

# Coverage-Time Optimization for Clustered Wireless Sensor Networks: A Power-Balancing Approach

Tao Shu and Marwan Krunz

Department of Electrical and Computer Engineering

University of Arizona, Tucson, AZ 85721, USA

Email: {tshu, krunz}@ece.arizona.edu

**Abstract**—In this paper, we investigate the maximization of the coverage time for a clustered wireless sensor network (WSN) by optimal balancing of power consumption among cluster heads (CHs). Clustering significantly reduces the energy consumption of individual sensors, but it also increases the communication burden on CHs. To investigate this tradeoff, our analytical model incorporates both *intra-* and *inter-cluster* traffic. Depending on whether location information is available or not, we consider optimization formulations under both deterministic and stochastic setups, using a Rayleigh fading model for inter-cluster communications. For the deterministic setup, sensor nodes and CHs are arbitrarily placed, but their locations are known. Each CH routes its traffic directly to the sink or relays it through other CHs. We present a coverage-time-optimal joint clustering/routing algorithm, in which the optimal clustering and routing parameters are computed using a linear program. For the stochastic setup, we consider a cone-like sensing region with uniformly distributed sensors, and we provide optimal power allocation strategies that guarantee (in a probabilistic sense) an upper bound on the end-to-end (inter-CH) path reliability. Two mechanisms are proposed for achieving balanced power consumption in the stochastic case: a routing-aware optimal cluster planning and a clustering-aware optimal random relay. For the first mechanism, the problem is formulated as a signomial optimization, which is efficiently solved using generalized geometric programming. For the second mechanism, we show that the problem is solvable in linear time. Numerical examples and simulations are used to validate our analysis and study the performance of the proposed schemes.

**Keywords:** Generalized geometric programming, signomial optimization, linear programming, sensor networks, clustering, topology control, coverage time.

## I. INTRODUCTION

### A. Motivation

The rapid transition to nanoscale ICs has led to the integration of high-performance processors and high-speed digital wireless communication circuits. Coupled with advances in micro-electro-mechanical systems (MEMs), such integration has paved the way for the deployment of dense wireless sensor networks (WSNs). These networks are expected to play an important role in a wide range of civilian and military applications, including environment monitoring, seismic-

structure analysis, marine micro-organisms research, surveillance and reconnaissance, etc. [3]. For harsh, inaccessible deployment scenarios, sensors are necessarily powered by energy-constrained, often non-rechargeable batteries [18]. This makes energy consumption a critical factor in the design of a WSN and calls for energy-efficient communication protocols that maximize the lifetime of the network.

For a large WSN, sensors are often hierarchically organized into *clusters*, each having its own *cluster head* (CH) [18]. Within a cluster, sensors transmit data to their CH, which in turn forwards the data (or a fused version of it) to the *sink*, either directly or via a multi-hop path through other (intermediate) CHs. Such an architecture is adopted by recent standard specifications for sensor networks (e.g., the 802.15.4 standard [1] and the ZigBee Alliance specifications [2]). It significantly reduces the battery drainage of individual sensors, which only need to communicate with their respective CHs over relatively short distances. It also has other advantages in terms of simplifying network management, improving security, and achieving better scalability. On the other hand, the clustering paradigm increases the burden on CHs, forcing them to deplete their batteries much faster than non-CH nodes. The additional energy consumption is attributed to the need to aggregate *intra-cluster* traffic into a single stream that is transmitted by the CH and to relay *inter-cluster* traffic of other CHs. Such relaying is sometimes desirable because of its power-consumption advantage over direct (CH-to-sink) communication. Given the high density of sensors in common deployment scenarios, the traffic volume coming from a CH can be orders of magnitude greater than the traffic volume of an individual sensor. Even though the CH may be equipped with a more durable battery than the individual sensors it serves, the large difference in power consumption between the two can lead to shorter lifetime for the CH. Once the CH dies, no communications can take place between the sensors in that cluster until a new CH is selected.

For clusters with comparable area coverage and node density, the volume of intra-cluster traffic is roughly the same for all clusters. On the other hand, the traffic relayed by different CHs is highly skewed; the closer a CH is to the sink, the more traffic it has to relay, and thus the faster it drains its energy reservoir. Such an imbalanced power consumption situation is essentially caused by the many-to-one communication paradigm in WSNs, i.e., traffic from all sensors is eventually destined to the sink (see Figure 1). If we do not take measures

Part of this work was presented at the *ACM MobiHoc 2005 Conference*, May 25–28, 2005. This research was supported in part by NSF (under grants CNS-0721935, CNS-0627118, CNS-0325979, and CNS-0313234), Raytheon, and Connection One (an I/UCRC NSF/industry/university consortium). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

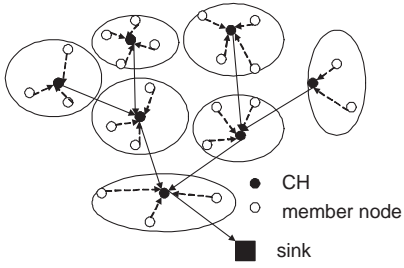


Fig. 1. Traffic implosion in WSNs.

to deliberately balance power consumption at different CHs, a “traffic implosion” situation will arise. More specifically, CHs that are closest to the sink will exhaust their batteries first. Re-assigning sensors to the next-closest CHs to the sink will simply increase the energy consumption of these CHs. As a result, they will eventually be the second batch of CHs to run out of energy. This process continues to the next level of CHs, propagating from inside out and eventually leading to early loss of coverage and partitioning of the topology. Our goal in this paper is to design optimal power allocation strategies that address this imbalance by maximizing the *coverage time*, defined as the time until one CH runs out of battery<sup>1</sup>. These strategies deliberately offset the impact of the skewed load by appropriately adjusting the transmission range (equivalently, transmission power, cluster size) of different CHs. Because the volume of relayed traffic is also affected by the underlying routing scheme, a joint routing/clustering design methodology is needed to achieve power balance among CHs.

### B. Related Work

Extensive research has been dedicated to the study of clustering algorithms for ad hoc and wireless sensor networks. Early clustering algorithms mainly focused on the connectivity issue (e.g., [6], [14], [20], [15], [26], [19], [25]), aiming at generating the minimum number of clusters that ensures network connectivity. In these algorithms, the election of the CH is done based on node identity [6], [14], [20], connectivity degree [15], or connected dominating set [26], [19], [25].

Recently, there has been increased interest in studying energy-efficient clustering algorithms, in the context of both ad hoc and sensor networks [17], [16], [4], [5], [10], [7], [8], [21]. In [17], the authors proposed the LEACH algorithm, in which the CH role is dynamically rotated among all sensors in the cluster. Energy is evenly drained from various sensors, leading to improved network lifetime. A similar CH-scheduling scheme was proposed in [21] for a time-slotted WSN. In this scheme, several disjoint dominating sets are found and are activated successively. Nodes that are not in the currently active dominating set are put to sleep. A distributed algorithm was proposed to obtain a set-schedule sequence for which the network lifetime is within a logarithmic factor of the maximum achievable lifetime. In general, rotation-based schemes require

<sup>1</sup>Other definitions for coverage time may also be used, such as the time until  $x\%$  of coverage is lost or the time until the network is partitioned. Such definitions will be considered in a future work.

excessive processing and communication overheads for CH re-election and broadcasting of the new CH information.

“Load-balanced” algorithms (e.g., [16], [4], [5]) focus mainly on balancing the intra-cluster traffic load, and ignore inter-cluster traffic. In [16], sensors are clustered according to “load-balancing” metrics, whereby the traffic volumes originating from various clusters are equalized. The authors in [4] extended the work in [16] by integrating the concept of load balancing into traditional node-id/connectivity-degree based clustering to produce a longer CH lifespan. In [5], the max-min  $d$ -cluster algorithm was proposed to extend the traditional 1-hop cluster to a  $d$ -hop cluster while generating load-balanced clusters. This extension achieves better load balancing using fewer clusters.

Distributed clustering algorithms were proposed in [7], [8], with the objective of minimizing the energy spent in communicating information to the sink. It should be noted that minimizing the total energy consumption is not equivalent to maximizing coverage time, as the former criterion does not guarantee balanced power consumption at various CHs. By shifting the load from over-power-drained CHs to under-power-drained CHs, coverage time can be maximized even though the total power consumption is not necessarily minimal.

In [10], the authors proposed clustering algorithms that maximize network lifetime by determining the optimal cluster size and optimal assignment of nodes to preselected CHs. Their exhaustive-search approach assumes full knowledge of the network topology (i.e., the location of each sensor node and each CH in the network). It also ignores inter-cluster traffic.

The scheme in [22] incorporates the impact of inter-cluster traffic in determining the optimal location of the sink so as to maximize the topological lifetime of the network. Power balance among CHs was not considered. To the best of our knowledge, no previous work has adequately addressed power balancing among CHs.

### C. Main Contributions and Paper Organization

The main contributions of this paper are as follows. First, as an alternative to previous “load-balanced” algorithms, we study a “power-balanced” approach that aims at *directly* optimizing coverage time by accounting for the interaction between clustering and routing, i.e., simultaneously taking into consideration the impacts of both intra- and inter-cluster traffic. Second, in contrast to previous algorithms, which are based on heuristics, ours is based on an analytical approach. Depending on the availability of location information, we consider in our analysis both deterministic and stochastic topology models. In the deterministic case, sensors and CHs are arbitrarily placed but their locations are known. The traffic of a CH (which includes intra-cluster traffic plus relayed traffic from other CHs) is delivered to the sink either directly or via other CHs. Using linear programming, we provide an algorithm for joint optimization of cluster sizes and the CH-to-CH routing matrix. More emphasis is then put on the stochastic case, where sensor locations are not available

beforehand. In this case, we consider a sensing region with uniformly distributed sensor nodes. Our analysis guarantees an upper bound on the reliability of the multi-hop path from the originating CH to the sink. Two schemes are proposed for achieving power-balanced communications: *routing-aware optimal cluster planning* and *clustering-aware optimal random relay*. The first scheme is essentially a clustering approach that is developed in the context of shortest-hop inter-CH routing. For this scheme, coverage-time maximization is formulated as a signomial optimization problem that is efficiently solved using generalized geometric programming (GGP) techniques. The optimal cluster sizes are obtained from this analysis. The second scheme is essentially a routing strategy for a given clustering approach, e.g., a “load-balanced” clustering, where all clusters are of the same size. According to this approach, a CH probabilistically chooses to relay the traffic to any neighboring “uplink” CH in the direction of the sink. The “optimal” relaying probabilities to various neighbors are derived through linear programming.

Numerical examples and simulations are used to validate our analysis and compare our proposed schemes with pure “load-balancing” algorithms. Our results indicate that by accounting for the interaction between clustering and routing, the proposed schemes achieve a significant reduction in energy consumption and an improved coverage time for the two considered network models.

The rest of this paper is organized as follows. In section II we consider the coverage-time maximization problem when location information is available (deterministic case). After defining the network and traffic models, we present a linear program for finding the optimal clustering and routing parameters. In section III we consider the stochastic case. The optimization is carried out under two different clustering/routing scenarios. In section IV we validate our analysis using numerical examples and computer simulations. Section V concludes the paper.

## II. COVERAGE-TIME OPTIMIZATION FOR DETERMINISTIC DEPLOYMENT

### A. Network Model

We consider a WSN that consists of two types of nodes: Type-I and Type-II nodes. Type-I nodes, which are called sensing nodes (SNs), are responsible for sensing activities. Such nodes are small, low cost, and disposable. They can be densely deployed across the sensing area. Neighboring SNs are organized into clusters. A type-II node has a more powerful energy source and a stronger computing capability, and is designated as a cluster head (CH). Type-II nodes are responsible for receiving and processing the sensing outcomes of SNs. A CH may collect data from intra-cluster SNs, conduct signal processing (a.k.a., data fusion) on these raw data to create an application-specific view of the cluster, and then relay the fused data to the sink through intermediate CHs.

Let the numbers of these two types of nodes be  $M$  and  $N$ , respectively. Suppose that the  $M + N$  nodes are arbitrarily placed but their locations are known. No assumptions are made on the shape of the sensing region. The availability of location

information is an appropriate assumption in many applications of WSNs (e.g., static WSNs in open regions). It can also apply to networks where sensors are first randomly deployed but later their locations become known, for example through GPS-assisted mechanisms. Each sensing node (for brevity, a sensor) is assigned to one CH. The sensor generates traffic at an average rate of  $\lambda$  bits/second, and sends it to its CH, which in turn delivers it to the sink (which we designate as the  $(N + 1)$ th CH) directly or through other CHs (see Figure 1). We assume that each sensor has sufficient energy to communicate directly with its CH. This could be done by either transmitting at a high enough transmission power or using a low enough transmission rate (and thus a longer duration for each transmitted bit). Furthermore, we assume that the CH depletes its energy at a much faster rate than the sensors it serves. This assumption is justified by the low data rate and duty cycle of commonly used sensors (i.e., for most of time, the sensor is put to sleep, in contrast to the CH, which is active most/all the time). Accordingly, we focus our attention on energy depletion at CHs. From a strategic point of view, a CH is more critical to the coverage of the network than individual sensors.

### B. Channel Model

We use a Rayleigh fading model to describe the channel between two CHs and also between a CH and the sink. At a transmitter-receiver separation  $x$ , the channel gain is given by

$$h(x) = L(d_0) \left( \frac{x}{d_0} \right)^{-n} \xi \quad (1)$$

where  $L(d_0) \stackrel{\text{def}}{=} \frac{G_t G_r l^2}{16\pi^2 d_0^2}$  is the path loss of the close-in distance  $d_0$ ,  $G_t$  is the antenna gain of the transmitter,  $G_r$  is the antenna gain of the receiver,  $l$  is the wavelength of the carrier frequency,  $n$  is the path loss exponent ( $2 \leq n \leq 6$ ), and  $\xi$  is a normalized random variable that represents the fluctuations in the fading process. Under the assumption of Rayleigh fading,  $\xi$  is exponentially distributed.

Because  $\xi$  is random, the received signal is also random<sup>2</sup>. Hence, correct reception of a signal can be guaranteed only on a probabilistic basis. In our work, we require that  $\Pr\{e_r \geq \tau\} \geq \delta_l$  for reliable reception, where  $e_r$  is the energy of the received signal,  $\tau$  is a predefined energy threshold, and  $\delta_l$  is a desired link reliability factor.

### C. Joint Clustering/Routing Optimization

For  $i = 1, \dots, N$ , let  $c_i$  be the total intra-cluster traffic collected by the  $i$ th CH (in bits/sec). The *clustering vector* is defined as  $\mathbf{c} = (c_1, \dots, c_N)$ . Note that by construction the size of cluster  $i$ , i.e., the number of sensors associated with CH  $i$ , is  $c_i/\lambda$ . For  $i \in \{1, 2, \dots, N\}$  and  $j \in \{1, 2, \dots, N + 1\}$ , with  $i \neq j$ , let  $v_{ij}$  be the inter-cluster traffic that is relayed from CH  $i$  to CH  $j$ . The *routing matrix*  $\mathbf{R}$  is the  $N \times (N + 1)$  matrix of elements  $v_{ij}$ ,  $i = 1, \dots, N$  and  $j = 1, \dots, N + 1$ . We let

<sup>2</sup>Cost and energy considerations in WSNs prohibit the use of fast (intra-packet) power control to combat the fluctuations in channel fading, as is typically done in cellular networks.

$v_{ii} \stackrel{\text{def}}{=} 0$ . Our goal is to determine the optimal clustering vector  $\mathbf{c}^o$  and routing matrix  $\mathbf{R}^o$  that maximize the coverage time.

Let  $P_i$  be the average power consumption of the  $i$ th CH.  $P_i$  can be written as

$$P_i = e_{rx} \left( c_i + \sum_{1 \leq j \leq N, j \neq i} v_{ji} \right) + e_{tx} \left( \sum_{1 \leq j \leq N+1, j \neq i} v_{ij} \right) + \sum_{1 \leq j \leq N+1, j \neq i} v_{ij} e_{tij}, \quad i = 1, \dots, N \quad (2)$$

where  $e_{rx}$  and  $e_{tx}$  are the per-bit energy consumed in the receive and transmit circuits, respectively, and  $e_{tij}$  is the over-the-air RF energy consumed when transmitting one bit from CH  $i$  to CH  $j$ . The three terms in (2) represent, respectively, the power consumption in the receive circuit, the transmit circuit, and the radio interface.

Let  $d_{ij}$  be the distance between CHs  $i$  and  $j$ . Given  $d_{ij}$  and the channel model in (1), the per-bit received energy  $e_{rij}$  is given by

$$e_{rij} = e_{tij} L(d_0) \left( \frac{d_{ij}}{d_0} \right)^{-n} \xi. \quad (3)$$

For a Rayleigh channel model, the link-reliability requirement can be expressed as

$$\begin{aligned} \delta_l &= \Pr\{e_{rij} \geq \tau\} \\ &= \Pr\left\{ \xi \geq \frac{\tau}{e_{tij} L(d_0)} \left( \frac{d_{ij}}{d_0} \right)^n \right\} = e^{-\frac{\tau d_{ij}^n}{e_{tij} L(d_0) d_0^n}} \end{aligned} \quad (4)$$

From (4) we can express  $e_{tij}$  as

$$e_{tij} = \beta d_{ij}^n, \quad i \neq j \quad (5)$$

where  $\beta \stackrel{\text{def}}{=} \frac{-\tau}{L(d_0) d_0^n \log \delta_l}$  is a constant. Accordingly, for  $i = 1, \dots, N$ , (2) can be written as

$$P_i = e_{rx} \left( c_i + \sum_{1 \leq j \leq N, j \neq i} v_{ji} \right) + \sum_{1 \leq j \leq N+1, j \neq i} v_{ij} (e_{tx} + \beta d_{ij}^n). \quad (6)$$

Note that the unknowns in the above equation are the  $c_i$ 's and  $v_{ij}$ 's.

Let  $E_i$  be the initial residual battery energy of the  $i$ th CH,  $i = 1, \dots, N$ . To maximize the expected coverage time, we need to solve the following optimization problem:

$$\text{maximize}_{\{\mathbf{c}, \mathbf{R}\}} \min \left\{ \frac{E_1}{P_1}, \frac{E_2}{P_2}, \dots, \frac{E_N}{P_N} \right\}. \quad (7)$$

When CHs are initialized with identical batteries, i.e.,  $E_i = E$  for all  $i$ , the optimization problem in (7) is equivalent to:

$$\text{minimize}_{\{\mathbf{c}, \mathbf{R}\}} \max\{P_1, \dots, P_N\}. \quad (8)$$

Hereafter, we focus on (8). The problem constraints are as follows. For CH  $i$ ,  $i = 1, \dots, N$ , the following *inter-cluster flow constraint* must be satisfied

$$a_i c_i + \sum_{1 \leq j \leq N, j \neq i} v_{ji} = \sum_{1 \leq j \leq N, j \neq i} v_{ij} + v_{i, N+1} \quad (9)$$

where  $0 < a_i \leq 1$  is the data aggregation (fusion) efficiency factor that represents the compression effect over the collected

intra-cluster traffic. In addition, the traffic collected by all the CHs over a given time duration must be equal to the traffic generated by all the sensors in the same time duration, i.e.,

$$\sum_{i=1}^N c_i = M\lambda. \quad (10)$$

Introducing an auxiliary variable  $t$ , where  $t \geq \max\{P_1, \dots, P_N\}$ , the objective function (8) and the constraints (9) and (10) can be transformed into the following linear programming (LP) problem in  $\mathbf{c}$ ,  $\mathbf{R}$ , and  $t$ :

$$\begin{cases} \text{minimize}_{\{\mathbf{c}, \mathbf{R}, t\}} t \\ \text{s.t.} \\ \sum_{1 \leq j \leq N, j \neq i} v_{ji} + a_i c_i - \sum_{1 \leq j \leq N, j \neq i} v_{ij} - v_{i, N+1} = 0, \\ \quad i = 1, \dots, N \\ \sum_{i=1}^N c_i = M\lambda \\ \sum_{1 \leq j \leq N, j \neq i} v_{ji} e_{rx} + c_i e_{rx} + \sum_{1 \leq j \leq N, j \neq i} v_{ij} (e_{tx} + \beta d_{ij}^n) \\ + v_{i, N+1} (e_{tx} + \beta d_{i, N+1}^n) - t \leq 0, \quad i = 1, \dots, N \\ v_{ij} \geq 0 \text{ and } c_i \geq 0, \quad i = 1, \dots, N; j = 1, \dots, N+1. \end{cases} \quad (11)$$

Using standard LP techniques, the above problem can be easily solved for  $\mathbf{c}^o$  and  $\mathbf{R}^o$ .

We now comment on the computational complexity of the above formulation. It is well known that, in general, the worst-case execution time of an LP problem is  $O(k^3)$ , where  $k$  is the number of variables in the problem. An inspection of the LP problem in (11) reveals that for a WSN of  $N$  CHs, the total number of variables is  $N^2 + N + 1$ . A closer examination of (11) reveals that this problem exhibits a *sparse* structure, which can be exploited by many LP solvers to significantly reduce the solution time. More specifically, for a WSN of  $N$  CHs, the total number of constraints in (11) is  $2N$ . In each constraint, the number of variables is not more than  $2N$ . The sparsity of the problem is clear because the number of constraints and the number of variables that appear in each constraint is orders of magnitude smaller than the total number of variables. This sparse structure implies that in practice the solution time of the problem is much shorter than the worst-case running time of  $O(k^3)$ .

#### D. Clustering Algorithm

Let  $\mathbf{c}^o = (c_1^o, \dots, c_N^o)$  be the resulting optimal clustering vector. For  $i = 1, \dots, N$ , CH  $i$  is assigned  $M_i^o = c_i^o / \lambda$  sensor nodes. Node assignment is done as follows. Sensor nodes are considered sequentially, one at a time. A given sensor is assigned to the closest CH, say  $i$ , provided that the number of assigned sensors to CH  $i$  does not exceed  $M_i^o$ . If it does, then the next-closest CH is considered, and so on. A pseudo code of the algorithm is given in Table 1. Note that depending on the order in which sensors are considered in the algorithm, different assignments (clusters) may be produced. These assignments achieve the same coverage time, i.e., each of them minimizes the maximum power consumption among CHs. However, they differ in the total energy consumption of *individual* sensors. The algorithm can be easily refined to reduce the total sensor-energy consumption of the initial node assignment. This is done as follows. First, we use the algorithm

in Table I to produce an initial assignment. Such an assignment is optimal with respect to (11), but is not necessarily unique. Then, we consider swapping the cluster assignment of all pairs of sensors that belong to different clusters if this swapping results in a reduction in the total *sensor-energy* consumption in the network. Note that such swapping does not change the energy consumption of the corresponding CHs or the values of the  $c_i^o$ 's. For each sensor in the first initial cluster, the number of pairs to be considered for swapping is  $M - M_1^o$ , for a total of  $M_1^o(M - M_1^o)$  for all the sensors in the first cluster. For the second cluster, there are  $M_2^o(M - M_1^o - M_2^o)$  pairs to consider. A simple combinatorial argument shows that the total number of pairs to consider is bounded by  $M^2$ , which is of low complexity.

<b>Input:</b>	$\mathbf{c}^o = (c_1^o, \dots, c_N^o)$
<b>Initialization:</b>	$\mathbf{U}_1 = \dots = \mathbf{U}_N = \emptyset$ (cluster sets)
<b>Begin:</b>	For $i = 1$ to $M$
	For $j = 1$ to $N$
	set $x_{ij}$ to distance
	between sensor $i$ and CH $j$
	endfor
<b>Loop:</b>	$k = \arg_{\{j\}} \min\{x_{ij}, j = 1, \dots, N\}$
	if $c_k^o > 0$
	$c_k^o = c_k^o - \lambda$
	$\mathbf{U}_k = \mathbf{U}_k + \{k\}$
	else
	$x_{ik} = \infty$
	goto <b>Loop</b>
	endif
	endfor
<b>Output:</b>	$\mathbf{U}_1, \dots, \mathbf{U}_N$

TABLE I  
PSEUDO-CODE FOR INITIAL NODE ASSIGNMENT (CLUSTERING)  
FOLLOWING THE COMPUTATION OF  $\mathbf{c}^o$ .

We use the following example to illustrate the above node swapping process. Consider a network of 3 CHs (CH1, CH2, CH3) and 4 SNs ( $A, B, C, D$ ). For SN  $X$ , we use the triple  $(i, j, k)$  to indicate the per-bit energy consumption of node  $X$  when  $X$  is assigned to one of the three CHs. For example,  $B(2, 4, 1)$  means that SN  $B$  requires 2, 4, and 1 Joules/bit to communicate with CH1, CH2, and CH3, respectively. Suppose that the per-bit energy consumptions of various SNs are:  $A(3, 4, 5)$ ,  $B(2, 4, 1)$ ,  $C(1, 3, 1)$ , and  $D(2, 5, 4)$ . Suppose that to achieve optimal clustering, CH1, CH2, and CH3 should be assigned 2, 1, and 1 SNs, respectively. According to the sequential node assignment procedure, the initial node-to-CH assignment is given by  $\{A, C\} \rightarrow \text{CH1}$ ,  $\{D\} \rightarrow \text{CH2}$ , and  $\{B\} \rightarrow \text{CH3}$ , yielding a total energy consumption of 10 Joules/bit. Now, we conduct node swapping between CH1 and other CHs. Four possible assignments can result from such swapping, as shown in Table II.

The second assignment produces the least energy consumption. Starting from this assignment, we then consider swapping

nodes between CH2 and CH3. Here, there is only one permutation to consider:  $\{C, D\} \rightarrow \text{CH1}$ ,  $\{B\} \rightarrow \text{CH2}$ ,  $\{A\} \rightarrow \text{CH3}$ , for a total energy consumption of 12 Joules/bit. Because the total SN energy consumption under this assignment goes up, we stay with the previous assignment (no. 2 in the table above). There is no more combinations to consider, so the assignment  $\{C, D\} \rightarrow \text{CH1}$ ,  $\{A\} \rightarrow \text{CH2}$ , and  $\{B\} \rightarrow \text{CH3}$  is finally adopted, yielding a total energy consumption of 8 Joules/bit, which amounts to 20% reduction compared with the initial node assignment.

### III. COVERAGE-TIME OPTIMIZATION FOR RANDOM DEPLOYMENT

We now consider the case when the locations of individual sensors are not known. The coverage time of the network is optimized by controlling the location and routing parameters of the CHs. To proceed with our analysis, some simplifying assumptions have to be made.

#### A. Network Model

We consider a cone-like sensing region  $\Psi$  of radius  $R$  and angle  $\phi$ . The sink is located at the vertex, as shown in Figure 2. The region  $\Psi$  may either be an isolated sensing field or a part of a larger sensing field of a general shape (see remark later in this section). The cone-like geometry, albeit idealistic, serves as a basis for understanding the intrinsic tradeoffs involved in a joint clustering/routing optimization framework. In addition, it still captures the fundamental traffic-implosion phenomenon in general WSNs. It has been widely used in the analysis of sensor networks. For example, a circular region (which is a special case of a cone) was recently used in [23], [11].

Sensors are uniformly distributed across  $\Psi$  with density  $\rho$ . Due to energy considerations, only those sensors within distance  $r_0$  from the sink can communicate directly with the sink; all other sensors are organized into clusters and they communicate their data through their respective CHs. Like many distance-based (or equivalently, received-signal-strength-based) cluster formation algorithms, we assume that each CH is located at the center of its cluster.

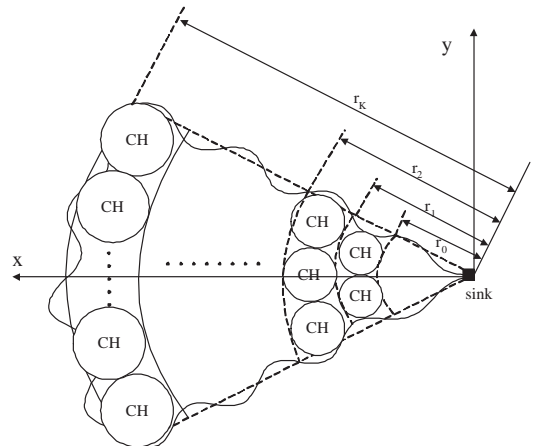


Fig. 2. Cluster formation in a cone-like region.

No.	Assignment	Energy Consumption (Joules/bit)
1	$\{A, D\} \rightarrow \text{CH1}, \{C\} \rightarrow \text{CH2}, \{B\} \rightarrow \text{CH3}$	9
2	$\{C, D\} \rightarrow \text{CH1}, \{A\} \rightarrow \text{CH2}, \{B\} \rightarrow \text{CH3}$	8
3	$\{A, B\} \rightarrow \text{CH1}, \{D\} \rightarrow \text{CH2}, \{C\} \rightarrow \text{CH3}$	11
4	$\{B, C\} \rightarrow \text{CH1}, \{D\} \rightarrow \text{CH2}, \{A\} \rightarrow \text{CH3}$	13

TABLE II  
POSSIBLE OUTCOMES OF NODE SWAPPING.

The procedure for cluster formation consists of two steps: the deployment of CHs and the assignment of sensors to CHs. Because of the symmetric nature of the area  $\Psi$  and the uniform distribution of sensors, the formation of clusters is also symmetric, i.e., any two CHs with the same distance to the sink should have the same coverage. Such clusters are said to be of the same type. Suppose there are  $K$  types of clusters in the network. We consider the following clustering approach: sensors whose distances to the sink fall in  $(r_{i-1}, r_i]$  are organized into clusters of the  $i$ th type, where  $1 \leq i \leq K$  and  $r_0 < r_1 < \dots < r_K = R$ . As a result, clusters of the  $i$ th type cover the  $i$ th ring, defined by the area  $\{(x, y) | r_{i-1}^2 < x^2 + y^2 \leq r_i^2, (x, y) \in \Psi\}$ . Accordingly, the CHs of the  $i$ th ring are placed along the circle  $\{(x, y) | x^2 + y^2 = d_i^2\}$  with equal spacing between consecutive CHs, where  $d_i = \frac{r_{i-1} + r_i}{2}$ . A sensor located in the  $i$ th ring is assigned to the nearest CH in the same ring. In the analysis, we assume that a sufficiently large number of CHs are placed in each ring such that the area covered by each CH can be approximated by a small circle, as shown in Figure 2. Such a ring-based model enables us to analytically capture the relationship between the traffic volume relayed by a CH and the CH's distance to the sink. Our subsequent optimization treatment is based on this relationship. In practice, the above clustering can be implemented as follows. First, CHs are placed according to the outcome of the optimization procedure. Then, each SN associates itself with the CH from which it receives the strongest beacon signal. In the simulations section, we test the validity of this implementation approach and show that the perfect-ring-model assumed in the analysis has a negligible impact on network performance. Unless indicated otherwise, we assume the same channel and energy models used in the previous section.

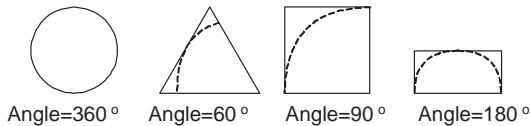


Fig. 3. Shapes that can be approximated by a cone.

**Remark:** Although our model assumes a cone-like sensing area and a two-tier network structure, the analysis adequately captures the intrinsic interaction between inter- and intra-cluster traffic. In addition, we note that this cone shape is general enough to approximate many other shapes. For example, as shown in Figure 3, a cone can approximate the shapes of circle, triangle, square, and rectangle, respectively,

when the angle  $\phi$  is properly set. The analysis can also be extended to handle a non-regularly-shaped region by covering it with a series of element shapes that can be approximated by cones, similar to the approach used in cellular networks (in cellular networks, the region is approximately covered by hexagons). A multi-layered organization of sensors, such as the “spine” hierarchy [12], can also be accommodated in our analytical framework. In this case, our analysis provides the optimal CH coverage time for the “base” layers and a sub-optimal coverage time for the whole network. The details of such extensions are beyond the scope of this paper and will be considered in a future work.

### B. Routing Models

Our coverage-time maximization is carried out under two different routing models, which are described below. Because the two models differ in the hop-count of the path from the source CH to the sink, it is more appropriate to reflect the quality of the communication in terms of a constraint  $\delta_p$  on the probability of a successful *end-to-end* reception. For a path of  $K$  links that experience independently and identically distributed (i.i.d.) fading<sup>3</sup>, the link reliability  $\delta_l$  should be at least  $\delta_p^{\frac{1}{K}}$ .

1) *Shortest-Distance Relay*: In this scenario, traffic is relayed through the closest CH in the adjacent ring towards the sink. More specifically, a CH in the  $i$ th ring receives traffic originating from its own cluster as well as traffic relayed from CHs in the  $(i + 1)$ th ring, and forwards the combined traffic to the closest CH in the  $(i - 1)$ th ring. Relaying continues hop-by-hop until the sink is reached.

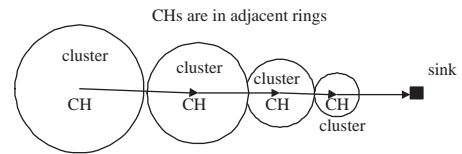


Fig. 4. Shortest-distance relay scheme.

For the shortest-distance relay, we consider a *routing-aware clustering* mechanism that balances power consumption at different CHs. Clearly, the *radius profile* of the clusters, given by  $\frac{1}{2}(r_1 - r_0), \dots, \frac{1}{2}(r_K - r_{K-1})$ , is critical to power consumption at different CHs. For example, reducing  $\frac{1}{2}(r_i - r_{i-1})$  results

<sup>3</sup>The assumption of i.i.d. link fading is justified by noting that the distance between consecutive CHs is much larger than the carrier wavelength for a system operating in the 2.4 GHz frequency region, which is typical in current WSN standards.



in smaller clusters in the  $i$ th ring, which leads to less local traffic from these clusters, shorter transmission distances to subsequent CHs in the  $(i-1)$ th ring, and a higher number of CHs in the  $i$ th ring. Because of the symmetry in the topology and traffic load, the traffic from the CHs in the  $(i+1)$ th ring will be evenly shared by a higher number of CHs in the  $i$ th ring, so the volume of the relayed traffic carried by individual CHs in the  $i$ th ring will decrease. All of these factors contribute to a reduced power consumption at the CHs in the  $i$ th ring. On the other hand, the reduction in the area of the  $i$ th ring must be compensated for by other clusters (e.g., the clusters in the  $j$ th ring), because of the fixed number of rings in the system. In an analogous manner, power consumption at CHs in ring  $j$  will increase. Therefore, by deliberately adjusting the cluster size in different rings, a more balanced power consumption at different CHs is achieved, leading to an increase in the coverage time. This is addressed in the routing-aware optimal cluster planning scheme presented in Section III-C.

2) *Random Relay*: In this scenario, a CH has the freedom to relay its data to the closest CH in any of the inner rings (this also includes the case of sending data directly to the sink). Let  $\alpha_{ij}$  be the fraction of the load that a CH in the  $i$ th ring transmits to the closest CH in the  $j$ th ring, where  $0 \leq j < i$  and  $j = 0$  denotes direct transmission to the sink. For a given clustering structure that contains  $K$  rings, the *relaying matrix*  $\mathbf{A}$  is defined as follows

$$\mathbf{A} = \begin{bmatrix} \alpha_{10} & 0 & 0 & \dots & 0 \\ \alpha_{20} & \alpha_{21} & 0 & \dots & 0 \\ \alpha_{30} & \alpha_{31} & \alpha_{32} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{K0} & \alpha_{K1} & \alpha_{K2} & \dots & \alpha_{KK-1} \end{bmatrix}. \quad (12)$$

where the  $i$ th row of  $\mathbf{A}$  represents the probabilities for relaying a packet at the  $i$ th ring to the closest CH in rings  $0, 1, \dots, i-1$ . The matrix  $\mathbf{A}$  plays a critical role in balancing power consumption at different CHs. For example, increasing  $\alpha_{ij}$  will reduce the relayed traffic carried by all CHs in rings  $j+1, j+2, \dots, i-1$ . But this comes at the expense of higher power consumption at the CHs in the  $i$ th ring, because of the longer transmission distance which, on average, increases from approximately  $\frac{1}{2}(r_i + r_{i-1}) - \frac{1}{2}(r_{\lfloor \frac{1}{2}(i+j) \rfloor} + r_{\lfloor \frac{1}{2}(i+j) \rfloor - 1})$  to  $\frac{1}{2}(r_i + r_{i-1}) - \frac{1}{2}(r_j + r_{j-1})$ . By deliberately adjusting the relay probabilities at different CHs, a more balanced power consumption at different CHs can be achieved.

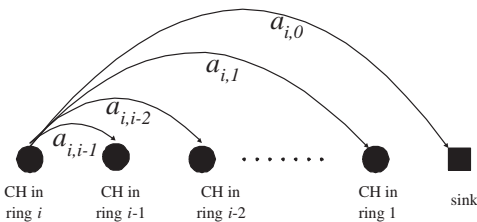


Fig. 5. Random relay scheme.

In section III-D, we propose a *clustering-aware optimal random relay* scheme that addresses the problem of finding

the optimal relay matrix for a given clustering structure. More specifically, we consider a homogeneous clustering structure, i.e.,  $r_1 - r_0 = r_2 - r_1 = \dots = r_K - r_{K-1}$ , so that all clusters roughly has the same number of sensors. This structure is exactly the “load balanced” clustering presented in [16]. It is highly desirable in practice because of its simplicity. Through numerical examples, we show that the proposed clustering-aware optimal random relay scheme achieves longer expected coverage time compared with pure “load balanced” clustering.

Let  $P_i$  be the average power consumption of a CH in the  $i$ th ring. For both routing strategies, we adopt the following energy model:

$$P_i = e_{rx}(\lambda_{oi} + \lambda_{ri}) + e_{tx}(\lambda_{oi} + \lambda_{ri}) + P_{Ti}(\lambda_{oi} + \lambda_{ri}, \Gamma) \quad (13)$$

where  $\lambda_{oi}$  is the expected intra-cluster bit rate (in bits/second),  $\lambda_{ri}$  is the expected bit rate of the incoming inter-cluster traffic that is to be relayed by the underlying CH, and  $P_{Ti}(\cdot, \cdot)$  is the RF transmission power expressed as a function of the outgoing bit rate and the employed routing scheme  $\Gamma$ . The quantities  $e_{rx}$  and  $e_{tx}$  were previously defined.

As in section II-C, under the assumption of equal initial battery energies, the coverage-time maximization problem is equivalent to the following problem:

$$\text{minimize } \max\{P_1, \dots, P_K\}. \quad (14)$$

where the optimization is carried out with respect to either the clustering or the relaying parameters, depending on which routing approach is employed, as explained next.

### C. Routing-Aware Optimal Cluster Planning Scheme

In this section, we formulate the optimal cluster organization problem in the context of shortest-distance (hop-by-hop) routing. Under this routing scheme, a CH in the  $i$ th ring transmits its data to the nearest CH in the  $(i-1)$ th ring. Let  $x_i$  be the physical distance between these two CHs. The expected transmission power is given by

$$P_{Ti} = e_{ti}(\lambda_{oi} + \lambda_{ri}) \quad (15)$$

where  $e_{ti}$  is the consumed transmission energy per bit for the underlying CH. Substituting (15) into (13), the expected communication power consumption of any CH in ring  $i$  is given by

$$P_i = (e_{rx} + e_{tx} + e_{ti})(\lambda_{oi} + \lambda_{ri}). \quad (16)$$

Given  $e_{ti}$ , the corresponding received energy  $e_{ri}$  is given by

$$e_{ri} = e_{ti}L(d_0) \left(\frac{x_i}{d_0}\right)^{-n} \xi. \quad (17)$$

The link-reliability requirement can be expressed as

$$\delta_l = \Pr\{e_{ri} \geq \tau\} = \Pr\left\{\xi \geq \frac{\tau}{e_{ti}L(d_0)} \left(\frac{x_i}{d_0}\right)^n\right\} = e^{-\frac{\tau x_i^n}{e_{ti}L(d_0)d_0^n}} \quad (18)$$

Under min-hop routing, the maximum number of links of an end-to-end path is  $K$ . So to guarantee the constraint  $\delta_p$  on path reliability, the minimum link reliability must be

$$\delta_l = \delta_p^{\frac{1}{K}}. \quad (19)$$

Equating (18) and (19), the minimum transmit energy per bit that satisfies the end-to-end reliability requirement is given by

$$e_{tx} = \frac{-K\tau x_i^n}{L(d_0)d_0^n \ln \delta_p}. \quad (20)$$

An approximation that provides an upper bound on the expected coverage time can be obtained by replacing  $x_i$  in (20) with a lower bound  $x_{i \min}$  that is given by:

$$x_{i \min} = \begin{cases} \frac{r_1+r_0}{2}, & \text{for } i = 1 \\ \frac{r_i-r_{i-2}}{2}, & \text{for } i = 2, \dots, K. \end{cases} \quad (21)$$

This lower bound represents the sum of the radius of a cluster in the  $i$ th ring and the radius of the nearest cluster in the  $(i-1)$ th ring. It is easy to see that the distance between the CHs of the corresponding two clusters is at least  $x_{i \min}$ .

Let  $\lambda_{\text{total}_i}$ ,  $i = 1, \dots, K$ , denote the bit rate of the aggregate traffic that originates from the clusters in rings  $i$  through  $K$ . Then,

$$\lambda_{\text{total}_i} = \pi(R^2 - r_{i-1}^2)\rho\lambda \frac{\phi}{2\pi}, \quad i = 1, \dots, K. \quad (22)$$

Because relaying is done hop-by-hop, the total traffic load carried by the CHs in the  $i$ th ring is equal to the total traffic volume originating from all clusters in rings  $i$  to  $K$ . Due to the symmetry of the rings and the uniform distribution of sensors, the traffic from the  $i$ th ring is evenly distributed among all CHs in that ring. The number of CHs in the  $i$ th ring is approximately given by

$$N_i \approx \frac{2\pi r_i}{r_i - r_{i-1}} \frac{\phi}{2\pi}. \quad (23)$$

The quality of this (and other) approximations is evaluated in section IV through a comparison with more realistic simulations.

Accordingly, the average traffic load at any CH in ring  $i$  is given by

$$\begin{aligned} \lambda_{oi} + \lambda_{ri} &= \frac{\lambda_{\text{total}_i}}{N_i} \\ &\approx \frac{(R^2 - r_{i-1}^2)(r_i - r_{i-1})}{2r_i} \rho\lambda. \end{aligned} \quad (24)$$

Substituting (24), (20), and (21) in (16), the expected power consumption of any CH in the  $i$ th ring can be approximated by *signomial functions*<sup>4</sup> of the radius profile  $\mathbf{r} \stackrel{\text{def}}{=} (r_1, r_2, \dots, r_K)$ . More specifically, they are given by

$$P_1 = \left[ e_{rx} + e_{tx} + \frac{K\tau}{-L(d_0)d_0^n \log \delta_p} \left( \frac{r_1 + r_0}{2} \right)^n \right] \times \frac{(R^2 - r_0^2)(r_1 - r_0)}{2r_1} \rho\lambda \quad (25)$$

and

$$P_i = \left[ e_{rx} + e_{tx} + \frac{K\tau}{-L(d_0)d_0^n \log \delta_p} \left( \frac{r_i - r_{i-2}}{2} \right)^n \right] \times \frac{(R^2 - r_{i-1}^2)(r_i - r_{i-1})}{2r_i} \rho\lambda, \text{ for } i = 2, \dots, K. \quad (26)$$

Our goal now is to determine the optimal  $\mathbf{r}$  that minimizes

the average maximum power consumption among all CHs. This optimization problem can be formulated as follows:

$$\begin{cases} \text{minimize}_{\{r_1, \dots, r_K\}} \{ \max \{ P_1, \dots, P_K \} \} \\ \text{s.t.} \\ r_0 < r_1 < \dots < r_K = R \end{cases} \quad (27)$$

where  $P_i$ ,  $i = 1, \dots, K$ , are given by (25) and (26).

By introducing the auxiliary variable  $t \geq P_i$  for  $1 \leq i \leq K$ , the optimization problem in (27) can be transformed into the following equivalent form:

$$\begin{cases} \text{minimize}_{\{\mathbf{r}, t\}} t \\ \text{s.t.} \\ t^{-1} P_i \leq 1, \quad i = 1, \dots, K \\ r_{i-1} r_i^{-1} < 1, \quad i = 1, \dots, K \\ r_K = R. \end{cases} \quad (28)$$

An examination of (28) reveals that its objective function is a *monomial*, the inequality constraints are *signomials*, and the equality constraint is a *monomial* of the variables  $(\mathbf{r}, t)$  (refer to the appendix for the concepts of monomial, posynomial, and signomial). Therefore, (28) is a signomial optimization problem of the standard form [13]. Its optimal solution can be efficiently found using GGP algorithms introduced in [13] and [24].

#### D. Clustering-Aware Optimal Random Relay Scheme

For a given clustering structure, i.e., under a given radius profile  $(r_0, r_1, \dots, R)$ , we now address the maximization of coverage time by determining the optimal relay probabilities at different CHs. Recall that in this scenario, a CH in the  $i$ th ring relays its traffic to the closest CH in the  $j$ th ring with probability  $\alpha_{ij}$ . To facilitate our analysis, we first introduce the variable  $f_{ij}$ ,  $i = 1, \dots, K$  and  $0 \leq j < i$ , which represents the aggregate traffic (in bits/sec) from the CHs in the  $i$ th ring to the CHs in the  $j$ th ring. The basic idea is to first formulate the optimization problem in terms of  $f_{ij}$ 's. After the optimal flow parameters  $f_{ij}^o$ 's are obtained, the optimal relay probabilities can be simply calculated as

$$\alpha_{ij}^o = \frac{f_{ij}^o}{\sum_{k=0}^{i-1} f_{ik}^o}, \quad i = 1, \dots, K. \quad (29)$$

For the  $i$ th ring, the aggregate traffic must satisfy the following flow-conservation constraints:

$$\sum_{j=i+1}^K f_{ji} + \lambda_{\text{ring}_i} = \sum_{k=0}^{i-1} f_{ik}, \quad i = 1, \dots, K \quad (30)$$

where  $\lambda_{\text{ring}_i}$  denotes the aggregate traffic that originates from the clusters in ring  $i$ . It is given by

$$\lambda_{\text{ring}_i} = \pi(r_i^2 - r_{i-1}^2)\rho\lambda \frac{\phi}{2\pi}. \quad (31)$$

In addition, all data should be finally sent to the sink, i.e.,

$$\sum_{j=1}^K f_{j0} = \sum_{i=1}^K \lambda_{\text{ring}_i}. \quad (32)$$

By construction, the traffic load of the  $i$ th ring is evenly distributed among all CHs in that ring. Therefore, the average

<sup>4</sup>See the appendix for the definition of signomial functions.



traffic load at any CH in the  $i$ th ring is given by

$$\lambda_{oi} + \lambda_{ri} = \frac{\sum_{k=0}^{i-1} f_{ik}}{N_i} \quad (33)$$

where  $N_i$  is the number of CHs in the  $i$ th ring, and its value is given in (23).

Substituting (33) into (15), the average transmission power for a CH in the  $i$ th ring is given by

$$P_{Ti} = \sum_{k=0}^{i-1} \frac{f_{ik}}{N_i} e_{tr(ik)} \quad (34)$$

where  $e_{tr(ik)}$ ,  $k = 0, 1, \dots, i-1$ , is the transmission energy per bit for relaying traffic from a CH in the  $i$ th ring to its nearest CH in ring  $k$ . Following a similar development to the one in section 3.3,  $e_{tr(ik)}$  is derived as follows:

$$e_{tr(ik)} = \frac{-(K-i+1+k)\tau y_{ik}^n}{L(d_0)d_0^n \ln \delta_p}, \quad \text{for } 0 \leq k < i \text{ and } i = 1, \dots, K \quad (35)$$

where  $y_{ik}$  is the shortest distance between a CH in the  $i$ th ring and its nearest counterpart in the  $k$ th ring. Keeping in mind the symmetry of our topology, a lower bound on  $y_{ik}$  is simply given by

$$y_{ik \min} = \begin{cases} \frac{1}{2}(r_i + r_{i-1}) - \frac{1}{2}(r_k + r_{k-1}), & \text{for } k \neq 0 \\ \frac{1}{2}(r_i + r_{i-1}), & \text{for } k = 0. \end{cases} \quad (36)$$

The above analysis applies to any clustering technique. In the special case of ‘‘load-balanced’’ clustering, i.e., each cluster is of the same size and thus each ring has the same ‘‘thickness’’, this lower bound reduces to:

$$y_{ik \min} = \begin{cases} \frac{R-r_0}{K}(i-k), & \text{for } k \neq 0 \\ r_i - \frac{R-r_0}{2K}, & \text{for } k = 0 \end{cases} \quad (37)$$

where now  $r_i = r_0 + i\frac{R-r_0}{K}$ , for  $i = 1, \dots, K$ . To obtain a lower bound on power consumption,  $y_{ik \min}$  can be used in place of  $y_{ik}$  in (35).

In (35), the factor  $(K-i+1+k)$  is used instead of  $K$  in (20) to accommodate a worst-case link reliability requirement. Recall that in deriving (20), we split the end-to-end path reliability  $\delta_p$  among  $K$  links, providing a conservative estimate of the link reliability for each of the  $K$  hops. In the case of the random relay scheme, the traffic that is relayed to a CH in the  $i$ th ring from outer rings may have traversed from one to  $K-i$  hops before reaching the  $i$ th ring. So if this traffic is to be transmitted from the  $i$ th ring to the  $k$ th sink, its maximum hop count would be  $K-i+1+k$ , which explains the appearance of this factor in (35).

Substituting (33) and (34) into (13), the expected power consumption of any CH in ring  $i$  is given by

$$P_i = \frac{1}{N_i} \sum_{k=0}^{i-1} f_{ik} (e_{rx} + e_{tx} + e_{tr(ik)}). \quad (38)$$

From (38), it is clear that for a given radius profile  $(r_0, r_1, \dots, r_K)$ , the expected power consumption at different CHs can be expressed as linear functions of the traffic flows  $f_{ij}$ . Our goal is to determine the optimal values for these flows that maximize the expected coverage time. This is equivalent

to the following min-max optimization problem:

$$\begin{cases} \text{minimize}_{\{f_{10}, \dots, f_{K0}, f_{21}, \dots, f_{KK-1}\}} \max\{P_1, \dots, P_K\} \\ \text{s.t.} \\ \sum_{j=i+1}^K f_{ji} + \lambda_{\text{ring}_i} = \sum_{k=0}^{i-1} f_{ik}, \quad i = 1, \dots, K \\ \sum_{j=1}^K f_{j0} = \sum_{i=1}^K \lambda_{\text{ring}_i} \end{cases} \quad (39)$$

where the  $P_i$ 's are given in (38).

By introducing the auxiliary variable  $t$ , (39) can be transformed into the following equivalent optimization problem:

$$\begin{cases} \text{minimize}_{\{f_{10}, \dots, f_{K0}, f_{21}, \dots, f_{KK-1}, t\}} t \\ \text{s.t.} \\ P_i - t \leq 0, \quad i = 1, \dots, K \\ \sum_{j=i+1}^K f_{ji} + \lambda_{\text{ring}_i} = \sum_{k=0}^{i-1} f_{ik}, \quad i = 1, \dots, K \\ \sum_{j=1}^K f_{j0} = \sum_{i=1}^K \lambda_{\text{ring}_i} \end{cases} \quad (40)$$

An examination of (40) and (38) shows that this is a standard linear programming problem, which can be solved using existing numerical algorithms such as Simplex. After obtaining the optimal  $f_{ij}$ , the optimal relaying matrix can be calculated according to (29).

**Remark:** As verified in section IV, in most cases, the objective functions in (40) and (28) are minimized when power consumptions at different CHs are equalized. This is because if there is a CH with power  $P_i$  that is larger than the power consumption of other CHs, then  $P_i$  can always be lowered without violating the constraints by decreasing  $r_i$  in (28) or  $f_{i,j-1}$  in (40), leading to an increase in the power consumption of some other CHs. As a result, the maximum power consumption will be minimized when a balance is reached across all CHs.

In addition, we note that the clustering-aware optimal random relay algorithm can be easily used to tackle the scenario in which the average traffic rate fluctuates over time (i.e., the traffic generation process is stationary only for a certain time interval, but becomes non-stationary over the lifetime of the network). A sliding-window mechanism can be used to decide the average traffic rate in each stationary time interval. Whenever there is a significant change in the average load, the optimization algorithm can be re-run to compute new values for the routing parameters of the random relay scheme. This way, the actions of each CH become adaptive to network dynamics.

## IV. NUMERICAL RESULTS AND SIMULATIONS

### A. Deterministic Scenario

We first consider a WSN for which the node-location information is available (section II). We start with a simple line topology (Figure 6) that is meant to demonstrate important aspects of the power-balancing approach. The network is composed of four CHs and 200 sensor nodes. The CHs and the sink are spaced out evenly with 10 meters between each other. Let  $\lambda = 5$  bits/second,  $a_1 = a_2 = a_3 = a_4 = 1$ ,  $e_{rx} = e_{tx} = 50$  nJoule/bit,  $d_0 = 10$  meters,  $G_t = G_r = 1$ ,  $n = 4$ ,  $\tau = 10^{-17}$  Joules,  $\delta_l = 0.99$ , and the carrier frequency be 2.4 GHz. Table III depicts the optimal clustering vector

CH id	Clustering Vector $\mathbf{c}^\circ$ (bits/s)		Power Consumption ( $\mu\text{W}$ )		Routing Matrix $\mathbf{R}^\circ$				
	PB	LB	PB	LB	CH1	CH2	CH3	CH4	CH5
CH1	369.1	250	37.3	101	0	0	0	0	369.1
CH2	321.1	250	37.3	75.8	0	0	0	0	321.1
CH3	205.5	250	37.3	50.5	0	0	0	0	205.5
CH4	104.3	250	37.3	25.3	0	0	0	0	104.3

TABLE III

COMPARISON BETWEEN POWER-BALANCING (PB) AND LOAD-BALANCING (LB) APPROACHES UNDER THE DETERMINISTIC SETUP (LINE TOPOLOGY).

and routing matrix, derived from (11). The proposed power-balancing (PB) clustering approach is compared with a load-balancing (LB) clustering approach [16] that uses hop-by-hop traffic relay between consecutive CHs. In the LB approach, the total intra-cluster traffic ( $200\lambda = 1000$  bits/sec) is split equally among the four CHs. Each CH relays its traffic to the sink hop-by-hop through intermediate CHs. For example, CH4 sends its 250 bits/second traffic to CH3, and CH3 in turn transmits 500 bits/second to CH2, and so on.

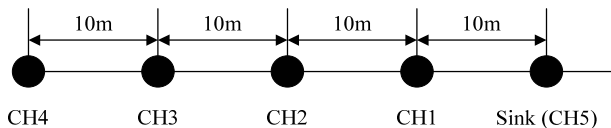


Fig. 6. Line topology for a deterministic WSN.

As expected, the PB approach produces unequal cluster sizes (second column in the table), but whose CH power consumptions are equal ( $37.3 \mu\text{W}$ ). Power balance is achieved by assigning more intra-cluster traffic (larger clusters) to CHs that are closer to the sink. In contrast, the LB approach produces equal-size clusters with variable CH power consumptions. Specifically, CH1 has the highest power consumption, so it is the first CH to run out of battery. Compared with the LB approach, the PB approach prolongs the coverage time by about 170%. The results in Table III indicate that for the PB approach, direct CH-to-sink communication is preferable (in terms of coverage time) over multi-hop communications. At first this may be surprising, as the channel nonlinearity suggests that a multi-hop path with short distance per hop is more energy-efficient than a single-hop path with a long transmission distance. However, the optimal structure of the above example can be explained by noting that the optimization is performed under min-max power consumption criterion. As a result, if a given solution requires some traffic to be relayed between intermediate CHs, then we can always construct another solution that requires a smaller maximum power consumption than the original one. In fact, it is easy to show theoretically that under the min-max power consumption criterion, for any line topology with no imposed limit on the cluster size, direct CH-to-sink communication is the optimal strategy.

The optimization in section II-C was carried out without imposing an upper bound on the number of sensors that can belong to a cluster. In practice, MAC considerations may require imposing such a bound. To test the impact of imposing such a bound, we consider a variant of the optimization

procedure of section II-C, in which we let  $c_i \leq c_{\max}$  for all  $i$ . Table IV depicts the resulting optimal clustering vector and routing matrix for the same line topology and using  $c_{\max} = 300$  bits/second (60 sensors/cluster). In this case, we notice that for the farthest CH (CH4), some traffic is “optimally” delivered using multi-hop forwarding via CH1 and CH2.

### B. Stochastic Scenario

We now consider the stochastic scenario for a circular ( $\phi = 360^\circ$ ) sensing region. We study the performance of the optimal cluster planning and optimal random relay schemes, and contrast them with the LB clustering approach. To get a clear picture of the advantages of adjusting the routing parameters, we use LB clustering for the random relay scheme. Recall that the analysis in section III was conducted under some simplifying assumptions (e.g., circular clusters, lower bounds on CH-to-CH distances, etc.). To validate the adequacy of our analytical results, we contrast them with simulations conducted under a more realistic setup (explained below). For the two proposed schemes, we use the analytical results to compute the optimal radius profile  $\mathbf{r}^\circ$  and optimal relaying matrix  $\mathbf{A}^\circ$ . We use these optimal values to *drive* the simulations of the two proposed schemes. Our main performance metric is the maximum expected power consumption of a CH,  $P_{\max} \stackrel{\text{def}}{=} \max\{P_1, \dots, P_K\}$ . The smaller the value of  $P_{\max}$ , the longer is the coverage time. We set the radius of the circular sensing region to  $R = 200$  meters. Sensors are uniformly distributed throughout this region at density  $\rho = 1$ , i.e., the number of sensors in any area  $S$  follows a spatial Poisson distribution with parameter  $\rho S$ . The number of CHs in both the analysis and the simulations is set to  $\sum_{i=1}^K N_i$ , where  $N_i$  is obtained from (23) and  $K$  is given. The location of these CHs is also taken to be the same for the analysis and the simulations. However, in the simulations, clusters are not necessarily circular, and the notion of rings is not strictly followed. Instead, each sensor in a given simulation run is assigned to the nearest CH. As a result, two CHs that have the same distance to the sink may have different cluster sizes. Each sensor generates data according to a Poisson process of rate  $\lambda = 10$  bits/second<sup>5</sup>. Because of the randomness in the traffic and node locations, the powers consumed by different CHs that have the same distance to the sink may be different in the simulations. In this case,  $P_{\max}$  is taken as the maximum

<sup>5</sup>The choice of the traffic model has no impact on the relative performance of the investigated schemes. For this reason, we opted for a simple traffic model.

CH id	Clustering Vector $\mathbf{c}^o$ (bits/s)		Power Consumption ( $\mu\text{W}$ )		Routing Matrix $\mathbf{R}^o$				
	PB	LB	PB	LB	CH1	CH2	CH3	CH4	CH5
CH1	300	250	39.5	101	0	0	0	0	391
CH2	300	250	39.5	75.8	0	0	0	0	340.1
CH3	217.6	250	39.5	50.5	0	0	0	0	217.6
CH4	182.4	250	39.5	25.3	91	40.1	0	0	51.3

TABLE IV

COMPARISON BETWEEN PB AND LB APPROACHES UNDER THE DETERMINISTIC SETUP WITH  $c_{\max} = 300$  BITS/SECOND (LINE TOPOLOGY).

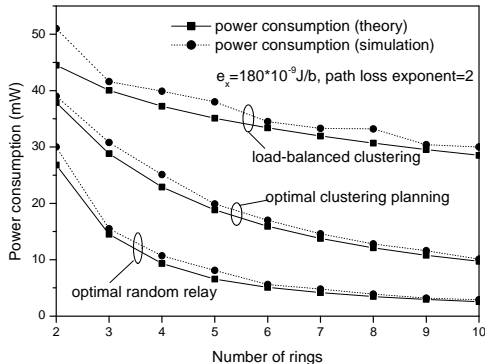


Fig. 7.  $P_{\max}$  vs. number of rings ( $e_x = 180$  nJ/bit,  $n = 2$ ).

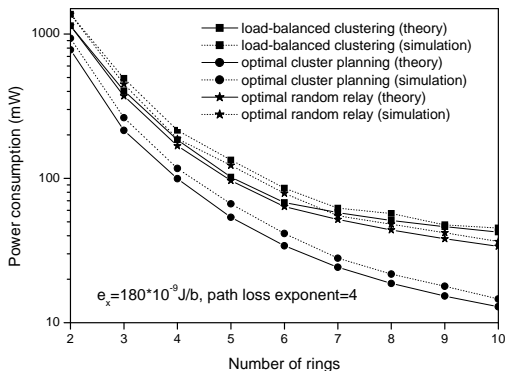


Fig. 8.  $P_{\max}$  vs. number of rings ( $e_x = 180$  nJoule/bit,  $n = 4$ ).

of  $P_{avg,1}, \dots, P_{avg,K}$ , where  $P_{avg,i}$  is the average power of a CH in the  $i$ th ring. We take  $r_0 = 10m$ ,  $G_t = G_r = 1$ ,  $\tau = 10^{-17}$  Joules, and  $\delta_p = 0.99$ .

Figures 7 and 8 depict  $P_{\max}$  versus the number of rings ( $K$ ) for two path loss factors:  $n = 2$  and  $n = 4$ . The transmit-plus-receive per-bit circuit energy is set to  $e_x \stackrel{\text{def}}{=} e_{tx} + e_{rx} = 180$  nJoule/bit. It is observed that the gap between the (approximate) analytical results and the simulations is reasonably small for all examined schemes, with the simulation results being slightly more conservative than the analysis. The disparity between the two is attributed in part to the approximate nature of the analysis and in part to the randomness in the packet generation process and the distribution of sensors within a CH. When  $n = 2$ , both the optimal cluster planning and the optimal random relay schemes result in significantly longer coverage times (smaller  $P_{\max}$  values) than the LB scheme. For  $n = 4$  (Figure 8), the optimal cluster planning scheme maintains its advantage, but the optimal random relay scheme is shown

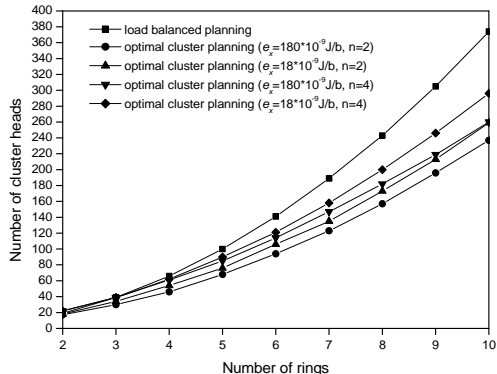


Fig. 9. Number of clusters vs. number of rings.

to achieve only limited power efficiency over LB clustering. This phenomenon can be explained by comparing the optimal relaying matrices for  $n = 2$  and  $n = 4$ . An example of these relaying matrices when  $K = 5$  is listed below. It can be observed that when  $n = 4$ , the optimal random relay scheme prefers to relay most traffic to the CHs in the next ring towards the sink (the values along the diagonal of  $\mathbf{A}_{n=4}^o$  are close to one). This is because now the total power consumption is dominated by the transmission power ( $P_{T_i}$ ), which is highly nonlinear in the transmission distance. As a result, for the random relay scheme, only a small portion of the traffic at each CH is transmitted across intermediate hops; the rest is sent hop-by-hop, making the scheme's behavior quite similar to the LB scheme. Therefore, when  $n$  is large, the flexibility in choosing the next-hop CH offers little performance benefit.

$$\mathbf{A}_{n=2}^o = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.8529 & 0.1471 & 0 & 0 & 0 \\ 0.5117 & 0.4221 & 0.0661 & 0 & 0 \\ 0.6075 & 0 & 0.2598 & 0.1328 & 0 \\ 0.9749 & 0 & 0 & 0 & 0.0251 \end{bmatrix}$$

$$\mathbf{A}_{n=4}^o = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.0064 & 0.0430 & 0.9506 & 0 & 0 \\ 0.0048 & 0.0120 & 0.0829 & 0.9003 & 0 \\ 0.0054 & 0.0154 & 0.0280 & 0.1241 & 0.8271 \end{bmatrix}$$

Figure 9 depicts the total number of formed clusters ( $\sum_{i=1}^K N_i$ ) versus the number of rings ( $K$ ) for the optimal cluster planning and the LB schemes. In addition to achieving a lower  $P_{\max}$  value (longer coverage time), optimal cluster planning also results in a smaller number of clusters, and hence reduced network-management overhead. The reduction

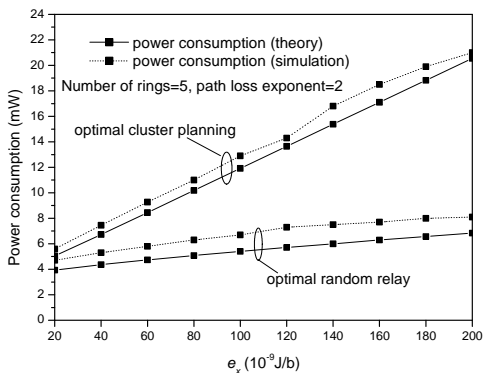


Fig. 10.  $P_{\max}$  vs. circuit energy efficiency  $e_x$  ( $K = 5$ ,  $n = 2$ ).

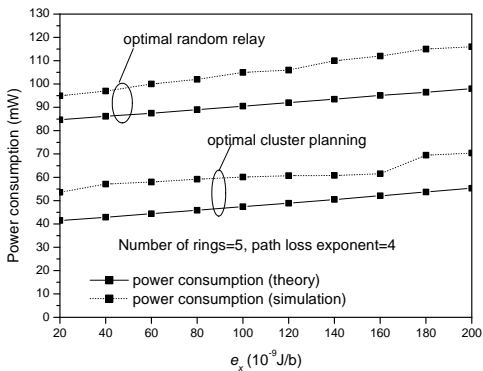


Fig. 11.  $P_{\max}$  vs. circuit energy efficiency  $e_x$  ( $K = 5$ ,  $n = 4$ ).

in the number of clusters comes from the improved energy utilization of under-drained CHs, i.e., in order to balance the power consumption of different CHs, an under-drained CH tends to carry more intra-cluster traffic, hence expanding the size of the cluster and reducing the number of clusters required to cover the sensing region.

In Figures 10 and 11, we study the effects of  $e_x$  and  $n$  when  $K = 5$ . As shown in these figures, when  $n$  is small, optimal random relay generally achieves better coverage time (smaller  $P_{\max}$ ) than optimal cluster planning. As  $e_x$  increases, the relative difference between these two schemes becomes more significant. On the other hand, when  $n$  is large, the optimal cluster planning scheme becomes superior to optimal random relay. This phenomenon can be explained as follows. When  $n$  is small, the circuit power in transmitting and receiving data is comparable with the communication power consumption ( $P_{T_i}$ ). Because optimal cluster planning relies solely on shortest-distance hop-by-hop routing, whereas optimal random relay sometimes bypasses intermediate hops and uses long-distance communication, the latter scheme reduces the circuit power overhead at intermediate CHs. When  $e_x$  increases, circuit power becomes dominant, and multi-hop routes become less energy-efficient. On the other hand, when  $n$  is large, the total power consumption at a CH is dominated by the communication power consumption, which is highly nonlinear in the transmission distance. As a result, short-distance communication becomes more energy-efficient. This drives the optimal random relay to use hop-by-hop routing in sending data. Thus its actual routing style becomes less random and

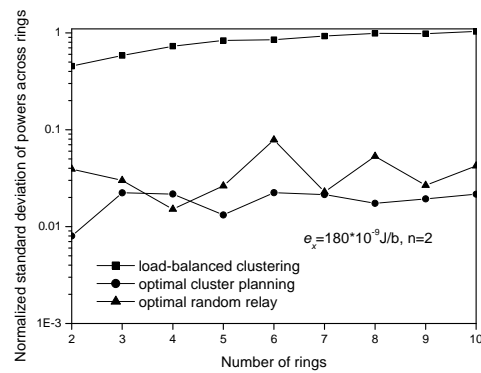


Fig. 12. Normalized standard deviation of power consumption vs.  $K$  ( $e_x = 180\text{nJ/b}$ ,  $n = 2$ ).

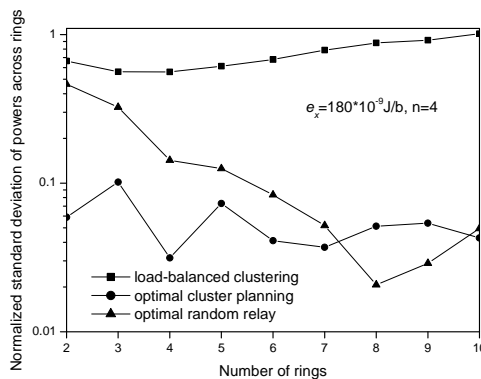


Fig. 13. Normalized standard deviation of power consumption vs.  $K$  ( $e_x = 180\text{nJ/b}$ ,  $n = 4$ ).

closer to that of optimal cluster planning. In this situation, the latter scheme has an extra benefit in optimally organizing its clusters, thus achieving better energy performance.

In Figures 12 and 13, we study via simulations the effect of balancing the powers across different rings. We measure the effectiveness in the power balance using  $\eta \stackrel{\text{def}}{=} \frac{\text{Std}(P_{\text{avg},1}, \dots, P_{\text{avg},K})}{\text{Avg}(P_{\text{avg},1}, \dots, P_{\text{avg},K})}$ . The smaller the value of  $\eta$ , the more balanced is power consumption across different CHs (and the larger is the coverage time). The figures indicate that in most cases, our analysis-based optimization of the radius profile and relay probabilities leads to a small  $\eta$  (e.g., less than 0.1). However, Figure 13 shows that for a small  $K$  and  $n = 4$ , the optimal random relay scheme exhibits a relatively large  $\eta$  (comparable with the value of  $\eta$  for LB clustering). This can be explained by noting that for a small  $K$ , the length of each CH-to-CH hop is considerably larger than the distance between the sink and a CH in the first ring. Under a highly nonlinear channel attenuation model ( $n = 4$ ), even if  $\alpha_{i,0} = 0$  (i.e., no traffic is sent directly to the sink), the power consumption for CH-to-CH relaying is still much larger than the power consumption of a CH in the first ring. Consequently, no power balance can be reached in this scenario. As we increase  $K$ , the distance of each hop decreases, so the power tradeoff between relay and direct transmission becomes dominant in the optimization, leading to a better power balance.

## V. CONCLUSIONS AND FUTURE WORK

We considered the problem of coverage-time optimization by balancing power consumption at different CHs in a clustered WSN. Stochastic as well as deterministic network models were investigated in our analysis. Our study demonstrates the significance of simultaneously accounting for the impacts of intra- and inter-cluster traffic in the design of routing and clustering strategies. For the deterministic-topology scenario, we presented a joint clustering/routing optimization based on linear programming. For the stochastic scenario, two mechanisms for balancing power consumption were studied: the (routing-aware) optimal cluster planning and the (clustering-aware) optimal random relay. The control parameters in both mechanisms (radius profile and relay probabilities) were optimized with respect to the maximum power consumption of a CH. The optimization problems were formulated as signomial optimizations and linear optimization, which were efficiently solved using generalized geometric programming and linear programming, respectively. For tractability purposes, our analysis for the stochastic model is necessarily approximate, as it relies on several simplifying assumptions. Simulations were conducted to verify the adequacy of this analysis and demonstrate the substantial benefits of the proposed schemes in terms of prolonging the coverage time of the network.

For simplicity, in our simulations we assumed a TDMA-like MAC. The implications of various types of MACs (e.g., CSMA/CA, TDMA, hybrid TDMA/CDMA, etc.) on our algorithms is an important issue and will be investigated in our future work. We will also consider extending the analysis to hierarchically clustered WSNs (e.g., the “spine” hierarchy).

## REFERENCES

- [1] IEEE 802.15 wpan task group 4 (tg4). available online at [www.ieee802.org/15/pub/TG4.html](http://www.ieee802.org/15/pub/TG4.html).
- [2] Zigbee alliance. available online at [www.zigbee.org](http://www.zigbee.org).
- [3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, pages 102–114, Aug. 2002.
- [4] A. D. Amis and R. Prakash. Load-balancing clusters in wireless ad hoc networks. In *Proc. 3rd IEEE Symposium on Application-Specific Systems and Software Engineering Technology*, pages 25–32, 2000.
- [5] A. D. Amis, R. Prakash, T. H. P. Vuong, and D. T. Huynh. Max-min d-cluster formation in wireless ad hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, pages 32–41, 2000.
- [6] D. J. Baker and A. Ephremides. The architectural organization of a mobile radio network via a distributed algorithm. *IEEE Transactions on Communications*, 29(11):1694–1701, Nov. 1981.
- [7] S. Bandyopadhyay and E. J. Coyle. An energy efficient hierarchical clustering algorithm for wireless sensor networks. In *Proceedings of the IEEE INFOCOM Conference*, volume 3, pages 1713–1723, 2003.
- [8] S. Bandyopadhyay and E. J. Coyle. Minimizing communication costs in hierarchically-clustered networks of wireless sensors. *Elsevier Journal of Computer Networks*, 44(1):1–16, Jan. 2004.
- [9] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.
- [10] C. F. Chiasserini, I. Chlamtac, P. Monti, and A. Nucci. An energy efficient method for nodes assignment in cluster-based ad hoc networks. *ACM/Baltzer Journal of Wireless Networks*, 10(3):223–231, May 2004.
- [11] C. F. Chiasserini and M. Garetto. Modeling the performance of wireless sensor networks. In *Proceedings of the IEEE INFOCOM Conference*, 2004.
- [12] B. Das, R. Sivakumar, and V. Bharghavan. Routing in ad hoc networks using a spine. In *Proceedings of IEEE International Conference on Computer Communications and Networks*, pages 34–39, 1997.

- [13] J. G. Ecker. Geometric programming: methods, computations and applications. *Society of Industrial and Applied Mathematics, SIAM Review*, 22(3):338–362, 1980.
- [14] A. Ephremides, J. E. Wieselthier, and D. J. Baker. A design concept for reliable mobile radio networks with frequency hopping signaling. *Proceedings of IEEE*, 75(1):56–73, 1987.
- [15] M. Gerla and J. Tsai. Multicluster, mobile, multimedia radio network. *ACM/Baltzer Journal of Wireless Networks*, 1(3):255–265, 1995.
- [16] G. Gupta and M. Younis. Load-balanced clustering of wireless sensor networks. In *Proceedings of the IEEE ICC Conference*, volume 1848–1852, 2003.
- [17] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4):660–670, Oct. 2002.
- [18] J. M. Kahn, R. H. Katz, and K. S. Pister. Next century challenges: mobile networking for smart dust. In *Proceedings of the ACM MobiCom Conference*, pages 271–278, 1999.
- [19] F. Kuhn, T. Moscibroda, and R. Wattenhofer. Initializing newly deployed ad hoc and sensor networks. In *Proceedings of the ACM MobiHoc Conference*, 2004.
- [20] C. R. Lin and M. Gerla. Adaptive clustering for mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 15:1265–1275, Sept. 1997.
- [21] T. Moscibroda and R. Wattenhofer. Maximizing the lifetime of dominating sets. In *Proceedings of IEEE WMAN Conference*, 2005.
- [22] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen. Topology control for wireless sensor networks. In *Proceedings of the ACM MobiCom Conference*, pages 286–299, 2003.
- [23] S. J. Park, R. Vedantham, R. Sivakumar, and I. F. Akyildiz. A scalable approach for reliable downstream data delivery in wireless sensor networks. In *Proceedings of the ACM MobiHoc Conference*, pages 78–89, 2004.
- [24] E. L. Peterson. Geometric programming. *Society of Industrial and Applied Mathematics, SIAM Review*, 18(1):1–51, Jan. 1976.
- [25] P. J. Wan, K. M. Alzoubi, and O. Frieder. Distributed construction of connected dominating set in wireless ad hoc networks. In *Proceedings of the IEEE INFOCOM Conference*, 2002.
- [26] J. Wu and H. L. Li. On calculating connected dominating set for efficient routing in ad hoc wireless networks. In *Proc. of the 3rd ACM International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 7–14, 1999.

## APPENDIX: GENERALIZED GEOMETRIC PROGRAMMING

A function  $h$  is a *monomial* in the variables  $x_1, x_2, \dots, x_n$  if it can be written as  $h(x_1, \dots, x_n) = x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$  for any real-valued exponents  $a_1, \dots, a_n$ . Furthermore, a function  $f$  is a *posynomial* in the variables  $x_1, x_2, \dots, x_n$  if it can be written as

$$f(x_1, \dots, x_n) = \sum_{j=1}^L c_j g_j(x_1, \dots, x_n) \quad (41)$$

where for  $j = 1, \dots, L$ ,  $c_j \geq 0$  and  $g_j$  is a monomial in  $x_1, x_2, \dots, x_n$ .

Let  $\mathbf{x} \stackrel{\text{def}}{=} (x_1, x_2, \dots, x_n)$  be a vector of  $n$  variables and let  $M_1$  and  $M_2$  be any two positive integers. A standard geometric program is an optimization problem of the form:

$$\begin{cases} \min f_0(\mathbf{x}) \\ \text{s.t.} \\ f_i(\mathbf{x}) \leq 1, \quad i = 1, \dots, M_1 \\ h_l(\mathbf{x}) = 1, \quad l = 1, \dots, M_2 \end{cases} \quad (42)$$

where  $f_0, f_1, \dots, f_{M_1}$  are posynomials in  $\mathbf{x}$  and  $h_1, \dots, h_{M_2}$  are monomials in  $\mathbf{x}$ .

A geometric program in the standard form is not a convex optimization problem. However, with the change of variables  $y_i \stackrel{\text{def}}{=} \log x_i$  and  $b_i \stackrel{\text{def}}{=} \log c_i$ , it can be transformed into the

following convex form:

$$\begin{cases} \min \left\{ p_0(\mathbf{y}) \stackrel{\text{def}}{=} \log \sum_j \exp(\mathbf{a}_{0j}^T \mathbf{y} + b_0) \right\} \\ s.t. \\ p_i(\mathbf{y}) \stackrel{\text{def}}{=} \log \sum_j \exp(\mathbf{a}_{ij}^T \mathbf{y} + b_i) \leq 0, \quad i = 1, \dots, M_1 \\ q_l(\mathbf{y}) \stackrel{\text{def}}{=} \mathbf{a}_l^T \mathbf{y} + b_l = 0, \quad l = 1, \dots, M_2 \end{cases} \quad (43)$$

where  $\mathbf{a}_{ij} = (a_{ij1}, a_{ij2}, \dots, a_{ijn})^T \in \mathbf{R}^n$  is the exponent vector of the  $j$ th monomial in the  $i$ th posynomial and  $\mathbf{y} \stackrel{\text{def}}{=} (y_1, \dots, y_n)^T$  is the optimization variable. The logarithm of a sum of exponentials is a convex function. Thus, (43) is a convex optimization problem that can be efficiently solved using numerical algorithms such as the interior point method [9].

A signomial is a more generalized form of a posynomial, whereby the coefficients  $c_j$ ,  $j = 1, \dots, L$ , can have any real values. If in (42) the constraints consist of signomials, the formulation is called a *signomial program* or *generalized geometric programming*.

Any signomial program can be transformed into an equivalent program of the form

$$\begin{cases} \min g_0(\mathbf{x}) \\ s.t. \\ g_k(\mathbf{x}) \leq 1, \quad k = 1, \dots, p \\ g_k(\mathbf{x}) \geq 1, \quad k = p + 1, \dots, q \end{cases} \quad (44)$$

where  $g_k(\mathbf{x})$  is a posynomial for  $k = 0, 1, \dots, q$ . The form (44) is called a reversed posynomial program.

One approach for solving signomial problems is to “condense” the posynomial in each reversed constraint (i.e., approximate the sum of monomials by using their geometric average, leading to another monomial) and obtain a posynomial program that approximates the original signomial program. Upon solving the posynomial program by any convex optimization algorithm, the solution is used to generate a better approximation. For example, suppose a program  $\mathbf{S}$  of the form (44) contains a single reversed constraint

$$g_l(\mathbf{x}) \geq 1. \quad (45)$$

Let  $\hat{g}_l(\mathbf{x})$  be the monomial obtained by condensing  $g_l$  with an arbitrary set of weights  $\epsilon$  using the arithmetic-geometric mean inequality. Let  $\hat{\mathbf{S}}$  denote the program obtained from  $\mathbf{S}$  where (45) is replaced by

$$\hat{g}_l(\mathbf{x})^{-1} \leq 1. \quad (46)$$

Since  $\hat{g}_l(\mathbf{x})$  is a monomial, (46) is a standard posynomial constraint and  $\hat{\mathbf{S}}$  is a posynomial program that approximates the signomial program  $\mathbf{S}$ . Moreover, the arithmetic-geometric inequality implies that  $\hat{g}_l(\mathbf{x}) \leq g_l(\mathbf{x})$ . Thus, if  $\mathbf{x}$  is feasible for  $\hat{\mathbf{S}}$ , then it is feasible for  $\mathbf{S}$ . The minimum value for  $\hat{\mathbf{S}}$ ,  $M(\hat{\mathbf{S}})$ , is an upper bound on the minimum value for  $\mathbf{S}$ ,  $M(\mathbf{S})$ . Suppose that  $\hat{\mathbf{x}}$  is optimal for  $\hat{\mathbf{S}}$ . Define a new set of weights

$$\epsilon_i = \frac{f_i(\hat{\mathbf{x}})}{g_l(\hat{\mathbf{x}})}. \quad (47)$$

Using these weights, one can define a new condensed posynomial  $\check{g}_l(\mathbf{x})$  and form the program  $\check{\mathbf{S}}$  where  $\check{g}_l(\mathbf{x}) \geq 1$  replaces  $g_l(\mathbf{x}) \geq 1$  in  $\mathbf{S}$ . Since  $\check{g}_l(\hat{\mathbf{x}}) = g_l(\hat{\mathbf{x}})$  and  $\hat{\mathbf{x}}$  is feasible for  $\mathbf{S}$ , it follows that  $\hat{\mathbf{x}}$  is feasible for  $\check{\mathbf{S}}$ . The minimum value for  $\check{\mathbf{S}}$ ,

$M(\check{\mathbf{S}})$ , therefore satisfies

$$M(\mathbf{S}) \leq M(\check{\mathbf{S}}) \leq M(\hat{\mathbf{S}}). \quad (48)$$

This defines an iterative process for generating a sequence of posynomial programs whose minimum values are monotonically decreasing upper bounds of the desired minima of  $\mathbf{S}$ .

Since  $\mathbf{S}$  is non-convex in general, it may have local minima that are not global minima and the above process may converge to such a point. Additional efforts have been made in the literature to enhance the above algorithm so that it converges to a global minima for non-convex signomial programs. The detailed algorithmic description of signomial programming is out of the scope of this work. A comprehensive survey on algorithms for generalized geometric programming is given in [13].

PLACE  
PHOTO  
HERE

**Tao Shu** received the B.S. and M.S. degrees in electronic engineering from the South China University of Technology, Guangzhou, China, in 1996 and 1999, respectively, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2003. Currently he is a Ph.D. student at the electrical and computer engineering department at the University of Arizona, Tucson, USA. His research interests include resource allocation in wireless cellular and sensor networks, optimization of physical and MAC layers in wireless

communication systems, security analysis for wireless networks, and queueing theory.

PLACE  
PHOTO  
HERE

**Marwan Krunz** is a professor of electrical and computer engineering at the University of Arizona. He directs the wireless and networking group in the ECE Department. He is also the UA site director for Connection One, a joint NSF/state/industry IUCRC cooperative center that focuses on RF and wireless communication systems and networks. At present, the center consists of five participating universities and 17+ industrial affiliates. Dr. Krunz received his Ph.D. degree in electrical engineering from Michigan State University in 1995. He joined the University of Arizona in January 1997, after a brief postdoctoral stint at the University of Maryland, College Park. He previously held visiting research positions at INRIA, HP Labs, University of Paris VI, and US West (now Qwest) Advanced Technologies. His research interests lie in the fields of computer networking and wireless communications. His current research is focused on cognitive radios and SDRs; distributed radio resource management in wireless networks; channel access and protocol design; MIMO and smart-antenna systems; UWB-based personal area networks; energy management and clustering in sensor networks; media streaming; QoS routing; and fault monitoring/detection in optical networks. He has published more than 150 journal articles and refereed conference papers, and is a co-inventor on US patents. M. Krunz is a recipient of the National Science Foundation CAREER Award (1998). He currently serves on the editorial boards for the IEEE Transactions on Mobile Computing and the Computer Communications Journal. He previously served on the editorial board for the IEEE/ACM Transactions on Networking (2001-2008). He was a guest co-editor for special issues in IEEE Micro and IEEE Communications magazines. He served as a technical program chair for various international conferences, including the IEEE WoWMoM 2006, the IEEE SECON 2005, the IEEE INFOCOM 2004, and the 9th Hot Interconnects Symposium (2001). He has served and continues to serve on the executive and technical program committees of many international conferences and on the panels of several NSF directorates. He gave several tutorials and participated in various panels at premier wireless networking conferences. He is a consultant for a number of companies in the telecommunications sector.

communication systems, security analysis for wireless networks, and queueing theory.