Energy Balanced Data Collection in Wireless Sensor Networks

Ning Jin, Kaiji Chen, Tao Gu
Department of Mathematics and Computer Science
University of Southern Denmark
Email: {njin,chen,gu}@imada.sdu.dk

***Abstract*—In this paper, we investigate the energy balanced data collection problem in WSNs, aiming to balance the energy consumption among all the sensor nodes in the data propagation process. Energy balanced data collection can potentially save energy consumption and prolong the network lifetime, and hence it has many practical implications for WSN design and deployment. The traditional hop-by-hop transmission model allows a sensor node to propagate its packets in a hop-by-hop manner towards the sink, resulting in poor energy balance for the entire network. To address the problem, we apply a slice based energy model, and divide the energy balanced data collection problem into inter- and intra-slice energy balance problems. We then propose a novel Inter-slice Mixed Transmission strategy and an Intra-slice Forwarding technique to address each of the problems. Finally, we design an Energy-balanced Transmission Protocol (ETP) to combine both techniques to achieve total energy balance in data collection. Through extensive simulation studies, we demonstrate that, while ETP achieves energy balanced data collection, the network lifespan is increased by 10 times and the network delay is decreased by more than 70% compared to the hop-by-hop transmission in a general square area WSN.**

I. Introduction

A Wireless Sensor Network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion and pollutants, and to cooperatively pass their data through the network to a base station (a.k.a. *sink*). Wireless Sensor Networks have been used in many application domains such as habitat monitoring [1], environmental surveillance [2], scientific observation [3], infrastructure management [4], health care [5], and etc.

The recent advances in RF communication and microelectronics have led to inexpensive, small sized, and battery-powered wireless sensors. These sensors typically have limited resources such as computational capacity, data storage, communication quality, and most crucially, their available amount of energy. Therefore, saving energy and prolonging the network lifetime is an important goal in designing various techniques in WSNs such as routing protocols [6], data collection [7], clustering [8], and duty cycle [9]. However, existing techniques do not take care of possible overuse of certain sensor nodes in the network. For example, in the hop-by-hop transmission strategy, sensor nodes lying closest to the sink tend to be utilized exhaustively since all the data pass through them. Thus, these sensor nodes may die out much more quickly than other nodes in the network, resulting in network collapse although there may be still significant amounts of energy in the nodes which are far away from the sink. In another example where each node in a WSN communicates directly to the sink, sensor nodes lying far from the sink consume their energy much faster than the nodes near the sink because transmitting data over a longer distance requires more energy (the energy needed to send a message directly from node u to node v is proportional to d^2, where d is the distance between u and v [10]–[12]). In addition, the imbalance use of the battery power in sensor nodes may lead to a practical issue for the real deployment and management of WSNs—the battery replacement has to be done frequently since the sensor nodes in the network drain their batteries at different rates.

We investigate the problem of energy balance during data collection in WSNs, aiming to balance the energy consumption among all the nodes in a network, and substantially increase the lifespan of the network. Several solutions [11], [13]–[15] have been proposed to address this issue. However, they essentially leverage on the following model—once a data message reaches a sensor node, it is either propagated one-hop (the neighbour nodes within its transmission range R) closer to the sink or it is sent directly to the sink. They basically combine the hop-by-hop transmission and the direct transmission, and apply a randomized mixed strategy to balance the energy consumption among nodes. However, this model is impractical for real applications since the range of a WSN will be limited by the transmission range of a sensor node. It may result that either a WSN only covers a very limited range or only the hop-by-hop transmission is allowed in a WSN intending to cover a large geographic area.

In the existing hop-by-hop transmission model, we investigate that nodes with different distances from the sink have unequal rates of energy consumption. In this paper, we propose a transmission model under which the sensor nodes are able to adjust their transfer range at m levels (i.e., transfer range can be $R, 2R, ..., mR$, where m is the maximum level, and R is the minimum transfer range). This model is more realistic and practical than both the direct transmission and the hop-by-hop transmission models. Under this model, we propose a probability-based transmission strategy. We partition a WSN area into n "slices" as shown in Figure 1 (following [11], [14]–[16]) according to the distance which each node is away

978-1-4673-2447-2/12/$31.00 © 2012 IEEE

from the sink. To balance the energy consumption during data collection, we aim to balance the energy consumption among these "slices", both inter-slice and intra-slice.

To achieve inter-slice energy balance, we propose a probability-based Inter-slice Mixed Transmission (IMT) strategy to let each sensor makes an opportunistic choice of its radio power level for sending its packets. A Linear Programming model is used to solve the transmission probabilities. Using this strategy, sensor nodes close to the sink can conserve their energy by propagating their packets using a lower power level (i.e., with a smaller hop, 1-hop for example), and sensor nodes far away from the sink tend to consume more energy by sending their packets using a higher power level (i.e., with a larger hop, m-hop for example) to achieve the balanced energy consumption among different slices. We analyze the necessary condition to achieve inter-slice energy balance through both mathematical analysis and numerical computation, and we show that the maximum level m (where m indicates the transfer limitation of each sensor node) should satisfy $m > 0.42n$ (where n is the network size) in a general sector-shaped network, and $m > \sqrt{2n}$ in a specific chain-shaped network. As compared to the traditional hop-by-hop data collection strategy, IMT has two main advantages: First, it allows longer distance transmission, resulting in less propagation delay; Second, since each node has a choice to select its transmission power, the total energy consumption can be spread evenly across the entire network.

We also analyze intra-slice energy consumption and discover that the energy of the nodes within one slice may drain at different rates (e.g. in a slice, some nodes may receive more packets from their children nodes than others). To achieve intra-slice energy balance, we propose an intra-slice forwarding technique which first allows nodes with lower energy to forward its packets to the nodes with higher energy in the same slice, and then the packets are propagated to the next slice towards the sink. By combining the inter-slice mixed-transmission strategy and the intra-slice forwarding technique, we design an Energy-balanced Transmission Protocol (ETP) to achieve total energy balance during data collection, and hence prolong the network lifetime.

To evaluate ETP, we uses two metrics—Relative Standard Deviation and Gini coefficient. Through comprehensive experiments, we show that 1) ETP can balance the energy consumption among all the nodes; 2) ETP prolongs the network lifetime by more than 10 times and decreases the transmission delay by more than 3 times as compared to the hop-by-hop data propagation method.

In summary, this paper makes the following contributions.

- We study the problem of energy balanced data collection in WSNs systematically, and propose a mixed transmission method in which sensor nodes are allowed to adjust their radio power to achieve total energy balance.
- We apply a Linear Programming model to study inter-slice energy balance, and propose an Inter-slice Mixed Transmission strategy to achieve energy balance across different slices. In addition, we find mathematically the necessary condition between the network size n and the maximum level m to achieve energy balance for two representative WSN topologies.
- We analyze the intra-slice energy balance problem, and propose an intra-slice forwarding technique to achieve intra-slice energy balance. By combining it with Inter-slice Mixed Transmission strategy, we design the Energy-balanced Transmission Protocol (ETP) to achieve total energy balance during data collection.
- We conduct comprehensive simulation studies and demonstrate that ETP achieves better energy balance, longer lifetime and less network delay, as compared to the existing methods.

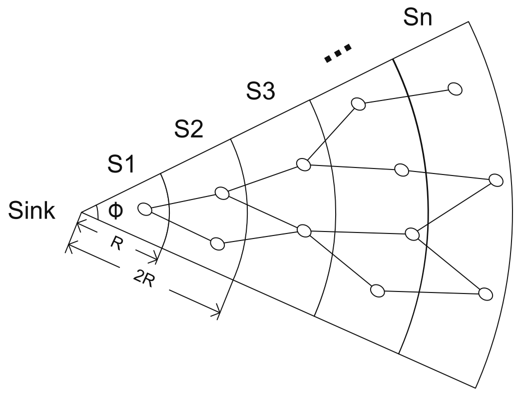

Fig. 1. Slice Model

The paper is structured as follows. Section II formally describes the problem and the "Slice" model which partitions the problem to be two parts, inter-slice and intra-slice. In section III, we study the inter-slice energy balance problem using a Linear Programming model. Section IV investigates the intra-slice energy balance problem. Section V presents the simulation results. Section VI describes the related work, and finally, Section VII concludes the paper.

II. Problem Definition

In this section, we will first describe the energy balance data collection problem in a WSN. We are given a region in which sensor nodes are scattered randomly and uniformly. Events may occur in the region and can be detected by the sensor nodes nearby. The function of the WSN is to collect and transfer the event data through the network to the sink (a sink is usually located at the the network edge). Before defining the problem, we make the following assumptions.

1) Sensor Deployment: We assume a homogeneous WSN where the same type of hardware and software is used. The sensor nodes are deployed in a geographic area randomly and uniformly.

2) Event Detection: Events occur randomly in the region, and each event results in the same amount of sensor data to be transferred to the sink.

3) Energy Consumption: Since events occur randomly and uniformly, we assume the energy consumed by detecting such an event is the same for all the nodes. In addition, we assume synchronous duty-cycled WSNs where sensor nodes wake up to transmit packets in the same period, and hence the energy spent in idle listening for each node is the same.

The energy consumption of a node in a WSN consists of the energy spent for event detection, sending/receiving packets, and idle listening, respectively. With the above assumptions, we focus our study on the energy consumption for sending and receiving packets only, which is believed the most important factor affecting the network lifespan [11], [17]. We consider both sending and receiving unlike the existing work [11], [14], [16] which typically ignore the energy consumption for receiving packets.

We define the energy balanced data collection problem as follows. During data collection, we aim to ensure all the sensor nodes in a WSN drain their batteries at the same rate in a relatively long period of time, in another word, all the sensor nodes will remain an equal energy level. In the remaining of this chapter, we describe the slice model.

A. Slice Model

We virtually "cover" the network area by a disk sector of an angle ϕ, as shown in Figure 1. The disk sector is divided into n ring sectors or "slices". The total number of slices in the WSN is denoted as n. This slice model can cover the entire network area, taking a proper large angle ϕ.

De nition Let $S_i(i = 1, 2, ..., n)$ be the i-th slice of the network. The first slice S_1 has radius R (i.e., the sensor s transmission range in the hop-by-hop network). Slice i $(i = 1, 2, ..., n)$ is defined by two successive disk sectors, one of radius iR and the other of radius $(i-1)R$. Particularly, we define S_0 as the *sink*.

By the slice model, we basically convert the euclidean distance between two sensor nodes into hop count. Sensors in S_i are i hop away from the $sink$ and $|i-j|$ hop away from the nodes in S_j.

De nition Let A_i be the area size of slice S_i $(i = 1, 2, ..., n)$ in the network.

As events occur in the network area randomly and uniformly, the probability that an event occurs in a slice is proportional to the slice s size.

De nition Let λ_i be the probability that an event occurs in the slice S_i. By assuming a random uniform generation of events in the network area, we have

$$\lambda_i = \frac{A_i}{\sum_{i=1}^{n} A_i} \tag{1}$$

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{2}$$

With the above model, we partition the WSN area into n "Slices", and then we aim to balance the energy consumption in two steps, inter-slice and intra-slice.

III. Inter-Slice Energy Balance

In this section, we study the inter-slice energy balance problem, aiming to balance energy consumption among different slices.

A. Inter-slice Mixed Transmission

In our model, each sensor node in the WSN is able to communicate with its neighbours within m-hops where m is the maximum transmission hop (level). In particular, the sensor nodes in the i-th slice can propagate their packets to any nodes in the next m slices (S_{i-1}, S_{i-2} to S_{i-m}) towards the sink. m is limited by the sensor hardware, and it can not be arbitrarily large. When $m = n$, every node can communicate directly to the *sink*. When $m = 1$, it is basically the hop-by-hop transmission.

Based on a probability, each node in the network determines its next slice to propagate its data packets towards the sink. The probabilistic decision doesn t conclude which node to propagate the packets to (different from routing), but it determines the hop distance the transmission should propagate.

De nition Let $p_{i,k}(i = 1, 2, ..., n, k = 1, 2, ..., m)$ be the probability that a sensor node belonging to S_i sends its packets to a sensor node belonging to the "next" slice S_{i-k} with k hops. We then have

$$\sum_{k=1}^{m} p_{i,k} = 1, \forall n = 1, 2, ..., n \tag{3}$$

B. Energy Consumption

For sending a packet, the longer the distance the packet propagates, the more energy it consumes. Existing studies in [10], [11] demonstrate that the energy required to send a packet directly from u to v is proportional to d^2, where d is the distance between u and v . Therefore, we define our energy consumption model for data transmission as follows. A sensor node consumes C_r energy units when receiving a data unit (e.g., a packet), and $k^2 \cdot C_s$ energy units when sending a data unit to its k-hop neighbour. For example, when a sensor node in slice S_i sends data to slice S_j $((i-j) = 1, 2, ..., m)$, it will consume $(i-j)^2 C_s$ energy units for sending a data unit, and C_r energy units for receiving a data unit.

C. Inter-slice Energy Balance Problem

We wish to solve the inter-slice energy balance problem, i.e., propagating data to the sink in a way that the "average" energy dissipation of each sensor node is the same. The "average" energy dissipation per sensor node is measured by the fraction of the total energy consumed by the sensor nodes in a given slice over the number of sensor nodes in that slice. In another word, each slice in the WSN "consumes" its energy at the same rate against the total energy in the slice, and eventually, all the slices drain their energy simultaneously.

De nition We define Z as the total number of events occur in the entire lifespan of a WSN. For simplicity, we assume each event generates an equal amount of data unit (i.e., one data unit per event).

The larger the number of events a WSN supports, the longer the network survives. Therefore, Z can be viewed as an indication for the network lifespan.

De nition We define $f_{i,j}$ as the total number of data units transmitting from S_i to S_j in the entire lifespan of a WSN.

With the ow count $f_{i,j}$ during the lifespan of the WSN, we can now calculate the total energy consumed by one slice. The energy consumed in receiving data by S_i is computed as follows.

$$\sum_{j=1}^{n} f_{j,i} \cdot C_r \tag{4}$$

and the energy consumed in sending data is computed as follows.

$$\sum_{j=0}^{n} f_{i,j} \cdot (i-j)^2 C_s \tag{5}$$

Since sensor nodes are randomly and uniformly spread in the network area, initially the available energy for a slice is proportional to its area size.

De nition We define b_i as the available energy in slice S_i, and b_i is proportional to A_i (γ is a constant).

$$b_i = \gamma \cdot A_i \tag{6}$$

To achieve inter-slice energy balance, all the slices in the network have to consume all their available energy at the end of the network lifespan, such that,

$$\sum_{j=1}^{n} f_{j,i} C_r + \sum_{j=0}^{n} f_{i,j}(i-j)^2 C_s = b_i, \forall i = 1,2,...,n \tag{7}$$

Solving the above problem is equivalent to find the probabilities $p_{i,k}$ for all $i = 1,2,...,n$ and $k = 1,2,...,m$ so that we achieve (7) at the end of the network lifespan.

D. LP Model

We observe that the inter-slice energy balance problem is very similar to the well-known max- ow problem in Linear Programming (LP). Therefore, we re-define this problem in LP formation. We then study the two typical WSN topologies, and we discuss the relation between maximum transmission range m and network size n. We first give the LP problem as follows.

De nition *The energy-balanced max-lifespan problem*

$$\text{Maximize Z} \tag{8}$$

S.T.

$$\sum_{j=1}^{n} f_{j,i} + Z \cdot \lambda_i = \sum_{j=0}^{n} f_{i,j}, \forall i = 1,2,...,n \tag{9}$$

$$\sum_{j=1}^{n} f_{j,i} C_r + \sum_{j=0}^{n} f_{i,j}(i-j)^2 C_s = b_i, \forall i = 1,2,...,n \tag{10}$$

$$f_{i,j} = 0, \forall i = 1,2,...,n \text{ and } (i-j > m \text{ or } j \geq i) \tag{11}$$

$$f_{i,j} \geq 0, \forall i = 1,2,...,n \text{ and } j = 1,2,...,n \tag{12}$$

The objective function Z is the total number of events in the entire lifespan of the WSN. If the events are supposed to be generated at a constant rate, maximizing the network lifespan is equivalent to maximizing Z. Then maximizing the lifespan is what we want to achieve in energy-restrained WSNs. (9) is the ow constraint, same as in a max- ow problem. The ow produced by the node itself plus the incoming ow should be equal to the outgoing ow of every node in the network. This constraint guarantees that every packet in the WSN will eventually be delivered to the sink. (10) is the energy balance constraint which we have defined in last section. It ensures that the WSN will finally achieve inter-slice energy balance. (11) is the transmission specification in our model. It limits the maximum transmission range for each node to be m-hop. Additionally, packets should not be propagated backwards from the sink and only transmission from a further slice to a closer slice is allowed. This constraint helps eliminating overhead of unnecessary transmission in the WSN. Lastly, (12) specifies that all the ows should be non-negative.

With a LP-solver, we can solve the *energy-balanced max-lifespan problem* with a result Z and all the variables $f_{i,j}$ for particular instances (where the parameters n, m, λ_i, b_i, C_r and C_s are all known). Then the ratio of the ow $f_{i,j}$ from S_i to S_j over the total ow out of S_i is just the expected proportion for the node in S_i to send the packets with $(i-j)$ hop. If the sensor nodes propagate their packets in a probability manner as the same ratio as the ow in the energy-balanced max-lifespan problem, the total energy consumed by the WSN will achieve energy balance as expected. Thus, we can figure out the probability from the following equation:

$$p_{i,k} = \frac{f_{i,i-k}}{\sum_{j=i-1}^{i-m} f_{i,j}}, \forall i = 1,2,...,n, k = 1,2,...,m \tag{13}$$

E. More Discussions

A hop-by-hop transmission model may hardly achieve energy balance since the sensor nodes closer to the sink are always supposed to transmit more packets than those lying far away from the sink. In our model, the result will be the same as the hop-by-hop model when the maximum transmission hop m is restricted by one-hop. Hence, there exists a relation between m and n such that in the certain condition, the inter-slice energy balance problem has a feasible solution. Finding this relation has many practical implications for WSN deployment and management, such as selecting the appropriate sensor with proper radio hardware to achieve energy balance data collection, and making a proper decision for routing protocols when m and n are determined by the application.

De nition For a given WSN, where the parameters except m are known, we define m_x as the smallest m so that the inter-slice energy balance problem has a feasible solution.

Lemma III.1. *If there exists a m_x so that the LP* ***energy-balanced max-lifespan problem*** *has a feasible solution, then for every $m > m_x$, the inter-slice energy balance problem has a feasible solution.*

Proof: Assume we have found a feasible solution $\mathcal{S}$ when $m = m_x$. Then for every other $m > m_x$, the solution $\mathcal{S}$ is also a feasible solution (not optimal) to the **energy-balanced max-lifespan problem**. Thus the **energy-balanced max-lifespan problem** has a feasible solution, then from (13), the corresponding energy balance problem has a feasible solution. ■

Lemma III.2. *If the* ***energy-balanced max-lifespan problem*** *has a feasible solution, the following inequations should hold.*

$$Z \leq \sum_{i=1}^{m} \frac{Cr \cdot Z \cdot \lambda_i + b_i}{i^2 C_s + C_r} \tag{14}$$

$$b_n \leq C_s \cdot Z \cdot \lambda_n \cdot m^2 \tag{15}$$

Proof: By summing up the equations in (9) for $i = 1, 2, ..., n$, we have

$$Z = \sum_{i=1}^{m} f_{i,0} \tag{16}$$

For each $i = 1, 2, ..., n$, by $(-C_r) \times (9) + (10)$, we get

$$\sum_{j=0}^{n} f_{i,j}[(i-j)^2 C_s + C_r] = C_r \cdot Z \cdot \lambda_i + b_i, \forall i = 1, 2, ..., n \tag{17}$$

For $i = 1, 2, ..., m$ and 17 divided by $i^2 C_s + C_r$, we get

$$\sum_{j=0}^{i-1} \frac{f_{i,j}[(i-j)^2 C_s + C_r]}{i^2 C_s + C_r} = \frac{C_r \cdot Z \cdot \lambda_i + b_i}{i^2 C_s + C_r}, \forall i = 1, 2, ..., m \tag{18}$$

$$f_{i,0} + \sum_{j=1}^{i-1} \frac{f_{i,j}[(i-j)^2 C_s + C_r]}{i^2 C_s + C_r} = \frac{C_r \cdot Z \cdot \lambda_i + b_i}{i^2 C_s + C_r}, \forall i = 1, 2, ..., m \tag{19}$$

with (12), we have

$$f_{i,0} \leq \frac{C_r \cdot Z \cdot \lambda_i + b_i}{i^2 C_s + C_r}, \forall i = 1, 2, ..., m \tag{20}$$

Summing up for $i = 1, 2, ..., m$ and from (16), we have

$$Z = \sum_{i=1}^{m} f_{i,0} \leq \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \lambda_i + b_i}{i^2 C_s + C_r} \tag{21}$$

Then we get (14).

In (9) and (10), for $i = n$, we have

$$Z \cdot \lambda_n = \sum_{j=n-1}^{n-m} f_{i,j} \tag{22}$$

$$\begin{aligned} b_n &= \sum_{k=1}^{m} f_{i,i-k} k^2 C_s \\ &\leq \sum_{k=1}^{m} f_{i,i-k} m^2 C_s \\ &\leq Z \cdot \lambda_n \cdot m^2 C_s \end{aligned} \tag{23}$$

Then we get (15). ■

We now study two typical WSN topologies, and analyze the relation between m and n.

1) Sector-shaped Network: We first consider a general sector-shaped network [11], [14]–[16]. As discussed in Section II, a sector-shaped network can cover the entire network area with a sufficiently large angle α, as shown in Figure 1. S_i is defined by two successive disk sectors, one of radius iR and the other of radius $(i-1)R$. We have

$$\begin{aligned} A_i &= \frac{\phi}{2}[(iR)^2 - (i-1)^2 R^2] \\ &= \frac{\phi}{2} R^2 (2i-1) \end{aligned}$$

especially,

$$A_1 = \frac{\phi}{2} R^2$$

Then the area of different slices hold

$$A_i = (2i-1) A_1, \forall i = 1, 2, ..., n \tag{24}$$

Since events occur randomly and uniformly, from (1)

$$\lambda_i = \frac{A_i}{\sum_{i=1}^{n} A_i} = \frac{2i-1}{n^2}, \forall i = 1, 2, ..., n \tag{25}$$

from (6) and (24), the energy for each slice hold

$$b_i = \frac{2i-1}{2n-1} b_n \tag{26}$$

From inequality (14) and (15), we get

$$\begin{aligned} Z &\leq \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \lambda_i + b_i}{i^2 C_s + C_r} \\ &\leq \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \lambda_i + \frac{2i-1}{2n-1} b_n}{i^2 C_s + C_r} \\ &\leq \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \lambda_i + \frac{2i-1}{2n-1} C_s \cdot Z \cdot \lambda_n \cdot m^2}{i^2 C_s + C_r} \\ &\leq \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \frac{2i-1}{n^2} + \frac{2i-1}{2n-1} \frac{2n-1}{n^2} C_s \cdot Z \cdot m^2}{i^2 C_s + C_r} \\ &\leq \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \frac{2i-1}{n^2} + \frac{2i-1}{n^2} C_s \cdot Z \cdot m^2}{i^2 C_s + C_r} \end{aligned}$$

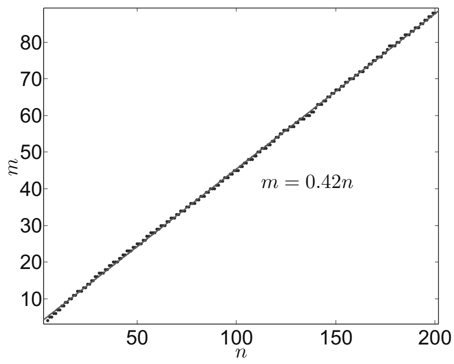

Fig. 2. Relation between m_x and n in sector-shaped network

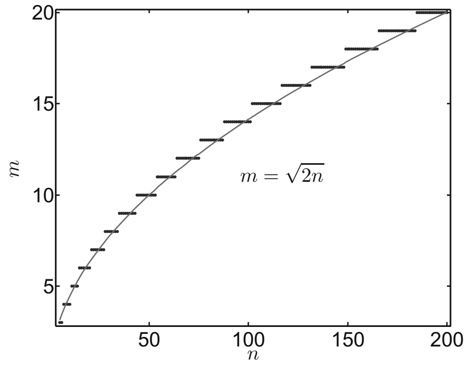

Fig. 4. Relation between m_x and n in chain-shaped network

$$n^2 \le \sum_{i=1}^{m} \frac{C_r \cdot (2i-1) + C_s \cdot (2i-1) \cdot m^2}{i^2 C_s + C_r}$$

$$n^2 \le \sum_{i=1}^{m} \frac{2i-1}{i^2 + C_r/C_s}(C_r/C_s + m^2) \tag{27}$$

In (27), when $C_r/C_s > 0$, the increase by $\sum_{i=1}^{m} \frac{2i-1}{i^2+C_r/C_s}$ is not sufficient large compared to m^2 [18] and we consider it as a small constant. Therefore, approximately, a linear relation exists between m_x and n. For a more precise result, we set $C_r/C_s = 1$ which means the cost in transmiting and receiving are equal in the one-hop transmission. Through experiment, we find the relation fitting to a line with $y = 0.42n$, as shown in Fig. 2. From the figure, we see that $y = 0.42n$ fit well to m_x with $m_x = 0.42n \pm 1$. This small uctuation is usually insignificant in large scale WSNs where $n \ge 100$. In practice, we can decide $m > 0.42n$ is the least requirement to achieve inter-slice energy balance.

2) Chain-shaped Network: We then consider a chain-shaped network, as shown in Fig. 3. Such networks are often deployed in applications such as structure management for bridges [19], environment monitoring for rivers [20] and mine-environment-based surveillance [21]. The area in such a network typically has a long narrow strip, and the size of each slice is almost the same. Thus, we have

$$A_i = A_j, \forall i, j = 1, 2, ..., n \tag{28}$$

Since events occur randomly and uniformly, we then have

$$\lambda_i = \frac{A_i}{\sum_{i=1}^{n} A_i} = 1/n \tag{29}$$

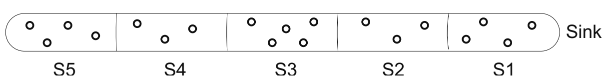

Fig. 3. Chain-shaped network

Since the available energy in each slice is proportional to its area, from (6), we have

$$b_i = b_j = b_n, \forall i, j = 1, 2, ..., n \tag{30}$$

From inequality (14) and (15), we get

$$\begin{aligned} Z &\le \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \lambda_i + b_n}{i^2 C_s + C_r} \\ &\le \sum_{i=1}^{m} \frac{C_r \cdot Z \cdot \lambda_i + C_s \cdot Z \cdot \lambda_n \cdot m^2}{i^2 C_s + C_r} \end{aligned} \tag{31}$$

with (29), we finally obtain

$$n \le \sum_{i=1}^{m} \frac{1}{i^2 + C_r/C_s}(C_r/C_s + m^2) \tag{32}$$

In (32), for $C_r/C_s > 0$ and we then have $\sum_{i=1}^{m} \frac{1}{i^2+C_r/C_s} \le \sum_{i=1}^{m} \frac{1}{i^2} \le \sum_{i=1}^{\infty} \frac{1}{i^2}$. From [18], we know that it will converge to a constant $\pi^2/6$. Then, approximately, a linear relation exists between m_x^2 and n. Similar to the sector-shaped network, we set $C_r/C_s = 1$ and use experiment to find out the approximate relation coefficient. As shown in Figure 4, we find the relation fitting to $y = \sqrt{2n}$ with a small error $m_x = \sqrt{2n} \pm 2$. In practice, we can use $m \ge \sqrt{2n}$ as the least requirement to achieve inter-slice energy balance in WSNs.

IV. Intra-Slice Energy Balance

We now focus on intra-slice energy balance in this section. We first observe the energy imbalance problem within a slice. To illustrate, we choose two nodes, A and B, in slice i where B is closer to the sink, as shown in Figure 5. From the figure, within the transfer distance R, node A has more adjacent neighbours from slice $i+1$. Therefore, node A has a higher chance than B for receiving the packets from its adjacent children slices, leading to energy consumption imbalance within slice i. When considering all the children slices (slice $i+1$ to slice $i+m$), we will obtain the similar result.

To address this issue, we propose an Intra-slice Forwarding technique which works as follows. First, each node keeps the

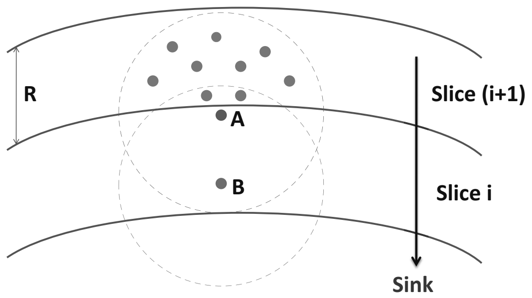

Fig. 5. Intra-slice Imbalance Problem Illustration Example

battery power level of its one-hop neighbours in the same slice. This information can be easily marked and maintained by using some bits in the control messages (note that these control packets will not effect energy balance because they cause the same amount of energy consumed at each node, and the overhead of these control packets can usually be ignored compared to data packets). When a sensor node receives a packet which can be either generated by itself or propagated from its adjacent slice, the node first compares its available energy to the average available energy in its slice. If it is higher, it will deliver the packet using the Inter-slice Mixed Transmission approach described in the previous section. If it is lower, the node will forward the packet to a neighbouring node with most available energy within the same slice, and then that node will forward the packet to the next slice. To avoid loop in transmission, we limit the number of times a packet can be forwarded to one within a slice, i.e., one packet can only be forwarded at most once in a slice.

V. Evaluation

We now move to evaluate ETP using simulation experiments. In this section, we first describe how we set up the simulation, and then present the experiment results.

A. Simulation Setup

Our simulation runs on two representative deployment areas – a square area and a chain-shaped area. We choose the square area to be $100m \times 100m$ and the chain-shaped area to be $200m \times 1m$. The sensor nodes are randomly deployed in the WSN region, there are 2000 nodes in the square area and 1000 nodes in the chain-shaped area. The sink is located in the border of the area, i.e., a corner for the square area, and a head for the chain-shaped area. The transfer range R is set to 10, and the maximum transfer range m of the node is set to 8. We assign each node an initial battery level of 10000 energy units and set $Cr = 1$ and $Cs = 1$. The simulation runs by generating event packets randomly on each node.

In the simulation, we evaluate and compare the following data collection approaches:

Hop-by-Hop Transmission only allows sensor nodes to propagate its packets to its one-hop neighbour towards the sink. This is the baseline solution.

Inter-slice Mixed Transmission (IMT) uses a series of probabilities computed from the Linear Programming model described in Section III. This technique is designed to achieve inter-slice energy balance.

Energy-balanced Transmission Protocol (ETP) combines IMT with the *Intra-slice forwarding* technique to achieve total energy balance during data collection in WSNs.

B. Network Lifetime

The lifetime of a WSN is the time span from the deployment to the instant when the network is considered non-functional [22]. When a network should be considered non-functional is, however, application-specific. Event delivery ratio is defined as the number of events successfully received by the sink above the total number of events occur in one period of time. We use this metric in this work, and define the lifetime of a WSN as the time when the event delivery ratio drops below 95%.

We plot the event delivery ratio for the square-area network in Figure 6. The result shows that the hop-by-hop transmission only propagates about 12k events, and then the event delivery ratio starts to drop rapidly, while both IMT and ETP maintain much longer lifespan. The event delivery ratio of IMT drops below 95% after about 141k events, and about 192k events for ETP, respectively. The event delivery ratio maintains quite well (more than 90%) and decreases very slowly even after the lifetime instant. For the chain-shaped area network, Figure 9 shows the result that the lifetime for the hop-by-hop transmission is about 7k events, 9k events for IMT, and 13k events for ETP, respectively. In such a network, ETP works much better in prolonging the network lifespan than IMT and the hop-by-hop transmission, and it is able to maintain a high event delivery ratio for a very long period of time.

C. Energy Balance

To measure energy balance during data collection in a WSN, we use two metrics —- Relative Standard Deviation (RSD) and Gini coefficient. In probability theory and statistics, relative standard deviation is a normalized measure of dispersion. RSD is defined as the absolute value of the ratio of the standard deviation to the mean. A low RSD value indicates that the data points tend to be very close to the mean, whereas a high RSD value indicates that the data points are spread out over a large range of values. Gini coefficient is also a measurement of statistical dispersion and often used in economy to measure the inequality of the income distribution of a country. The value of Gini is between 0 and 1, and the larger value implies higher inequality.

Figure 7 and Figure 10 shows the RSD values of energy consumption in the network lifetime. From the figures, the RSD values of IMT and ETP are much smaller than that of the hop-by-hop transmission, demonstrating better energy balance property. Figure 8 and Figure 11 show the Gini values of energy consumption in the two WSNs. Similar to RSD, the Gini results also demonstrate that IMT and ETP are much better than the hop-by-hop transmission, and ETP achieves the best energy balance among the three approaches. Through

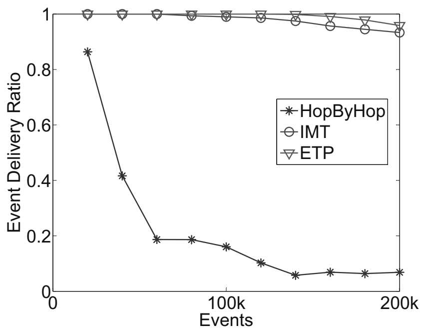

Fig. 6. Event delivery ratio in square area

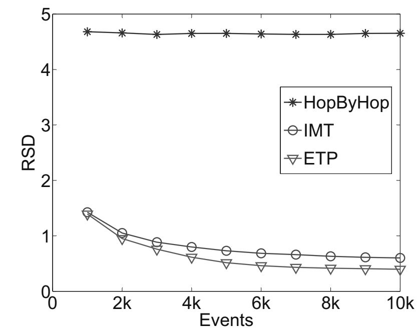

Fig. 7. Relative standard deviation of energy consumption against the time in square area

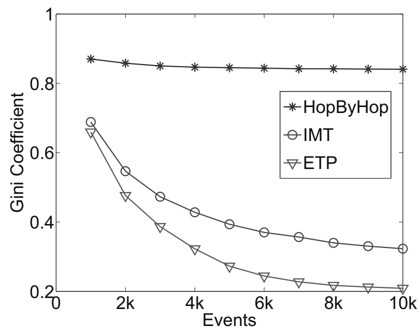

Fig. 8. Gini coefficient of energy consumption against time in square area

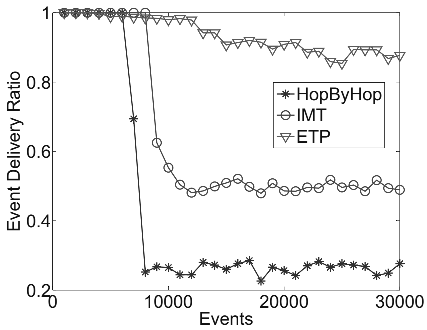

Fig. 9. Event delivery ratio in chain-shaped area

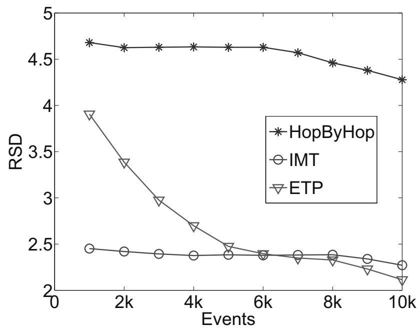

Fig. 10. Relative standard deviation of energy consumption against the time in chain-shaped area

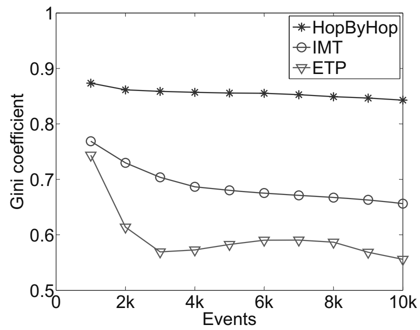

Fig. 11. Gini coefficient of energy consumption against time in chain-shaped area

this experiment, we show that IMT and ETP achieve much better energy balance during data collection than hop-by-hop transmission, resulting in a much longer lifespan.

D. Delay

We now evaluate and compare the packet delay for the three approaches. In this paper, we use the total number of hop count computed from the source node to the sink. In duty-cycled WSNs, hop count is often in positively correlation with the real time delay because sensor nodes need to wait the entire cycle to make one successful transmission. Therefore, less hop counts used in the transmission period implies smaller packet delay. Both Figure 12 and Figure 15 plot the Cumulative Distribution Function (CDF) of the hop count for all the successfully delivered packets in the network lifetime. From the figures, we can see in both square- and chain-shaped area networks, the hop-by-hop transmission has the largest hop count because the packets can only be propagated one-hop towards the sink for each transmission. With a mix transmission strategy, ETP and IMT result in much smaller hop count, and they have almost the same delay as expected.

Figure 18 shows that in the square area network, the hop-by-hop transmission has an average delay of 7.0 hop counts while IMT has 2.2 hop counts and ETP has 2.1 hop counts, respectively. The delay for ETP is less than one third of the hop-by-hop transmission. In the chain-shaped area network, the average hop count for the hop-by-hop transmission is about 10.0, and 7.2 hop counts for IMT and 6.6 hop counts for ETP, respectively. The delay is decreased by about 30%, as compared to that of the hop-by-hop transmission.

E. Energy Ef ciency

Figure 13 and Figure 16 plot the CDF of the remaining battery power at each node at the end of the network lifespan. The results clearly demonstrate that the energy utilization for the entire network by the hop-by-hop transmission is poor (i.e., the nodes closest to the sink drain their energy very fast and the nodes far away from the sink drain their energy very slow), while IMT and ETP effectively use the energy of the nodes far away from the sink by allowing longer transmission range to achieve energy balance. Figure 14 and Figure 17 show the energy utilization —the portion of the total energy utilized by the network during data collection — for the three approaches. The results demonstrate both ETP and IMT achieve much better energy utilization than the hop-by-hop transmission.

We also demonstrate the energy usage in each slice in the square-area WSN. Figure 19 plots the average energy consumed in data collection by each slice. From the figure, it is clearly shown that the hop-by-hop transmission overuses the sensor nodes close the to sink but fails to utilize the energy of the nodes in far away slices. On the other hand, both ETP and IMT use the energy from all the slices more evenly. Figure20

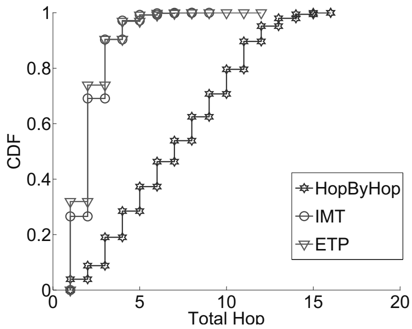

Fig. 12. The CDF of total hop in square area

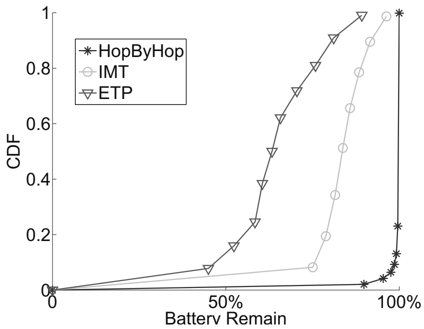

Fig. 13. Energy remain ratio at the end of the lifetime in square area

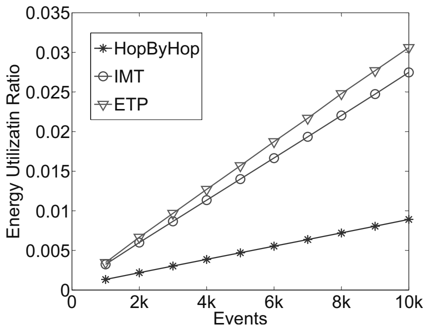

Fig. 14. Energy utilization in square area

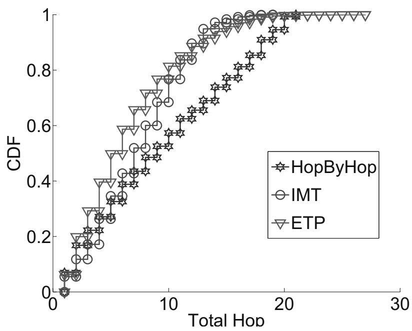

Fig. 15. The CDF of total hop in chain-shaped area

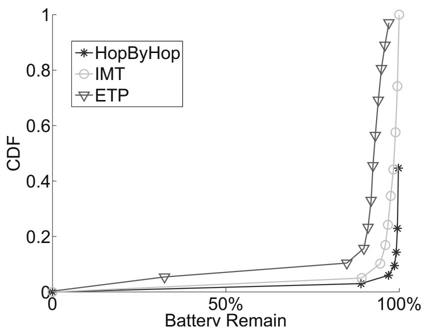

Fig. 16. CDF of energy remain ratio at the end of the lifetime in chain-shaped area

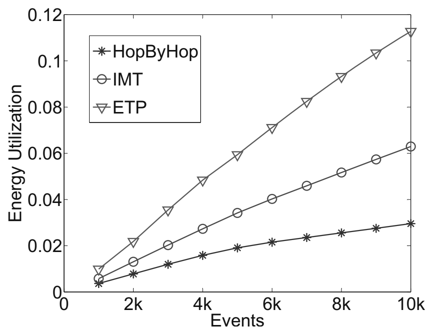

Fig. 17. Energy utilization in chain-shaped area

plots the standard deviation of energy consumption in each slice, and the results show that the hop-by-hop transmission tends to drain a few nodes in the slice which easily become the bottleneck, while both ETP and IMT keep a balanced use of each node in the slice. Furthermore, ETP works better to achieve energy balance inside each slice.

VI. Related Work

The energy balance problem in wireless sensor networks was first introduced in [23] where the authors studied the energy balance property, and proposed an energy-optimal and energy balanced algorithm for sorting in WSNs. Inspired by [23], several work [11], [14]–[16] have been done to extend the approach to a general multi-hop network and study the problem of energy balanced data propagation. However, they made an assumption that energy for each node in the network can propagate data to the sink either by direct transfer or through hop-by-hop transmission. As compared to this assumption, our mixed transmission model, which allows each sensor node to adjust its transfer range, is more realistic and practical. In [15], the authors proposed a hybrid transmission model based on estimated probabilities to achieve the energy balance property in WSNs. In [13], the authors proposed a mix transmission strategy for sensor nodes within the same slice model — in the first period the nodes send the data directly to the sink while in the second period the nodes forward the data to the next slice. The ratios between the two periods of time are computed in order to balance the energy consumption among all the nodes. In [14], the authors used a similar slice model to solve the energy balance problem in a stochastic manner. Their approach is to model the dynamics of energy consumption of sensors as a random walk. They extended the previous work [13], [15] by allowing adaptive energy assignment and proposed two efficient distributed algorithms to achieve energy balance. In [11], the authors proposed a ow model using Linear Programming to solve the mixed strategy data propagation, and showed that maximizing the lifespan, balancing the energy among individual sensor nodes, and maximizing the message ow in the network are equivalent. The mixed routing algorithm proposed allows each sensor node to either send a message to one of its immediate neighbours, or to send it directly to the sink. The routing decision is made based on a potential function depending on the remaining energy of the node. In our work, while we use the same slice model as in the existing work, we provide a more comprehensive solution to achieve both inter- and intra-slice energy balance. In addition, we analyze the necessary condition to achieve total energy balance for two representative WSN topologies.

VII. Conclusion

In this work, we study the energy balanced data collection problem in WSNs. By using the slice model, we address the problem by solving both the inter- and intra-slice energy balance problems. We propose an Energy-balanced Transmission Protocol by combining both the inter-slice mixed transmission strategy and the intra-slice forwarding technique to achieve

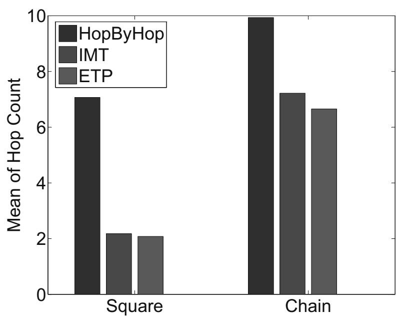

Fig. 18. Average delay in square and chain-shaped areas

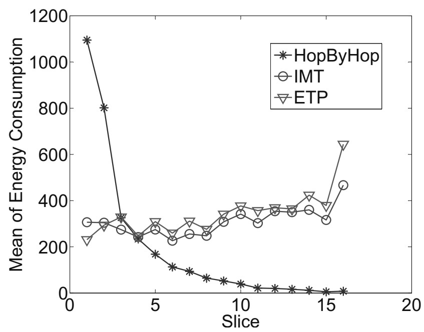

Fig. 19. Mean of energy consumption in each slice

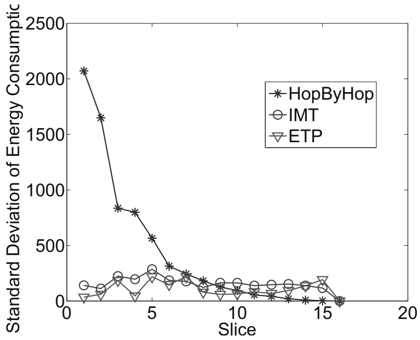

Fig. 20. Standard deviation of energy consumption in each slice

total energy balance in the WSN.

In the analysis of inter-slice energy balance, we find that to achieve total energy balance, the transmission range of a sensor node should be large enough with respect to the WSN area size, and they satisfy the necessary condition we derive. However, there always exists a trade-off between better energy balance performance and the cost of WSN deployment (i.e., using less expensive sensors with a shorter transmission range).

The existing design of ETP requires prior knowledge of events and energy distribution in the area. However, in the dynamic WSNs, such information may be hard to obtain because network condition keeps changing all the time. Hence, an advanced data collection protocol which works in a distributed manner is needed which we leave for our future work. In addition, link quality and interference can also affect the energy consumption during data transmission, especially for long-range data transfer. Although we can utilize multi-channel communications to minimize the negative impact, it is still interesting to investigate such challenging problems which we plan for our future work.

REFERENCES

[1] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, ser. WSNA 02. New York, NY, USA: ACM, 2002, pp. 88–97. [Online]. Available: http://doi.acm.org/10.1145/570738.570751

[2] K. Lu, Y. Qian, D. Rodr´guez, W. Rivera, and M. Rodriguez, "Wireless sensor networks for environmental monitoring applications: A design framework," in *GLOBECOM*, 2007, pp. 1108–1112.

[3] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in *Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation*. Ieee, 2005, pp. 719–724.

[4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[5] B. Lo, S. Thiemjarus, R. King, and G. Yang, "Body sensor network–a wireless sensor platform for pervasive healthcare monitoring," in *The 3rd International Conference on Pervasive Computing*, vol. 13, 2005.

[6] Y. Yu, R. Govindan, and D. Estrin, "Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks," *UCLA Computer Science Department Technical Report, UCLA-CSD TR-01-0023*, 2001.

[7] S. Jain, R. Shah, W. Brunette, G. Borriello, and S. Roy, "Exploiting mobility for energy efficient data collection in wireless sensor networks," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 327–339, 2006.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *HICSS*, 2000.

[9] T. van Dam and K. Langendoen, "An adaptive energy-efficient mac protocol for wireless sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, ser. SenSys 03. New York, NY, USA: ACM, 2003, pp. 171–180. [Online]. Available: http://doi.acm.org/10.1145/958491.958512

[10] P. Santi, "Topology control in wireless ad hoc and sensor networks," *ACM Computing Surveys (CSUR)*, vol. 37, no. 2, pp. 164–194, 2005.

[11] A. Jarry, P. Leone, O. Powell, and J. D. P. Rolim, "An optimal data propagation algorithm for maximizing the lifespan of sensor networks," in *DCOSS*, 2006, pp. 405–421.

[12] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad hoc networks*, vol. 3, no. 3, pp. 325–349, 2005.

[13] W. Guo, Z. Liu, and G. Wu, "An energy-balanced transmission scheme for sensor networks," in *SenSys*, 2003, pp. 300–301.

[14] P. Leone, S. E. Nikoletseas, and J. D. P. Rolim, "Stochastic models and adaptive algorithms for energy balance in sensor networks," *Theory Comput. Syst.*, vol. 47, no. 2, pp. 433–453, 2010.

[15] C. Efthymiou, S. E. Nikoletseas, and J. D. P. Rolim, "Energy balanced data propagation in wireless sensor networks," *Wireless Networks*, vol. 12, no. 6, pp. 691–707, 2006.

[16] S. E. Nikoletseas, "On the energy balance problem in distributed sensor networks," *Computer Science Review*, vol. 4, no. 2, pp. 65–79, 2010.

[17] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient mac protocol for wireless sensor networks," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 2002, pp. 1567–1576.

[18] G. Strang, *Calculus*. Wellesley-Cambridge Press, 1991.

[19] N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A wireless sensor network for structural monitoring," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*. ACM, 2004, pp. 13–24.

[20] J. D. Yu, K. T. Kim, B. Y. Jung, and H. Y. Youn, "An energy efficient chain-based clustering routing protocol for wireless sensor networks," in *AINA Workshops*, 2009, pp. 383–388.

[21] L. Kai, "A mine-environment-based energy-efficient routing algorithm for wireless sensor network," in *2009 Ninth International Conference on Hybrid Intelligent Systems*. IEEE, 2009, pp. 215–218.

[22] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," *Communications Letters, IEEE*, vol. 9, no. 11, pp. 976–978, 2005.

[23] M. Singh and V. K. Prasanna, "Energy-optimal and energy-balanced sorting in a single-hop wireless sensor network," in *PerCom*, 2003, pp. 50–59.