented a simple analysis to show when multi-hop routing is preferable over direct communication with the objective of minimizing energy for transmission.

The above two analyses provide a good theoretical basis for using data-centric multi-hop routing in wireless sensor network. However, a rigorous theoretical model is needed to estimate energy and to find sensor placement strategy to minimize energy. In the next section, we provide a model for such analysis.

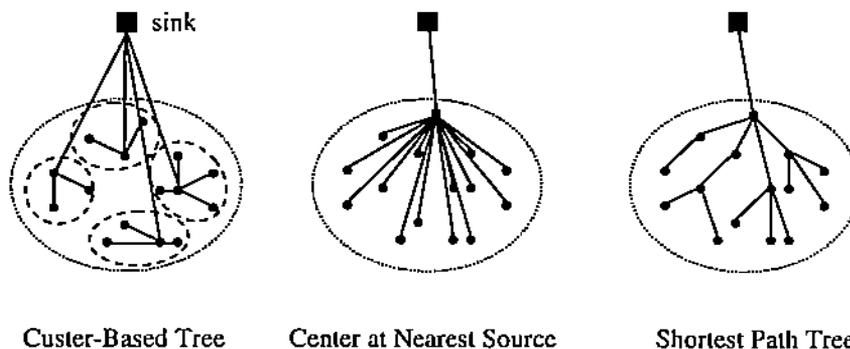


Figure 1. Various schemes for data aggregation and routing. The sink can be inside or outside of the sensing area.

Several schemes have been proposed for data-centric routing in sensor networks. Cluster-based [2, 7], center at nearest source [5], and shortest path tree [5] are important among them.

Cluster-Based Tree (CBT): In this scheme, the sources send data to the associated cluster head. Cluster head aggregate data and send to the sink. Multiple levels of cluster hierarchy [2] can be another option.

Center at Nearest Source (CNS): In this scheme, the source which is nearest the sink acts as the aggregation point. All other sources send their data directly to this source which then sends the aggregated information on to the sink.

Shortest Paths Tree (SPT): In this scheme, each source sends its information to the sink along the shortest path between the two. Where these paths overlap for different sources, they are combined to form the aggregation tree.

SPIN [10], Directed Diffusion [6], GEAR [11], and Rumor Routing [12] are the recently proposed routing protocols for sensor networks to disseminate query and/or data. All of these

algorithms rely on flooding techniques, and use different heuristics to minimize flooding and setup directed paths.

3 Routing in A simple Linear Network

Unattended operations and limited energy of the sensor nodes demands a routing scheme, which minimizes energy consumptions for routing data from the sources to the sink. At this point, several research questions arise: to minimize energy/cost,

- 1) How the sensors should be placed (sensor distribution)?
- 2) How many sensors should be deployed (density)?
- 3) Which routing scheme requires minimum energy?
- 4) What is the expected energy required for a given routing protocol when the sensors are placed randomly?

To answer the above questions, we begin with a simple one-dimensional sensor array shown Figure 2. For the purpose of analysis, a transmission step and a transmission phase are defined as follows.

Definition 1. Transmission Step. A transmission step is the time duration in which a node begins and completes transmitting data to the next hop.

Definition 2. Transmission Phase. A transmission phase a collection of successive transmission steps that begins with the steps when the sources start sending data and ends with the step when the sink receives data from all of the sources.

Considering data-centric routing the following assumptions are made.

Assumption 1. Each node waits until it receives data from all of its descendants and upon receiving, aggregates received data (including its own) then sends to the next hop.

Assumption 2. In each transmission phase, each source sends data exactly once.

Lemma 1. If n be the number of sensors nodes, the total number of transmissions in each transmission phase is n .

Proof: Assumption 1 and 2 lead us to the conclusion that each sensor node needs to transmit exactly once in each transmission phase. That is, the number of total transmission is n . □

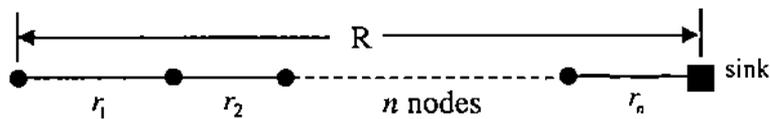


Figure 2. A simple linear network - one-dimensional sensor array.

Theorem 1. Considering a simple linear network, let the sink receive data from a distance R using n hops placed in a straight line. The total transmission energy is minimal when nodes are

equally spaced, i.e. when the distance between any two neighbors is $\frac{R}{n}$, and the minimum energy consumption $E_{\min} = \frac{cR^2}{n}$, where c is the constant energy factor.

Proof: Let r_i be the distance from node i to the next hop (Figure 2). Energy required to transmit to a distance r_i is $E_i = cr_i^m$, where $m \geq 2$ is a constant. Total energy consumption $E = \sum_{i=1}^n cr_i^m$. We minimize $\sum_{i=1}^n cr_i^m$ with the constraint $\sum_{i=1}^n r_i = R$. An equivalent expression to minimize is

$$L = \sum_{i=1}^n cr_i^m - \lambda \left(\sum_{i=1}^n r_i - R \right), \text{ where } \lambda \text{ is a Lagrange's multiplier}$$

$$\text{now, } \frac{\delta L}{\delta r_i} = cmr_i^{m-1} - \lambda = 0, \text{ i.e. } r_i = \left(\frac{\lambda}{cm} \right)^{\frac{1}{m-1}}.$$

$\left(\frac{\lambda}{cm} \right)^{\frac{1}{m-1}}$ is a constant (independent of i), that is $r_1 = r_2 = \dots = r_n = \frac{R}{n}$. Minimum energy

consumption, $E_{\min} = \sum_{i=1}^n cr_i^m = cn \left(\frac{R}{n} \right)^m$. Considering $m = 2$, $E_{\min} = \frac{cR^2}{n}$.

□

Using Theorem 1, we realize that energy consumption is minimal when nodes are equally spaced and observe that transmission energy is a strictly monotone decreasing function of number of nodes. This observation suggests using as much sensor as possible to minimize energy consumption. However, there is a constant (independent of distance) amount of energy requirement for each transmission at the receiving end and also may be at the sending end to processing data (data aggregation and processing data packets). Furthermore, cost of sensor nodes need to be considered to find the optimal number sensor nodes. Theorem 2 gives the optimal number of sensors for a simple linear network.

Theorem 2. To minimize the total cost, the optimal number of sensors nodes in a simple linear

network is $n_{opt} = R \sqrt{\frac{cTC_e}{C_s + TE_c C_e}}$, where C_s and C_e are the costs per sensor and per unit

energy, and c , T , and E_c are the constant energy factor, the estimated number of transmissions phases in the lifetime of the network, and the constant amount of energy associated with per transmission, respectively.

Proof: Total variable transmission cost $= ncr^m \times T \times C_e = \frac{cTC_s R^m}{n^{m-1}}$, using $r = \frac{R}{n}$.

Total fixed transmission cost (independent of transmission distance) $= nE_c TC_e$.

Cost of the sensor nodes $= nC_s$.

$$\text{Total cost } C_t = \frac{cTC_s R^m}{n^{m-1}} + nE_c TC_e + nC_s.$$

$$\frac{dC_t}{dn} = -(m-1) \frac{cTC_s R^m}{n^m} + E_c TC_e + C_s, \text{ solving } \frac{dC_t}{dn} = 0, n_{opt} = \left\{ \frac{(m-1)cTC_s R^m}{E_c TC_e + C_s} \right\}^{\frac{1}{m}}$$

$$\text{With } m = 2, n_{opt} = R \sqrt{\frac{cTC_e}{C_s + TE_c C_e}}.$$

□

Corollary 2.1. Ignoring the cost of the sensor nodes $n_{opt} = R \sqrt{\frac{c}{E_c}}$.

In many applications, we may not be able to position nodes at the exact positions. Instead, we might need to through the sensors randomly in the area of interest. In such a situation, although we are not able to get the minimum energy configuration, we are interested to find the expected energy required for transmission. The following lemmas and theorems give us way to compute expected energy for random distribution of the sensor nodes. An interesting result (Theorem 3 below) to observe is that the expected energy consumption in a linear network with randomly distributed nodes is bounded by $2E_{min}$, twice the energy consumption when the nodes are equally spaced.

Lemma 2. For the simple linear network model, when the nodes are randomly distributed, the probability density function that a node transmit to distance r (i.e. the next hop is at distance r) is

$$P(r) = \frac{n-1}{R} \left(1 - \frac{r}{R}\right)^{n-2}.$$

Proof: Consider any arbitrary node N at point A shown in Figure 3. Assuming uniform distribution, the probability that a particular node is on the line segment AB of length r is $\frac{r}{R}$.

The probability that the next hop N' is within distance r
= the probability that at least one of the $n-1$ nodes (except N) on AB

$$= 1 - \left(1 - \frac{r}{R}\right)^{n-1}$$

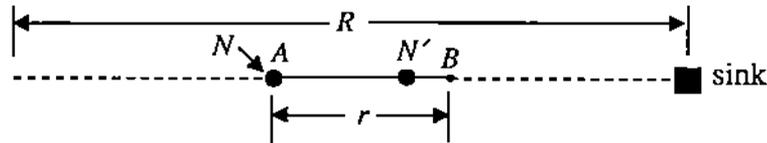


Figure 3. Randomly distributed sensor nodes

This is the cumulative distribution function. The derivative of this function is the probability density function. That is,

$$P(r) = \frac{d}{dr} \left\{ 1 - \left(1 - \frac{r}{R} \right)^{n-1} \right\} = \frac{n-1}{R} \left(1 - \frac{r}{R} \right)^{n-2}$$

□

Note that $\int_0^R P(r)dr = \int_0^R \frac{n-1}{R} \left(1 - \frac{r}{R} \right)^{n-2} dr = 1$ and the expected (average) distance to the next node is $\int_0^R rP(r)dr = \int_0^R r \frac{n-1}{R} \left(1 - \frac{r}{R} \right)^{n-2} dr = \frac{R}{n}$, which are obvious.

Theorem 3: The expected energy consumption in one transmission phase of a simple linear network with randomly distributed n nodes is $E_{\text{exp}} = \frac{2cR^2}{n+1} \leq 2E_{\text{min}}$.

Proof: Energy consumption for one transmission to distance r is cr^2 . Therefore, the expected energy consumption for one transmission is

$$\begin{aligned} & \int_0^R cr^2 P(r)dr \\ &= \int_0^R cr^2 \frac{n-1}{R} \left(1 - \frac{r}{R} \right)^{n-2} dr, \text{ [Lemma 2]} \\ &= \frac{2cR^2}{n(n+1)} \end{aligned}$$

There are n transmissions in one transmission phase [Lemma 1]. Hence, expected energy consumption in one phase is $E_{\text{exp}} = \frac{2cR^2}{n(n+1)} \times n = \frac{2cR^2}{n+1} \leq \frac{2cR^2}{n} = 2E_{\text{min}}$ [Theorem 1].

□

4 Routing in Two Dimensional Networks

In this section we examine network configuration to minimize energy consumption in routing data in a two dimensional network. First we analyze manual configuration of the network, where we are able to fix the nodes in the desired locations, followed by the analysis of the network with randomly distributed nodes.

4.1 Manual Placement of the Sensor Nodes

From the analysis of the simple linear network, we foresee that the energy consumption in a two dimensional network is also minimum when nodes are uniformly spaced. Equilateral triangle, square, and hexagon are such three possible uniform arrangements of the nodes (Figure 4). However, some coverage (the whole region of interest is covered) criterion must be satisfied by the sensor nodes. We define coverage as follow.

Definition 3. Coverage, d . We say, coverage by a set of sensor nodes in a region L is d if d is the minimum distance such that every point in the region L has at least one sensor node within distance d . More formally,

$$\text{Coverage, } d = \max_{p \in L} \{D(p, N_p)\}$$

Where, $D(p, N_p)$ is the distance of the nearest sensor node N_p from point p .

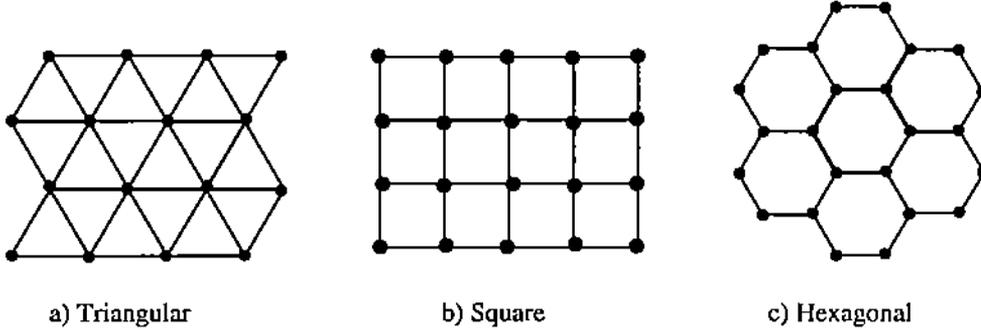


Figure 4. Uniform arrangements of the sensor nodes

Let us examine the triangular arrangement first. We assume that the area A of the region (L) under consideration, hence the required number of nodes, is large.

Let each side of a triangle is r i.e. that each node transmit to distance r to the next hop. Area of one triangle $= \frac{1}{2} r \cdot r \sin \frac{\pi}{3} = \frac{\sqrt{3}}{4} r^2$.

Each node shares 6 triangles (6 triangles meet at one point). Share of a node to one such triangle is $\frac{1}{6}$ node. For each triangle, there are 3 nodes at the 3 vertices. Therefore, the number of nodes per triangle is $\frac{1}{6} \times 3 = \frac{1}{2}$; that is, area per sensor node = area of two triangles $= \frac{\sqrt{3}}{2} r^2$.

$$\text{If there are } n \text{ nodes in area } A, n \frac{\sqrt{3}}{2} r^2 = A \Rightarrow r = \left(\frac{2A}{\sqrt{3}n} \right)^{\frac{1}{2}}.$$

$$\text{Radiation energy for } n \text{ transmissions, } e_n = ncr^m = nc \left(\frac{2A}{\sqrt{3}n} \right)^{\frac{m}{2}} = 2^{\frac{m}{2}} 3^{-\frac{m}{4}} cA^{\frac{m}{2}} n^{-\frac{m}{2}}.$$

The furthest point inside a triangle from its vertices is the centroid of the triangle; that is, the coverage in a triangular arrangement is the distance between a vertex and the centroid as shown in Figure 5.

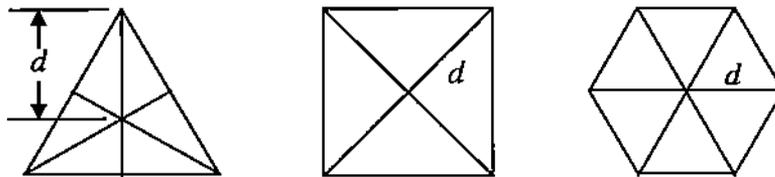


Figure 5. Coverage, d , in triangular, square, and hexagonal arrangement

Height of the triangle, $h = \frac{\sqrt{3}}{2}r$; coverage, $d = \frac{2}{3}h = \frac{1}{\sqrt{3}}r = \left(\frac{2A}{3\sqrt{3}n}\right)^{\frac{1}{2}} = 0.62\sqrt{\frac{A}{n}}$ (1)

In the similar fashions, we can calculate radiation energy and coverage in square and hexagonal arrangements. The values are shown in Table 1.

Table 1. Coverage and radiation energy for n nodes deployed in area A . The last two columns shows energy when m is 2 and 3 respectively.

Arrangement	Coverage, d	Energy, e_m	Energy with $m = 2$	Energy with $m = 3$
Triangular	$0.62\sqrt{\frac{A}{n}}$	$2^{\frac{m}{2}} \times 3^{\frac{m}{4}} cA^{\frac{m}{2}} n^{1-\frac{m}{2}}$	$1.15cA$	$1.24c\sqrt{\frac{A^3}{n}}$
Square	$0.71\sqrt{\frac{A}{n}}$	$cA^{\frac{m}{2}} n^{1-\frac{m}{2}}$	cA	$c\sqrt{\frac{A^3}{n}}$
Hexagonal	$0.88\sqrt{\frac{A}{n}}$	$2^m \times 3^{\frac{3m}{4}} cA^{\frac{m}{2}} n^{1-\frac{m}{2}}$	$0.77cA$	$0.66c\sqrt{\frac{A^3}{n}}$

We see that as the number of sides in the shape increases, energy consumption decreases but coverage becomes poorer. Next we analyze what are the energy consumptions to get a constant coverage in all arrangements. The energy consumptions and required number of nodes to have a coverage d is given in Table 2. The values are simply re-expressed in terms of d using the relationship between d and n (e.g., equation 1 for triangular arrangement).

Table 2. Required number of nodes and energy consumptions to satisfy coverage d . The last two columns shows energy when m is 2 and 3 respectively.

Arrangement	Required n	Energy, e_m	Energy with $m = 2$	Energy with $m = 3$
Triangular	$\frac{2A}{3\sqrt{3}d^2} = 0.38\frac{A}{d^2}$	$2 \times 3^{\frac{m-3}{2}} cAd^{m-2}$	$1.15cA$	$2cAd$
Square	$\frac{A}{2d^2} = 0.50\frac{A}{d^2}$	$2^{\frac{m}{2}-1} cAd^{m-2}$	cA	$1.41cAd$
Hexagonal	$\frac{4A}{3\sqrt{3}d^2} = 0.77\frac{A}{d^2}$	$\frac{4}{3\sqrt{3}} cAd^{m-2}$	$0.77cA$	$0.77cAd$

Keeping the coverage constant, as the number of sides in the shape increases, still radiation energy decreases but required number of nodes increases. Again, total energy consumption in digital and radio electronics increases with number of nodes. That is, radiation energy decreases and electronic energy consumptions increases with the number of sides of the shape. We conclude that there is an optimal arrangement for a specified d ; that is, there are some boundary values and ranges for coverage d , which determine the optimal arrangement to minimize total energy consumption.

Let e_c be the average electronic (digital and radio) energy consumption by a sensor node in one transmission. In one transmission phase (n transmissions), total energy consumptions = electronic energy + radiation energy. For triangular arrangements, total energy consumption,

$$E_{triangle} = ne_c + e_{nl} = \frac{2A}{3\sqrt{3}d^2} e_c + 2 \times 3^{\frac{m-3}{2}} cAd^{m-2}.$$

$$\text{Similarly, } E_{square} = \frac{A}{2d^2} e_c + 2^{\frac{m}{2}-1} cAd^{m-2}$$

$$\text{and } E_{hexagon} = \frac{4A}{3\sqrt{3}d^2} e_c + \frac{4}{3\sqrt{3}} cAd^{m-2}$$

Triangular arrangement is better than square arrangement when, $E_{triangle} < E_{square}$,

$$\text{i.e., } \frac{2A}{3\sqrt{3}d^2} e_c + 2 \times 3^{\frac{m-3}{2}} cAd^{m-2} < \frac{A}{2d^2} e_c + 2^{\frac{m}{2}-1} cAd^{m-2}$$

$$\Leftrightarrow d < \left(\frac{3\sqrt{3}-4}{3^{\frac{m}{2}} 4 - 2^{\frac{m}{2}} 3\sqrt{3}} \cdot \frac{e_c}{c} \right)^{\frac{1}{m}}$$

$$\text{Similarly, } E_{square} < E_{hexagonal}, \text{ when } d < \left(\frac{8-3\sqrt{3}}{2^{\frac{m}{2}} 3\sqrt{3} - 8} \cdot \frac{e_c}{c} \right)^{\frac{1}{m}}.$$

Let us consider a scenario with $m = 3$, $c = 200$ pJ/bit/m³, electronic power consumption = 50 mW, and effective data transmission rate = 10 Kbps. Then $e_c = 25 \times 10^{-3} / 10^4 = 25 \times 10^{-7}$ Joule. $E_{triangle} < E_{square}$ if $d < 13.49$ m and $E_{square} < E_{hexagonal}$ if $d < 17.36$ m. We envision that in most of the cases, desired coverage $d < 13.49$; that is, in those cases, triangular arrangement is optimal.

4.2 Random Distribution of the Sensor Nodes

In this section, we analyze routing and energy consumptions when the sensors are randomly (uniform) distributed in a two dimensional region. For the purpose of analysis, the following definitions and assumptions are made.

Definition 4. Routing Path. A routing path is the path along which a source sends data to the sink.

Definition 5. Routing Tree. The routing paths in a network form a tree when they satisfy the conditions a) a routing path does not contain any cycle and b) if two routing paths merge at some node, they never get separated. This tree is called a routing tree.

Assumption 3. A node can be connected to (can communicate with) the other nodes, which are within a specified distance.

Definition 6. Connectivity Graph. A connectivity graph, $G = (V, E)$, is the graph where V is the set of sensor nodes and for any two nodes u and v , weight of the edge (u, v) , $w(u, v) =$ the distance between u and v if u and v can communicate with each other (e.g., if they are within a specified distance), otherwise $w(u, v) = \infty$.

Theorem 4. Let $G = (V, E)$ be the connectivity graph of the sensor nodes. A routing tree with minimum energy consumption is a minimum spanning tree on G .

Proof: Let $w(u, v)$ be the weight of the edge (u, v) in G . Energy required for one transmission from u to v is $cw^2(u, v)$. Let G' be the graph with the same vertices and edges as in G but weight for edge (u, v) , $w'(u, v) = \infty$ if $w(u, v) = \infty$, otherwise $w'(u, v) = cw^2(u, v)$, i.e., energy required for one transmission from u to v . A minimum spanning tree T' on G' minimizes $\sum_{(u,v) \in T'} w'(u, v)$, which is $\sum_{(u,v) \in T'} cw^2(u, v)$, that is, T' minimizes energy consumption for a transmission phase.

Now we show that for all u and v , $(u, v) \in T'$ if and only if $(u, v) \in T$, where T is the minimum spanning tree on G . Consider Kruskal's algorithm [8] to find minimum spanning tree: the edges are sorted by non-decreasing weight, then, to form tree, edges are added one by one from the sorted list. An edge (u, v) is added to the tree if u and v are not connected using the edges already added. For any two edges (u_1, v_1) and (u_2, v_2) , $w'(u_1, v_1) \geq w'(u_2, v_2) \Leftrightarrow cw^2(u_1, v_1) \geq cw^2(u_2, v_2) \Leftrightarrow w(u_1, v_1) \geq w(u_2, v_2)$, that is, the both set of weights w' and w produce the same sorted order of the edges. As a result, the set of edges in T' is equal to the set of edges in T . Since T' minimizes energy consumption, hence, T does so. \square

The following theorem (Theorem 5) gives a lower bound on expected radiation energy in one transmission phase in a sensor network with randomly distributed nodes in a circular area. That is, there exists no routing scheme for which the expected radiation energy in one transmission phase is less than this lower bound.

Lemma 3. In a two dimensional network model, the probability density function that a node transmit to distance r (i.e. the nearest neighbor is at distance r) is $\frac{(n-1)2r}{R^2} \left(1 - \frac{r^2}{R^2}\right)^{n-2}$.

Proof: A particular node has its nearest neighbor within distance r (Figure 4) with probability $1 - \left(1 - \frac{r^2}{R^2}\right)^{n-1}$. The derivative of this function, $\frac{(n-1)2r}{R^2} \left(1 - \frac{r^2}{R^2}\right)^{n-2}$, is the probability density function. \square

Theorem 5. A lower bound on the Expected energy consumption in one transmission phase is cR^2 , where c is the constant energy factor and R is the radius of the area covered by the sensor nodes.

Proof: Consider each node send to the nearest neighbor. In any routing scheme, a node cannot send data to a closer distance than distance to its nearest neighbor. The probability density function that the nearest neighbor is at distance r is given by

$$P(r) = \frac{(n-1)2r}{R^2} \left(1 - \frac{r^2}{R^2}\right)^{n-2} \quad [\text{Lemma 3}].$$

Expected energy consumption in one transmission is

$$\geq \int_0^R cr^2 P(r) dr = \int_0^R cr^2 \frac{(n-1)2r}{R^2} \left(1 - \frac{r^2}{R^2}\right)^{n-2} dr = \frac{cR^2}{n}$$

Energy consumption in n transmissions $\geq n \frac{cR^2}{n} = cR^2$.

□

5 Self-Configuring Nearest Neighbor Tree

Computing a minimum spanning tree in a distributed fashion takes $O(n \log n)$. message passing among the nodes. If reconfiguration of the tree needs to be done frequently, the configuration overhead becomes significant. In such a situation, an approximation algorithm can serve better. We propose an approximation algorithm called the nearest neighbor routing.

A *nearest neighbor tree* (NNT) is tree where each node is connected to the nearest among its available neighbors. We say a neighbor is available for connection if and only if the neighbor is not connected yet. The detail of the algorithm, how a sensor network self-configures to form a nearest neighbor tree, is given below.

a. Building neighbor list: Each node broadcasts a signal with its ID and creates a list of the IDs of the nodes from which it receives signal. From the strength of the signals received, the distances of the nodes are estimated. Based on this distance, a sorted list of the neighbors is created.

b. Getting connected: Each node set a timer proportional to the remaining energy and number of neighbors. When the timer expires, a node selects it's nearest available node to get connected to. If node p gets connected to node q , p is no longer available for further connection and q is available if it is not connected to any other node in some previous step. A node closer to boundary or in a sparse region has lower timer value, hence gets connected early to avoid dead-end (described below). A node with higher remaining energy gets connected later that allow them to be closer to the root of the tree. A node closer to the root consumes more energy in longer idle listening and data aggregation.

c. Resolving dead end: Consider a scenario that node p is not connected yet and all neighbors of p are connected to some nodes, i.e., there is no available neighbor for p to get connected to; we say p is in *dead end*. If there is only one node is in dead-end, it is the root of the tree. If there are more than one node are in dead-end, those nodes increase their transmission distance, communicate with each other and select the node with highest remaining energy as the root, and other nodes get connected to the root.

6 Simulation Results

The goals of the experiments are to compare the required radiation energy in minimum spanning tree and nearest neighbor tree with those of three cluster-based trees generated by LEACH [9], Localized [17] and LLC [18] algorithms. The theoretical analyses of the random networks are also verified with the simulated data.

6.1 Experimental Setup

200 sensor nodes are randomly (uniform distribution) distributed in a square area 200m×200m and the data-sink (base station) is considered to be at 200m from the center of the sensor field. For the sake of fairness, every measured parameter is computed by averaging 50 different random distributions of the nodes. Considering that the self-configuration process needs to be repeated over the lifespan of the sensor networks and the nodes already lost a portion of their energy, the current energy of a node is randomly selected from the range of 30 to 50 Joule.

Various power and energy related specifications are collected from [9, 19, 20] and in some cases adjusted and normalized to keep consistent with the motes developed by University of California, Berkeley [21]. We compute the radio path loss with an empirical r^3 model. The energy consumption to transmit k bits to distance r is given by $k(E_{tx} + E_{rad}r^3)$, where E_{tx} is the energy consumed by the radio electronics to transmit one bit and E_{rad} is the radio path loss per bit per cubic meter. The simulation parameters are given in Table 3.

Table 3. Simulation Parameters

Parameters	Value
Digital Electronics	11 mW
Radio Receiver Electronics	13.5 mW
Radio Idle Listening	13.5 mW
Radio Trans. Electronics	24.8 mW
Radio Sleep Mode	15 μ W
Radio Path Loss	200 pJ/bit/m ³
Transmission Rate	1 Mbps

6.2 Results

The expected energy consumption in linear network (Theorem 3):

The theoretically calculated value is 7.96 Micro Joule and simulation value is 7.93 (by averaging 10000 repetitions) with $R = 2000$, $n = 200$.

The expected energy consumption in a two 2-d network (Theorem 5):

The theoretically calculated value is 2.0 Micro Joule and simulation value is 2.14 (by averaging 10000 repetitions) with $R = 200$, $n = 200$.

We see the theoretically calculated values are very close to the simulated results.

Radiation Energy for various routing trees:

Radiation energies required to transmit 1-bit of data by all of the source nodes to sink are compared among the routing schemes.

The best possible clustering by the algorithms LEACH, Localized, and LLC consumes 24.34 mJ, 12.42 mJ, and 12.42 mJ of energy respectively while energy consumption in minimum spanning tree (MST) is 1.66 mJ. Energy radiated in cluster-based tree (CBT) is 7.5 times larger than that in MST.

Radiation energy required in nearest neighbor tree (NNT) is 2.70 mJ which is 1.6 times larger than that in MST. We can assume that NNT formed by the given algorithm is 2-approximation to MST. Again, energy consumption in CBT is 4.6 times larger than that in NNT.

7. Conclusions

The analyses and experiments presented in this paper, leads us to the following conclusions. In a simple linear network, energy consumption is minimal when the nodes are equally spaced. In a manually configured two-dimensional sensor network, triangular arrangement is optimal when coverage requirement is very high i.e. every point in the region must have a sensor node within close proximity. It is proven that minimum spanning tree is the optimal data aggregation tree. The proposed approximation algorithm, nearest neighbor routing consumes less energy than twice the required energy in minimum spanning tree.

References

- [1] X. Lin and I. Stojmenovic. Power-Aware Routing in Ad Hoc Wireless Networks. In *SITE, University of Ottawa, TR-98-11*, Dec. 1998.
- [2] T. Meng and R. Volkan. Distributed Network Protocols for Wireless Communication. In *Proc. IEEE ISCAS*, May 1998.
- [3] T. Shepard. A Channel Access Scheme for Large Dense Packet Radio Networks. In *Proc. ACM SIGCOMM*, pages 219–230, Aug. 1996.
- [4] S. Singh, M.Woo, and C. Raghavendra. Power-Aware Routing in Mobile Ad Hoc Networks. In *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)*, Oct. 1998.
- [5] B. Krishnamachari, D. Estrin, and S. Wicker. Modelling Data-Centric Routing in Wireless Sensor Networks. In *Proceedings of IEEE INFOCOM*, 2002.
- [6] J. Heidemann, F. Silva, C. Intanagonwiwat, R. Govindan, D. Estrin, and D. Ganesan, "Building Efficient Wireless Sensor Networks with Low-Level Naming," *18th ACM Symposium on Operating Systems Principles*, October 21-24, 2001.
- [7] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *ACM/IEEE International Conference on Mobile Computing and Networks (MobiCom 2000)*, August 2000, Boston, Massachusetts.
- [8] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. The MIT Press, Cambridge, Massachusetts, ISBN 0-262-03141-8, 1990.
- [9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient Communication Protocol for Wireless Microsensor Networks," in *Proceedings of the 33rd International Conference on System Sciences (HICSS)*, January 2000.
- [10] W. R. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks," in *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99)*, Seattle, Washington, August 1999.
- [11] Y. Yu, R. Govindan, and D. Estrin, "Geographical and Energy Aware Routing: A Recursive Data Dissemination Protocol for Wireless Sensor Networks," *UCLA Computer Science Department Technical Report UCLA/CSD-TR-01-0023*, May 2001.
- [12] D. Braginsky and D. Estrin, "Rumor Routing Algorithm For Sensor Networks," In *Proceedings of the First Workshop on Sensor Networks and Applications (WSNA)*, Atlanta, GA, September 2002.

- [13] N. Bulusu, D. Estrin, L. Girod and J. Heidemann, "Scalable Coordination for Wireless Sensor Networks: Self-Configuring Localization Systems," in *Proceedings of the Sixth International Symposium on Communication Theory and Applications (ISCTA 2001)*, Ambleside, United Kingdom, July 2001.
- [14] W. Heinzelman, "Application-Specific Protocol Architectures for Wireless Networks," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, June 2000.
- [15] D. Tennenhouse, "Embedding the Internet: Proactive Computing", *Communications of the ACM*, 43(5), pp. 43-43, 2000.
- [16] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. D. Tygar, "SPINS: Security protocols for sensor networks," in *Proceedings of the Seventh Annual International Conference on Mobile Computing and Networks (MobiCom)*, pp.189-199, 2001.
- [17] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next Century Challenges: Scalable Coordination in sensor Networks," in *Proceedings of the Fifth Annual International Conference on Mobile Computing and Networks (MobiCom)*, Seattle, WA, August 1999.
- [18] M. Khan, B. Bhargava, S. Agarwal, L. Lilien, "Self-configuring Node clusters, Data Aggregation, and Security in Microsensor Networks," *Technical Report, CSD-TR 03-005*, Department of Computer Science, Purdue University, West Lafayette, IN, March 2003.
- [19] R. Min and A. Chandrakasan, "Energy-Efficient Communication for Ad-Hoc Wireless Sensor Networks", *35th Asilomar Conference on Signals, Systems, and Computers*, vol. 1, November 2001, pp. 139-143.
- [20] W. Ye, John Heidemann, and D. Estrin, "An Energy-Efficient MAC Protocol for Wireless Sensor Networks," in *Proceedings of the 21st International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, New York, NY, USA, June, 2002.
- [21] <http://www.cs.berkeley.edu/~awoo/smartdust>.