Analytical Modeling and Mitigation Techniques for the Energy Hole Problem in Sensor Networks

Jian Li and Prasant Mohapatra
Department of Computer Science
University of California, Davis
Email: {lijian, prasant}@cs.ucdavis.edu

Abstract

In this paper we investigate the problem of uneven energy consumptions in a large class of many-to-one sensor networks. In a many-to-one sensor network, all sensor nodes generate constant bit rate (CBR) data and send them to a single sink via multihop transmissions. This type of sensor networks has many potential applications such as environmental monitoring and data gathering. Based on the observation that sensor nodes sitting around the sink need to relay more traffic compared to other nodes in outer sub-regions, our analysis verifies that nodes in inner rings suffer much faster energy consumption rates (ECR) and thus have much shorter expected lifetime. We term this phenomenon of uneven energy consumption rates as the "energy hole" problem, which may result in severe consequences such as early dysfunction of the entire network. We proposed an analytical modeling for this problem, which can help understand the relevance of different factors on energy consumption rate. Using this model, we study the effectiveness of several existing approaches towards mitigating the "energy hole" problem, including deployment assistance, traffic compression and aggregation. We have used simulation results to validate our analysis.

Index Terms

Energy hole problem, many-to-one communication model, sensor networks, uneven energy consumption rate.

I. Introduction

Sensor networks have become a hot research topic in recent years [1], [2], [3]. Sensor networks are formed of a number of nodes that are deployed for monitoring specific activities, such as environmental monitoring, battlefield monitoring, and construction distortion detection. These nodes have limited resources in terms of computation power, memory, battery power, and transmission capability. Among these issues, the energy problem is of key concern. Low power devices can be used in order to reduce energy consumption. Low-power algorithms and protocols for sensor networks are also under intensive research investigation. Some efforts have been directed to the fundamental lifetime limits in sensor networks [4], [5], [6], [7]. In this paper, we will focus on the problem of uneven energy consumptions in a large class of many-to-one sensor networks.

First of all, what is a many-to-one sensor network? This categorization relates to traffic patterns in sensor networks. According to their traffic patterns, sensor networks can be divided into two major classes, namely, one-to-many and many-to-one networks. In a one-to-many sensor network, the sensing data obtained at one sensor node is disseminated into the network for multiple interested receivers. A more generic application scenario of the one-to-many communication model would be a many-to-many network, where multiple sensing nodes generate and disseminate their data into the network. In other words, in a many-to-one network, traffic flows between random pairs of source and destination nodes. In a many-to-one network, traffic from all sensor nodes is directed to a *single* sink (basestation) for further processing. Many-to-one sensor networks have various applications such as data gathering, monitoring and surveillance [7], [17], [18].

In very large scale monitoring sensor networks, clustering technique has been proposed to address routing scalability and energy conservation issues [23], [24]. Sensor nodes send data to a local cluster head, which may do some data fusion and aggregation and then forward the data to a central sink for further processing. In such cluster-based sensor networks, each individual cluster is a many-to-one subsystem. Meanwhile, cluster heads and the central sink form a upper level many-to-one communication model. Thus our work is not only applicable to flat, many-to-one sensor networks, but also to hierarchical, cluster-based large scale sensor networks.

Although in-network information processing and in-network reasoning are expected to take important roles in large scale sensor networks (e.g. see [25], [26], [27]), some form of central functionality such as decision-making and action commanding is required in many applications, which normally depend on a many-to-one communication model to collect information from the network. Thus, it is very important to understand the characteristics of energy consumption model in many-to-one sensor networks.

Energy conservation in sensor networks has two aspects. First, both the devices and the protocols or algorithms in use should be highly energy efficient, which means to operate with the limited source of battery power. Second, ideally, energy consumption rates in different parts of the network should be even or almost even. Thus all nodes throughout the network area have about the same lifetime. Otherwise, some parts of the network may die much sooner than the others. In some cases as illustrated in Section III-A, if some critical parts of the network run out of battery early, it may lead to early dysfunction of the entire network, even if the other parts of the network still have a lot of residual energy.

In this paper, we are more concerned with the second aspect. We analyze the severe problem of uneven energy consumption rates in many-to-one sensor networks. In particular, sensor nodes around the sink suffer much faster energy dissipation rates, which we term as the "energy hole" problem. We develop an analytical model to investigate this problem, and identify key factors that contribute to the uneven energy consumption rates. Based on the characteristics of the "energy hole" model, we study the effectiveness of several existing approaches in the literature towards mitigating the "energy hole" problem, including mobile sink, deployment assistance, and traffic compression and aggregation. Some of these techniques have been proposed to address networking problems such as clustering and routing scalability as well as energy conservation. However, here we attempt to investigate the effectiveness of these approaches for solving the "energy hole" problem based on our analytical model. We have done extensive simulations on the performance of different approaches, and the results verify our analysis. A preliminary version of our work was reported in [28]. In this paper, we extend our earlier work in several aspects. We add more details in the development of our analytical model. We take more metrics into consideration when analyzing the effectiveness of different approaches. We also present more new simulation results to verify our analysis.

The organization of the rest of this paper is as follows. Preliminary issues and related assumptions are given in Section II. In Section III, we describe the "energy hole" problem and propose our analytical modeling. The effectiveness analysis of several existing approaches to address this problem are presented in Section IV, followed by Section V on simulation validation. Related work is discussed in Section VI. The paper is then concluded in Section VII.

II. Preliminaries

In this section we introduce some background knowledge in sensor network models, sensor node and system lifetime, the energy model we adopt as well as the assumptions we have made to facilitate the discussion of this paper.

A. Sensor Network Models

An in-depth taxonomy on sensor network models can be found in [22]. Here we only discuss the categorization of sensor networks based on their traffic patterns and traffic source characteristics.

Based on traffic patterns, sensor networks can be divided into one-to-many and many-to-one models, as discussed in the previous section.

We can also do the classification of sensor networks based on other features such as traffic source characteristics. Considering the traffic source characteristics, sensor networks can be categorized into two major classes: clock-based and trigger-driven sensor networks. In clock-based sensor networks, each active sensor node continuously generates constant bit rate data and sends it to a basestation(s) for further processing. In trigger-driven sensor networks, only when certain event happens, like a tank entering the battlefield, or a query being broadcasted, the sensor nodes are triggered by the event of interest and reacts accordingly.

When taking both traffic source characteristics and traffic patterns into account, we can have different combinations. For example, a sensor network is used in battleground to detect enemy tank intrusion, and alarming information needs to be disseminated to many soldiers. This scenario falls into the trigger-driven, many-to-many category. Consider another example, a continuous temperature monitoring network falls into the clock-based, many-to-one category. Our work in this article is focused on the latter type of sensor networks.

B. System Lifetime

System lifetime of a sensor network is concerned with the time period in which the network can maintain its desired functionality, such as maintaining enough connectivity, covering entire area, or keeping miss rate below a certain level. Note that system lifetime is related to, but different from nodal lifetime. Nodal lifetime is the lifetime of individual sensor nodes. It depends on both given battery capacity and energy consumption rate.

System lifetime of a sensor network has different definitions based on the desired functionality. It may be defined as the time till the first node dies. It may also be defined as the time till a proportion of nodes die. If the proportion of dead nodes exceeds a certain threshold, it may result in uncovered sub-regions, and/or network partitioning. The location of the failure nodes is also of importance. If the proportion of nodes that have run out of battery are located in some critical part of the network, e.g., connecting the central sink and the rest of the network, it may result in early dysfunction of the entire network. Although it is not our intention to give a formal definition of sensor network lifetime in this paper, our discussion in the rest of this paper should be taken in the spirit of the second "definition."

We would like to point out that we are not intended to proposed a new technique which is more energy efficient for sensor networks. Instead, we are interested in the uneven energy consumption issue in different parts of the network. Our goal is to develop an analytical model for this problem. This model can show what kinds of factors are significant to the uneven energy consumption issue. We also investigate the effectiveness of two generic approaches in existing literature, namely, hierarchical deployment assistance, and data compression and aggregation, in efforts towards alleviating the "energy hole" problem.

C. Assumptions

To facilitate the discussion in the rest of this paper, we make some reasonable assumptions on the class of many-to-one sensor networks, as follows:

- In a clock-based many-to-one sensor network, each sensor node continuously generates constant bit rate (CBR) data (b bits/sec) and sends to a common sink through multihop shortest routes.
- Nodes are uniformly and randomly distributed, so the node density is uniform throughout the entire network:

$$p = \frac{N}{A_{net}},$$

  where N is the total number of sensor nodes and A_{net} is the coverage area of the sensor network.
- Sensor nodes have the same, fixed transmission range of r meters.
- Ideal MAC layer, i.e., transmission scheduling is perfect such that there is no collision and restransmission.
- Sensor nodes use a location based greedy forwarding approach to transmit data packets to the sink. Quite a few such techniques have been proposed (for example, see [19]). In greedy forwarding, data packets are transmitted to a next-hop which is closest towards the destination.

- Initially the network is well connected. The problem of what node density can ensure network connectivity is investigated in [35].

D. Energy Model

A typical sensor node comprises of three basic units: sensing unit, processing unit, and transceivers. For the energy model, we consider power for sensing, power for receiving and power for transmitting. The processing energy is not accounted for here, which depends on the computation hardware architecture and the computation complexity.

The energy consumption formulas that we use in the analysis and simulations throughout the rest of this paper are as follows:

$P_{Sense} = \alpha_1 b$,

$P_{Tx} = (\beta_1 + \beta_2 r^n) b$,

$P_{Rx} = \gamma_1 b$,

where b (in bits/sec) is the data rate of each sensor node. According to [36], the term r^n accounts for the path loss, and the typical value for n is 2 or 4. According to [37], some typical values for the above parameters in current sensor technologies are as follows:

$\alpha_1 = 60 \times 10^{-9} J/bit$,

$\beta_1 = 45 \times 10^{-9} J/bit$,

$\beta_2 = 10 \times 10^{-12} J/bit/m^2$ (when $n = 2$),

or, $\beta_2 = 0.001 \times 10^{-12} J/bit/m^4$ (when $n = 4$),

$\gamma_1 = 135 \times 10^{-9} J/bit$.

III. The "Energy Hole" Problem and its Characterization

In many-to-one sensor networks, data from all sensor nodes are transmitted to the basestation through multihop routes. This multihop relaying results in the problem of uneven energy consumption. In this section we will analyze this problem mathematically and validate the model through simulation. Different approaches to address this problem are presented in Section IV.

A. Description of the Problem

First, let us see an example illustrated in Figure 1. A homogeneous sensor network is uniformly and randomly deployed in an $L \times L$ square area, where $L = M \times r$ meters. We assume the single sink node S is located in the center. We can divide the whole area into $\frac{M}{2}$ concentric bands with a step size of r meters. As in Figure 1, ring 0 is the small circle with radius r meters, and ring 1 is the shaded band with radius $2r$ meters. Since a greedy shortest hop routing policy is assumed to be in use, data packets from outer area hop from ring to ring towards the sink node S. Note that here we assume a packet can traverse each ring using only one hop transmission, although in reality a packet may be transmitted more than one time within the territory of a single ring.

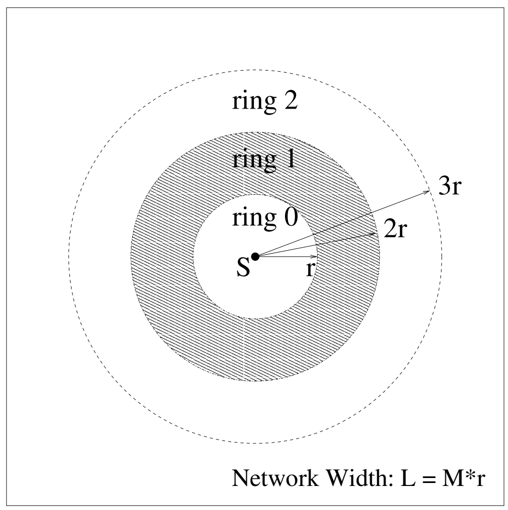

Fig. 1. The existence of "energy hole" around the sink node

Intuitively, the nodes around the sink have to relay more traffic compared to nodes that are farther away from the sink. Let us do a few calculations to analyze the details. Recall that $p = \frac{N}{A_{net}}$ is the node density, and b is the per node bit-rate. Note that any data traffic generated from the outer rings has to reach a node in ring 0 first, and then the node in ring 0 relay the data to the basestation. Besides this relaying traffic, a node in ring 0 also needs to transmit its own sensed data. So, the per node traffic load in ring 0 is:

$$\begin{aligned} Load_{ring\,0} &= \frac{\text{total traffic in the network}}{\text{num of nodes in ring 0}} \\ &= \frac{p(Mr)^2 b}{p\pi r^2} \\ &= \frac{M^2}{\pi} b. \end{aligned} \tag{1}$$

We would like to point out that the above calculation is based on our assumption of a perfect MAC layer (i.e., no collision and retransmission). Here we also assume that the wireless transceivers of nodes that are not on the transport path would be in *sleep mode* (for example, see [10]) which has very little energy consumption. Only the destinated next-hop node will be active and receive the data packet.

Similarly, we can obtain the per node traffic load in the other rings:

$$\begin{aligned} Load_{ring\,1} &= \frac{\text{total traffic from outside ring 0}}{\text{num of nodes in ring 1}} \\ &= \frac{p\left((Mr)^2 - \pi r^2\right)b}{p\left(\pi(2r)^2 - \pi r^2\right)} \\ &= \frac{(\frac{M^2}{\pi} - 1)}{3} b, \end{aligned} \tag{2}$$

and more generally,

$$\begin{aligned} Load_{ring\,i^{th}} &= \frac{p\left((Mr)^2 - \pi(ir)^2\right)b}{p\left(\pi\left((i+1)r\right)^2 - \pi(ir)^2\right)} \\ &= \frac{(\frac{M^2}{\pi} - i^2)}{2i+1} b,\ where\, i = 0,\ 1,\ ...,\ (\frac{M}{2} - 1). \end{aligned} \tag{3}$$

From these formulas we can observe that there is considerable difference between the per node traffic load in different rings. The analytical results for the two cases where $M = 8$ and $M = 6$ are illustrated in Figure 2. Note that here we call it "normalized" traffic load simply because we use the amount of sensing data generated by a single node as the "unit". So the nodes in ring 3 have traffic load as one unit because they are in the most outer ring and do not need to relay traffic for other nodes.

In both cases, the per node traffic load in ring 0 is three times higher than that of ring 1. The ratios between ring 0 and remote rings (ring 2 or ring 3) are even greater. Since wireless transmission/reception is the major source of energy dissipation, the nodes in inner rings are expected to consume much more energy than nodes in the outer rings.

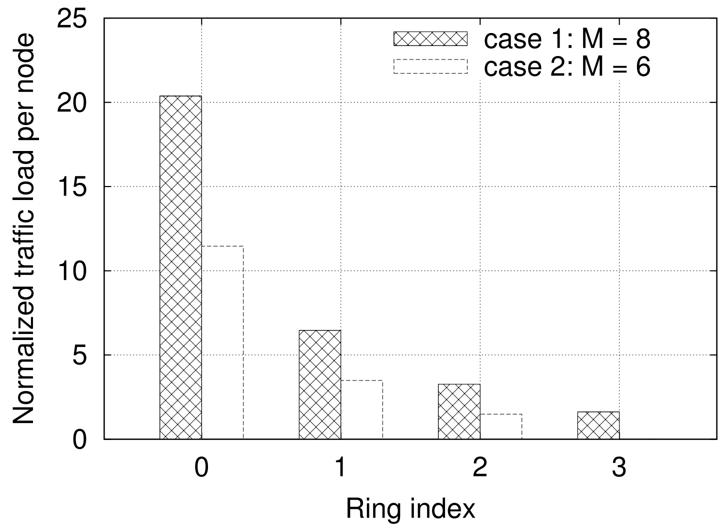

Fig. 2. Per node traffic load in different rings (when $M = 8$ or $M = 6$)

B. Characterization of the Model

The nodes in ring 0 have to relay the traffic from outer rings, in addition to sensing and transmitting their own data. The per node relaying load in ring 0 is:

$$\begin{aligned} RelayLoad_{ring\,0} &= \frac{p\left((Mr)^2 - \pi r^2\right)b}{p\pi r^2} \\ &= (\frac{M^2}{\pi} - 1)b. \end{aligned} \tag{4}$$

So the per node energy consuming rate (ECR) in ring 0 is:

$$ECR_{ring\,0} = \alpha_1 b + \gamma_1(\frac{M^2}{\pi} - 1)b + (\beta_1 + \beta_2 r^n)\frac{M^2}{\pi}b. \tag{5}$$

which accounts for three parts of energy consumptions for a single node in ring 0: sensing b amount of data, receiving $(\frac{M^2}{\pi} - 1)b$ amount of data from nodes in ring 1, and transmitting $\frac{M^2}{\pi}b$ amount of data to the sink. Similarly, we can derive the per node energy consumption rates in other rings:

$$ECR_{ring\,1} = \alpha_1 b + \gamma_1\frac{(\frac{M^2}{\pi} - 4)}{3}b + (\beta_1 + \beta_2 r^n)\frac{(\frac{M^2}{\pi} - 1)}{3}b, \tag{6}$$

and more generally,

$$\begin{aligned} ECR_{ring\, i^{th}} &= \alpha_1 b + \gamma_1 \frac{\left(\frac{M^2}{\pi} - (i+1)^2\right)}{2i+1} b + (\beta_1 + \beta_2 r^n) \frac{(\frac{M^2}{\pi} - i^2)}{2i+1} b, \\ where\, i &= 0,\, 1,\, ...,\, (\frac{M}{2} - 1). \end{aligned} \tag{7}$$

In order to verify our analytical model, We have done simulations with a 2000×2000 meters network area, which is the case where $M = 8$. In our analytical model, we assume that a perfect MAC layer is in use and that sensor nodes may run in sleep mode to conserve energy when they are not on transport path. In our simulations, we do not consider energy consumption due to multiple reception issue. We vary the node density in a vast range, and run the simulation with different random seeds for multiple times under each node density. Each run of simulation last for 2000 seconds. Finally we obtain the average results over all simulation runs of all node densities.

Both the calculated and simulated energy consumptions in different rings are shown in Figure 3. We observe that the simulated results match well with the analysis results, although all simulation results are a bit higher. We believe the reason behind this slight increase is that, in analysis we always assume an ideal case that a packet is transmitted from one ring to another using one hop. In simulation, however, greedy geo-forwarding technique [19] is used to route data packets to the sink. In such a manner, a packet may travel more than one hop within a single ring in some situations. Just as expected, nodes in ring 0 consumes much more energy compared to nodes in the outer rings. We term this phenomena as the "energy hole" problem.

Due to the "energy hole" problem, the nodes in ring 0 will expect a much shorter lifetime compared to the nodes outer rings, given that all mobile nodes are equipped with the same battery source. What is even worse, once the nodes in ring 0 are depleted of energy, the sink is disconnected from the rest of the sensor network!

Remarks In the above analysis, we consider a square area and the sink node is located in the center. We would like to point out that, this problem exists in other scenarios as long as the sink has a fixed location. Additionally, although we assume a clock-based network (i.e., each sensor node generates CBR traffic), the result is also valid in event-driven many-to-one sensor networks. If the probability of event

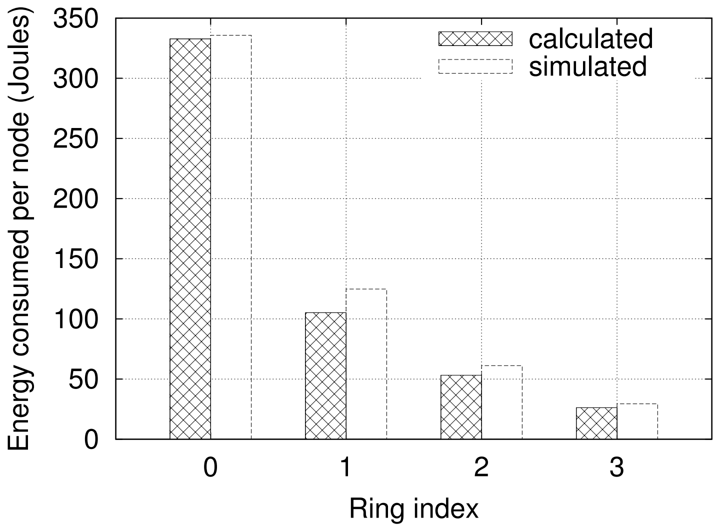

Fig. 3. Energy consumptions in different rings (when $M = 8$)

occurrence is uniformly and randomly distributed across the entire network and the sink is stationary, we can derive the same phenomena over a long time scale.

IV. Different Energy Conservation Approaches

In the existing literature, several approaches have been proposed toward energy conservation in sensor networks. In this section, we study the effectiveness of these approaches in addressing the "energy hole" problem.

The main reason for the "energy hole" problem is that some nodes have to relay a lot of traffic for other nodes in multihop transmissions. To avoid the multihop relaying, we can use a mobile sink, which has the capability to move around to collect data from sensor nodes. The alternative of a mobile sink would be a virtually moving sink [20], [21]. The main advantage of these approaches is the flexible control on what and how sensing data is collected. It is easy to see that the mobile sink approach requires additional memory/storage in individual sensor nodes. The process may also incur longer delay to collect all data before a snapshot of the entire network area can be formed, which is not desirable for real time monitoring in many applications. In this paper we will not include further discussion on this approach.

In the following discussion, we will focus on two energy conservation approaches, deployment assistance, and traffic compression and aggregation.

A. Deployment Assistance

In practice, we can exploit deployment assistance in order to overcome the "energy hole" problem. In addition to low power devices and energy efficient protocols, deployment policy is a third major method to prolong system lifetime, especially when considering the conflict between off-the-shelf devices and application-specific lifetime requirements. First, different types of applications, or different situations of the same type of application, may have quite different lifetime requirements. Second, it is more cost efficient and thus more promising to build general (or multi-purpose) sensing devices, although it is possible to manufacture application specific sensing devices. Third, it is reasonable to envision that we will have *heterogeneous* devices, in terms of battery source, wireless bandwidth, transmission range, and/or processor speed. Given such off-the-shelf, multi-purpose, and heterogeneous devices, in order to achieve a variety of application-specific lifetime requirements, it is a natural idea to exploit different deployment policies when designing and building different sensor networks.

We can divide the large network area into small sub-regions by using a two-tier architecture. Two-tier architecture has been proposed in existing literature for energy conservation. For example, TTDD [33] attempts to use geographical grids to exploit high node densities. Using hierarchical *deployment* can also smooth out the uneven energy consumption rates in a sensor network. This approach of deployment assistance is also different from previous work based on clustering (chain forming) [23], [24] in the sense that it relies on deployment rather than clustering.

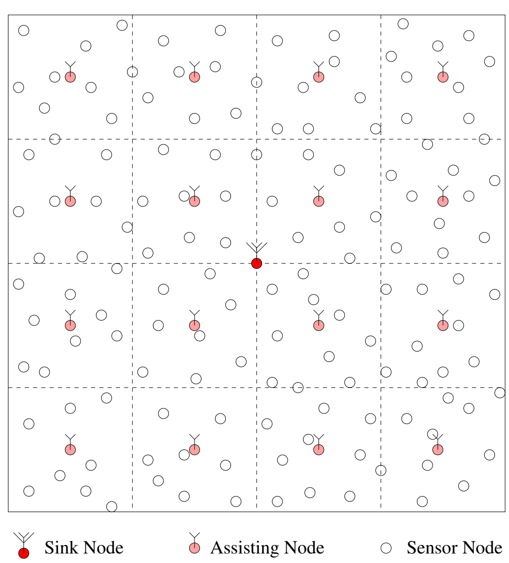

Fig. 4. An example of two-tier grid-based architecture

An example architecture is shown in Figure 4. We assume that there are two classes of nodes, namely, normal sensors and assisting nodes. Assisting nodes have much higher (compared to normal sensor nodes) battery capacities and have a larger transmission range. We can deploy a number of assisting nodes to form a relay layer on top of the normal sensors in the network. The trade-off here is to use some high capacity relaying nodes to help construct a large scale monitoring sensor network with sensing nodes which have very limited battery source. The relay network formed by the assisting nodes may use single hop or multihop transmission when sending data to the sink. We just take the relay layer as a blackbox for now. Normal sensors send their data to a closeby assisting node, and assisting nodes are in charge of forwarding the data to the basestation. Since the number of hops within each grid are smaller, the nodes sitting in ring 0 sub-regions in the grids have much lighter relaying burden compared to a flat network architecture. Hence the problem of "energy hole" is expected to be greatly alleviated.

In Figure 4, all assisting nodes are shown to be located in the centers of the small grids. However, considering the ad hoc deployment of sensor networks in practice, it is hard to deploy assisting nodes exactly where we want. To investigate the impact of position deviation of assisting nodes, we introduce a new variable in our simulation - PER (position error ratio), which is defined as:

$$PER = \frac{\text{distance from actual position to the ideal center}}{\text{transmission range}} \tag{8}$$

In our simulations, we will use various PER values representing different degrees of position deviation.

Remarks Besides hierarchical deployment, other deployment approaches include non-uniform deployment and incremental deployment. In non-uniform deployment, sensor nodes with higher battery capacity are deployed in inner rings. In incremental deployment, we can obtain a snapshot of current network energy map through techniques such as eScan proposed in [32], and then decide where to add new nodes as needed. An equivalent variant to incremental deployment is selective over-deployment, which deploys redundant nodes to the inner rings. When the network is in operation, only a proportion of nodes in these over-deployed areas are active while the other nodes are in sleep mode. Since the nodes in inner rings take turn to be in active mode, energy consumption rates of individual nodes will be reduced. While all these deployment techniques can help prolong the system lifetime of a sensor network, our focus in this work is to investigate how to alleviate the "energy hole" problem under a given deployment structure.

B. Traffic Compression and Aggregation

We can also use data compression and traffic aggregation techniques to mitigate the "energy hole" problem. As the data packets are relayed from outer rings towards the sink in the center, each ring can exploit data redundancy and spatial correlation to aggregate and compress the traffic. Much work has been done on data compression and aggregation in sensor networks (e.g., [8], [9]). Our work on a wavelet-based approach for time series compression and dissemination in sensor networks was summarized in [11]. Specific compression and aggregation techniques and related issues, such as data accuracy and error variance, are beyond the scope of this paper.

Let us assume that nodes in each ring can obtain a compression ratio $\alpha < 1.0$ on both pass-by traffic and self-sensed data. Please note that this is a simplified assumption. In real applications, the compression (aggregation) ratio at each ring is highly related to the specific application and the routing scheme. Here we assume a fixed compression ratio just to facilitate the analysis and obtain some understanding of the effectiveness of this technique in mitigating the "energy hole" problem. Since network width is $L = M \times r$, the network area consists of $m = \frac{M}{2}$ rings (plus four corners). So, we can obtain such an approximation of per node load in ring 0:

$$\begin{aligned} Load_{ring\,0} &\approx \alpha^{m-1}D_{m-1} + \alpha^{m-2}D_{m-2} + ... + \alpha D_1 + b \\ &= b + \sum_{i=1}^{m-1} \alpha^i D_i, \end{aligned} \tag{9}$$

where

$$\begin{aligned} D_i &= \frac{p\left(\pi(ir)^2 - \pi(ir-r)^2\right)b}{p\pi r^2} \\ &= (2i-1)b,\ where\ i = (m-1),\ (m-2),\ ...,\ 1. \end{aligned} \tag{10}$$

The physical meaning of D_i is the per node relaying traffic in ring 0 without compression. In other words, it is the amount of traffic that is generated from the i^{th} outer ring and imposed on a **single** node in ring 0. (If there are k nodes in ring 0, $k \times D_i$ would be equal to the total amount of data that is generated by nodes in ring i.) So, substitute (10) into (9) and we get:

$$Load_{ring\,0} \approx b + \sum_{i=1}^{m-1} \alpha^i (2i-1) b. \tag{11}$$

Because of the fact that:

$$\begin{aligned}
& \sum_{i=1}^{t} \alpha^i (2i-1) \\
= \; & \sum_{i=1}^{t} 2(i+1)\alpha^i - 3\left(\sum_{i=1}^{t} \alpha^i\right) \\
= \; & 2\left(\sum_{i=1}^{t} \alpha^{i+1}\right)' - 3\left(\sum_{i=1}^{t} \alpha^i\right) \\
= \; & 2\left(\frac{\alpha^2(1-\alpha^t)}{1-\alpha}\right)' - \frac{3\alpha(1-\alpha^t)}{1-\alpha} \\
= \; & \frac{\alpha+\alpha^2-(2t+1)\alpha^{(t+1)}+(2t-1)\alpha^{(t+2)}}{(1-\alpha)^2} \\
\ll \; & t^2+2t \qquad \text{(when } t \geq 1\text{)},
\end{aligned} \tag{12}$$

we obtain:

$$\begin{aligned}
Load_{ring\,0} & \approx b + \sum_{i=1}^{m-1} \alpha^i (2i-1) b \\
& \ll b + \left((m-1)^2 + 2(m-1)\right) b \\
& = m^2 b \\
& < \frac{M^2}{\pi} b,
\end{aligned} \tag{13}$$

where the last term is the traffic load in ring 0 when there is no compression, as shown in Equation (1). That means, the relaying burden in ring 0 is greatly reduced when some form of compression is in use (i.e., when $\alpha < 1.0$).

Remarks In (9) i times of compression are applied to the traffic D_i that is generated from the i^{th} ring. Please note that this expression does not exactly reflect the physical process. Instead, in real applications, as the data packets transfer from outer rings towards the sink, they will be combined with locally-sensed data at each intermediate ring, and some form of compression will be applied to the aggregated data before it is forwarded to the next ring. This physical process is different from applying multiple times of compression to one copy of data directly. Intuitively, our analytical model attempts to mimic the hop by hop transmission process in the physical network.

V. Simulation Validation

We have done extensive simulations using a customized program to verify our analysis on effectiveness of the existing solutions, namely, hierarchical deployment, and traffic compression and aggregation. In this section we present the simulation setups and results.

In all the simulations we assume that MAC layer is ideal, i.e., there is no collision and retransmission which can result in extra energy consumption. Since our goal is to investigate the "energy hole" problem, we assume that each link always has enough capacity to transfer the data. In our simulations, we adopt the energy model described in Section II-D. If not stated otherwise, we use 250 meters as transmission range and $n = 4$ is chosen as the path loss factor.

A. Impact of Node Density

As observed from Equation (7), network density does not affect the energy consumption rates. To verify this feature, we have done simulations with a 2000×2000 meters network area. The number of nodes varies from 500, 600, 700, 1000, 1500 to 2000, which represents different node densities. For each node density, we run the simulation with different random seeds for multiple times. The bit rate is 2000 bits/second. Each run of simulation last for 2000 seconds. Finally we obtain the average results, as shown in Figure 5.

From Figure 5 we observe that, for each ring, the energy consumption stays at a steady level under different node densities. That is, per node energy consumption is independent of node density (assuming node density is adequate to guarantee network connectivity), which justifies our earlier statement that we cannot prolong network lifetime by simply deploying more nodes.

B. Impact of Hierarchical Deployment

To investigate the impact of hierarchical deployment, we run simulations with a 3000×3000 meters network. We use different division granularity by dividing the area into 1×1, 2×2, 3×3, and 4×4 grids. Correspondingly, the grid width is 3000, 1500, 1000, and 750 meters. Correspondingly, there are 0, 4, 9,

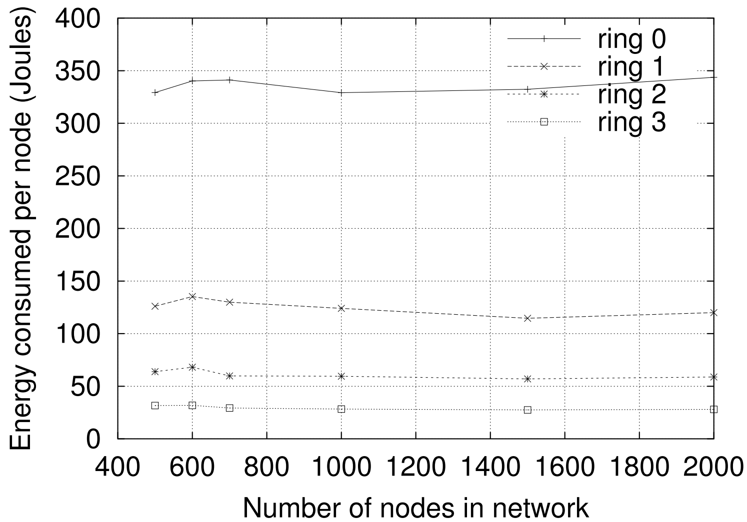

Fig. 5. Impact of different network node numbers

and 16 assisting nodes, plus the single sink node. Ideally, the assisting nodes are located in the centers of the grid areas. In practice, however, assisting nodes may be deployed to a place far away from its ideal position. As defined in Section IV-A, we use different PER values of 0%, 50%, and 150% in our simulation to represent such kind of position errors. Because sensor nodes always select the closest assisting node as its clusterhead, the network area is divided into a Voronoi diagram, as illustrated in Figure 6.

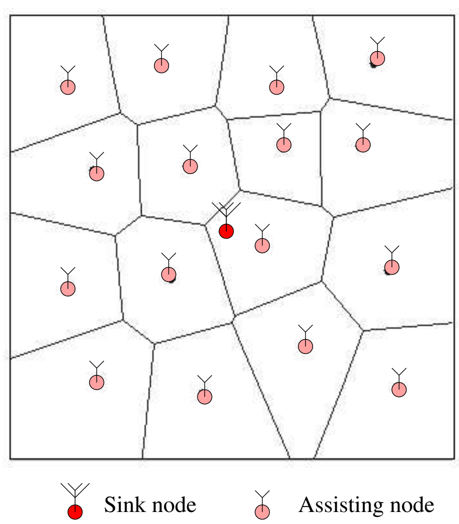

Fig. 6. Voronoi diagram due to deviated assisting nodes

For each kind of grid division, we do multiple runs of simulations with different node densities and random seeds and obtain average results, which are shown in Figure 7. We can observe that, under all the three PER values, the energy consumption rate in ring 0 is greatly reduced, even with a 2×2 division method.

We would like to point out that the energy consumption in ring 0 in Figure 7 is much higher than that in Figure 5. This is because we use a larger network width (3000 meters), compared to 2000 meters that is used in simulations for Figure 5.

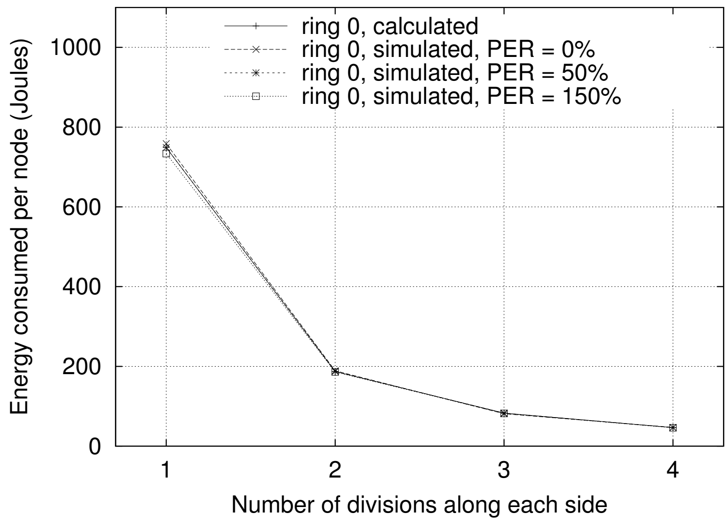

Fig. 7. Energy consumptions with sub-region division

Meanwhile, we observe in Figure 7 that the averaged per node energy consumption in ring 0 is almost the same for different values of PER. In other words, assisting nodes' deviation from ideal positions will not affect the average per node energy consumption in ring 0.

While the PER value does not affect the average per node energy consumption across the network, it indeed has effect on the Voronoi cell division as shown in Figure 6, which in turn will affect individual nodes' energy consumption rate. In order to understand the impact of position deviation, we measure the relative skewness of the consumed energy of different nodes in ring 0. We define the normalized energy skewness (NES) of per node energy consumption as the following:

$$NES = \sqrt{\frac{\sum_{i=1}^{k}(1 - \frac{E_i}{E_{avg}})}{k-1}}, \tag{14}$$

where E_i is each node's energy consumption, and E_{avg} is the average energy consumption over all k nodes in ring 0. The results are plotted as shown in Figure 8. We observe that, for each case of PER, NES decreases as the number of divisions increases. We also observe that the NES curve for PER = 150% case is above the other two cases. In summary, the larger the PER, the larger the NES. As each assisting node deviates further from the ideal grid center, the assisting network becomes more irregular which results in more uneven per node energy consumption rates.

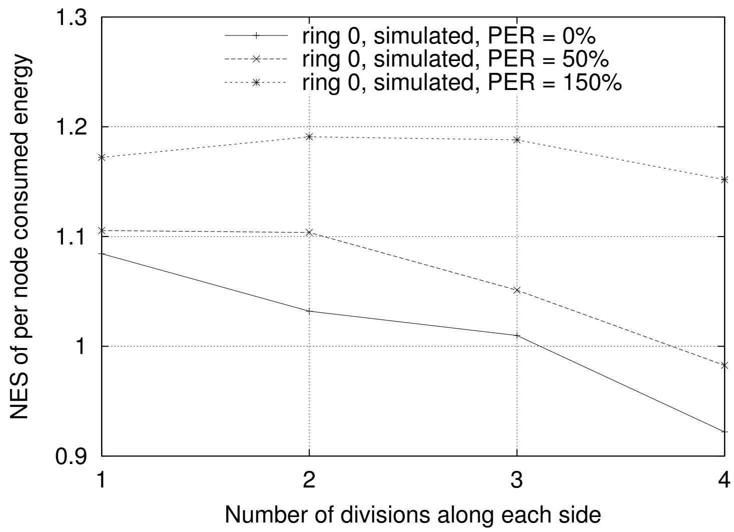

Fig. 8. Impact of PER on energy consumption rates

C. Impact of Source Bit Rate

From Equation (7), the energy consumption rate increases as the bit rate increases. To investigate the impact of source bit rate, we use a 2000×2000 meters network and vary the bit rates from 1000, 2000, 3000, to 4000 bits/second. For each bit rate, we run the simulation with different numbers of nodes from 500 to 2000. Each run of simulation last for 2000 seconds. The results are averaged over all runs of all scenarios.

Figure 9 shows the energy consumptions in different rings with varying bit rates. First, we can see that the simulated results match well with the analytical results. As in Figure 3, the simulation curves are always a bit higher than the calculated ones, and we believe this is due to the same reason: in simulation, a data packet may take more than one hop transmission to travel from ring i to ring $(i-1)$. In summary, this test shows that the expression in Equation (5), which is based on our simplification assumption, is a reasonably good approximation of the behavior of the system. Second, we observe that, as the bit rate increases, the energy consumption in ring 0 increases much faster than those in outer rings. This implies that, under the same network diameter, higher bit rates will worsen the “energy hole” problem.

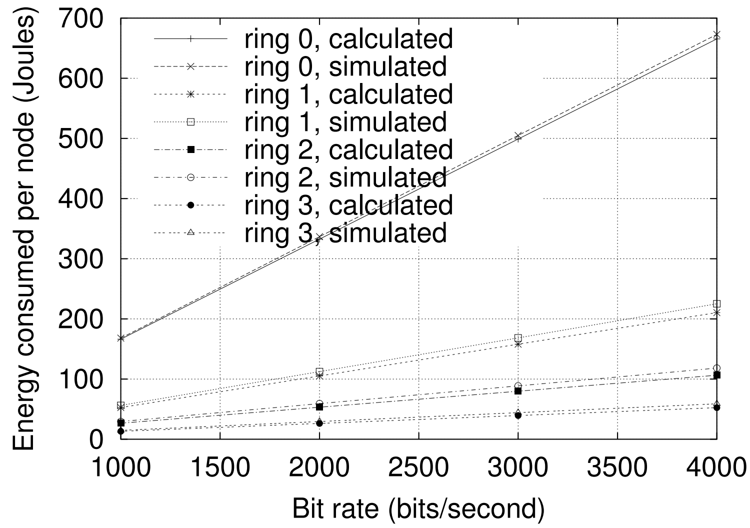

Fig. 9. Energy consumptions under different bit rates

D. Impact of Traffic Compression

In order to investigate the impact of traffic compression and aggregation, we use a 2000×2000 meters network with different node densities. The bit rate is 2000 bits/second and the packet size is 2000 bits. When a sensor node generates some data via sensing, or receives some data from other nodes, it will apply compression and aggregation technique to achieve a given compression ratio. We use 1.0, 0.9, 0.8, and 0.7 to represent different compression ratios, where a ratio equal to 1.0 means no compression is in use.

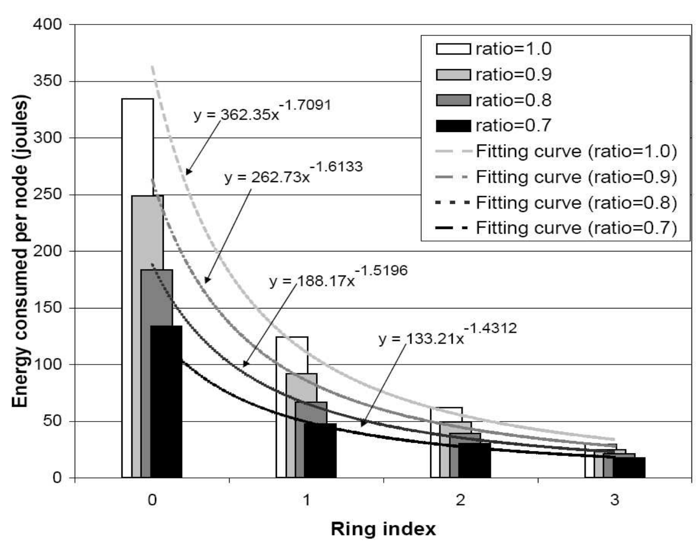

Fig. 10. Impact of traffic compression

The simulation results for different compression ratios are shown in Figure 10. We observe that, as the compression ratio increases, the energy consumption rate in each ring decreases. As shown in the figure, we do curve fitting for each compression ratio case. We observe that, as the compression ratio is reduced from 1.0, to 0.9, 0.8, till 0.7, the power index of the fitting curve decreases from 1.7091, to 1.6133, 1.5196, till 1.4312. We can say that, the greater the compression degree, the flatter the fitting curve. In other words, the decrease in ring 0 is relatively greater than that in the outer rings, which helps even out the consumption rates in different rings.

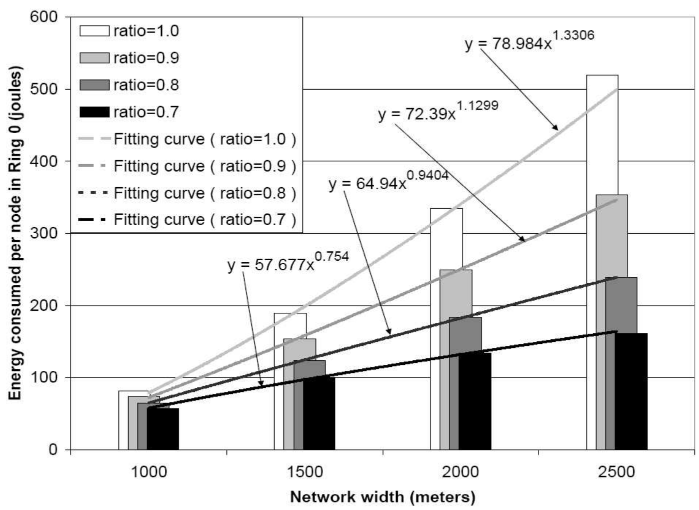

Fig. 11. Different compression ratios with different network sizes

We do further simulations to investigate the impact of compression ratio under different network sizes. The bit rate is fixed at 2000 bits/second. We vary the network width from 1000, 1500, 2000, to 2500 meters. For each network size, we do multiple runs of simulations with different node densities and random seeds, and obtain averaged results over all the runs. Simulation results are shown in Figure 11. We observe that, as the network width increases, the per node energy consumption rate increases under all compression ratios. When the compression degree is greater (with a smaller ratio), the acceleration is also smaller. Specifically, as the compression ratio is reduced from 1.0, to 0.9, 0.8, till 0.7, the power index of the fitting curve decreases from 1.3306, to 1.1299, 0.9404, till 0.754. With a relatively large compression degree, the network is more scalable in term of energy consumption rates in different rings, which justifies that compression and aggregation techniques can help alleviate the “energy hole” problem.

VI. Related Work

In this section we briefly discuss the related work on bounding the fundamental limits, capacity and lifetime, in ad hoc and sensor networks. We also talk about some related work on deployment assisted approaches in this field.

A. Bounding Lifetime

Bhardwaj et al have worked on upper bounds on the lifetime of sensor networks [4], [5]. In [4] the authors provided an analytical model for the lifetime issues based on trigger-based, many-to-one sensor networks. In [5], the authors further presented a role assignment technique in constructing the upper bounds on sensor network lifetime. To the best of our knowledge, these papers are among seminal efforts in this field. However, they do not identify the problem of uneven energy consumptions in many-to-one sensor networks.

A cell-based energy conservation technique was proposed in [33]. Nodes in the same cell are selectively turned on or off collaboratively in order to save energy. Blough et al investigated this technique's performance on energy conservation and lifetime extension [6]. Their simulation results showed that this cell-based technique can extend network lifetime greatly. In their work they assumed a uniform network density and random distributed peer-to-peer traffic, which is different from the many-to-one traffic pattern in our work.

In a more recent work [7], Duarte-Melo et al investigated extending sensor network lifetime by using hierarchical clustering technique. Based on a generic energy model and calculating mathematical expectation of sender-to-receiver distance, the authors gave MATLAB-based numerical results on estimated lifetime and optimal network cluster number. While the work in [7] is very similar to the deployment assistance approach in this paper, we use a totally different model for energy consumption analysis. Specifically, our work proposed an analytical model for the "energy hole" problem in the many-to-one traffic pattern. Our analytical reasoning attempts to mimic the hop by hop transmission in physical networks, and our simulation results were obtained with the widely-accepted NS-2 simulator.

B. Bounding Capacity

A lot of work has been done on bounding the capacity of ad hoc networks and sensor networks [12], [13], [14], [15], [16], [17]. All of these works assume that each node generates the same amount of data, i.e., the per node capacity is uniform throughout the entire network. Among them, [17] discussed the capacity issue in many-to-one sensor networks. The authors characterized the amount of data required to reconstruct a two-dimensional field, and the amount of data that can be transported per time slot in a sensor network. They also investigated the change rate of these variables as the number of nodes increases. Our work is complimentary to their results. Our analysis justified that in many-to-one sensor networks those nodes close to the sink have to relay more traffic than others in outer rings, and it would be a good idea to deploy more bandwidth capacity to inner sub-regions as needed. This type of architecture aware capacity planning needs further research efforts.

C. Deployment Assisted Approaches

Deployment assisted approaches have been previously proposed to improve the performance of ad hoc and sensor networks [29], [30], [31]. In [29], Ahmed et al proposed to deploy some assisting gateways in a mobile ad hoc network in order to provide better connectivity and facilitate scalability. Based on some assumptions they derived an approximate algorithm to compute the optimal positions where the gateways should be placed. In a more recent work [30], Ye et al proposed to deploy some reliable nodes in order to provide redundancy and better reliability in ad hoc routing protocols. Closer to our work, the authors in [31] investigated the infrastructure tradeoff in sensor network deployment.

VII. Concluding Remarks

In this paper we address the uneven energy consumption issue in sensor network. Our goal was not to develop a new energy conservation technique. Instead, we were to develop an analytical model for the “energy hole” problem in many-to-one sensor networks. Based on the understanding of the characteristics of the “energy hole” model, we study the effectiveness of various existing techniques towards mitigating this problem. These techniques can facilitate balanced energy consumption rates in different parts of a sensor network, and thus achieve more even lifetime across the network. Simulation results are used to verify our analysis and the proposed solutions.

We would like to point out that the “energy hole” problem is inherent in many-to-one sensor networks, and thus the best we can do is to reduce inner ring’s energy consumption and achieve more even energy consumption rates across different rings. Although many-to-one sensor networks suffer from the “energy hole” problem, it is a required communication model in many applications, as long as we want to facilitate some form of central decision-making functionality in the sensor networks.

In many-to-one sensor networks, energy consumption, lifetime planning and capacity planning are closely related to each other. One of our future work plan is to investigate possible combinations and trade-offs among these issues in sensor network design and deployment.

References

[1] D. Estrin, D. Culler, and K. Pister. Connecting the physical world with pervasive networks. IEEE Pervasive Computing, vol. 1, no. 1, pp. 59-69, Jan.-Mar., 2002.

[2] G. J. Pottie, and W. J. Kaiser. Wireless Integrated Network Sensors. Communications of the ACM, vol. 43, no. 5, pp. 51-58, May 2000.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cyirci. Wireless Sensor Networks: A Survey. Computer Networks, vol. 38, no. 4, pp. 393-422, Mar. 2002.

[4] M. Bhardwaj, T. Garnett, and A. P. Chandrakasan. Upper bounds on the lifetime of sensor networks. In Proc. of IEEE ICC 2001.

[5] M. Bhardwaj, A. P. Chandrakasan. Bounding the lifetime of sensor networks via optimal role assignments. In Proc. of IEEE InfoCom 2002.

[6] D. M. Blough, and P. Santi. Investigating upper bounds on network lifetime extension for cell-based energy conservation techniques in stationary ad hoc networks. In Proc. of IEEE/ACM MobiCom 2002.

[7] E. J. Duarte-Melo, M. Liu. Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks. In Proc. of IEEE GlobeCom 2002.

[8] B. Krishnamachari, D. Estrin, and S. Wicker. The impact of data aggregation in wireless sensor networks. In Proc. of International Workshop on Distributed Event-Based Systems, 2002.

[9] K. Kalpakis, K. Dasgupta, and P. Namjoshi. Maximum lifetime data gathering and aggregation in wireless sensor networks. In Proc. of the 2002 IEEE International Conference on Networking (ICN’02), Atlanta, Georgia, 2002.

[10] W. Ye, J. Heidemann, and D. Estrin. Medium access control with coordinated adaptive sleeping for wireless sensor networks. IEEE/ACM Trans. Netw., vol. 12, no. 3, pp. 493–506, 2004.

[11] H. Chen, J. Li, and P. Mohapatra. RACE: Time series compression with rate adaptivity and error bound for sensor networks. IEEE MASS 2004.

[12] P. Gupta, and P. R. Kumar. The capacity of wireless networks. IEEE Transactions on Information Theory, vol. 46, no. 2, March 2000.

[13] J. Li, C. Blake, D.S.J. De Couto, H.I. Lee, and R. Morris. Capacity of ad hoc wireless networks. In Proc. of IEEE/ACM MobiCom 2001.

[14] M. Gastpar, and M. Vetterli. On the capacity of wireless networks: the relay case. In Proc. of IEEE InfoCom 2002

[15] S. Toumpis and A. J. Goldsmith. Ad hoc network capacity. In Proc. ofAsilomar Conference of Signals, Systems and Computers, 2000.

[16] B. Liu, Z. Liu, and D. Towsley. On the capacity of hybrid wireless networks. In Proc. of IEEE InfoCom 2003.

[17] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff. On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. IPSN 2003, LNCS 2634, pp. 1-16, 2003.

[18] E. J. Duarte-Melo, and M. Liu. Data-gathering wireless sensor networks: organization and capacity. Computer Networks (COMNET) Special Issue on Wireless Sensor Networks, Vol 43, Issue 4, pp. 519-537, November 2003.

[19] B. Karp and H. T. Kung. GPSR: Greedy Perimeter Stateless Routing for Wireless Networks. Proc. of ACM MOBICOM 2000, pp.243-254.

[20] C. Gui, and P. Mohapatra. Virtual Patrol: A New Power Conservation Design for Surveillance Using Sensor Networks. In proceedings of 4th ACM/IEEE Conference on Information Processing in Sensor Networks (IPSN-2005), April 2005.

[21] W. Zhang, H. Song, S. Zhu, and Guohong Cao. Least Privilege and Privilege Deprivation: Towards Tolerating Mobile Sink Compromises in Wireless Sensor Networks. ACM MOBIHOC, May 2005.

[22] S. Tilak, N. Abu-Ghazaleh, and W. Heinzelman. A taxonomy of wireless micro-sensor network communication models. ACM Mobile Computing and Communication Review, Apr. 2002.

[23] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy efficient communication protocol for wireless micro-sensor networks. In Proc. of IEEE Hawaii Int. Conf. on System Sciences, Jan. 2000.

[24] S. Lindsey and C. Raghavendra. PEGASIS: Power-efficient gathering in sensor information systems. In Proc. of IEEE Aerospace Conference, March 2002.

[25] Y.J. Zhao, R. Govindan, and D. Estrin. Computing aggregates for monitoring wireless sensor networks. In Proc. of IEEE SNPA 2003.

[26] R. Kumar, V. Tsiatsis,and M. Srivastava. Computation hierarchy for in-network processing. In Proc. of ACM WSNA 2003.

[27] B.J. Bonfils, and P. Bonnet. Adaptive and decentralized operator placement for in-network query processing. In Proc. of IPSN 2003.

[28] J. Li and P. Mohapatra. An analytical model on the energy hole problem in many-to-one sensor networks. In Proc. of IEEE VTC Fall 2005 (to appear).

[29] M. Ahmed, S. Dao, and R. Katz. Positioning range extension gateways in mobile ad hoc wireless networks to improve connectivity and throughput. In Proc. of IEEE MilCom 2001.

[30] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi. A framework for reliable routing in mobile ad hoc networks. In Proc. of IEEE InfoCom 2003.

[31] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman. Infrastructure tradeoffs for sensor networks. ACM Workshop on Sensor Networks and Applications (WSN'02). Held in Conjunction with Mobicom'02.

[32] Y. Zhao, R. Govindan, and D. Estrin. Residual energy scans for monitoring wireless sensor networks. In Proc. of IEEE WCNC 2002.

[33] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang. A two-tier data dissemination model for large-scale wireless sensor networks. In Proc. of IEEE/ACM MobiCom 2002.

[34] L. M. Feeney, and M. Nilsson. Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In Proc. of IEEE InfoCom 2001.

[35] C. Bettstetter. On the minimum node degree and connectivity of a wireless multihop network. In Proc. of ACM MobiHoc 2002.

[36] T. Rappaport. Wireless communications: principles & practice. New Jersey: Prentice-Hall, Inc., 1996.

[37] W. Heinzelman. Application-specific protocol architectures for wireless networks. PhD thesis, Massachusetts Institute of Technology, 2000.