Energy Balanced Data Propagation in Wireless Sensor Networks *

Charilaos Efthymiou [†] Sotiris Nikoletseas [†] Jose Rolim[‡]

Abstract

We study the problem of *energy-balanced* data propagation in wireless sensor networks. The energy balance property guarantees that the average per sensor energy dissipation is the same for all sensors in the network, during the entire execution of the data propagation protocol. This property is important since it prolongs the network's lifetime by avoiding early energy depletion of sensors.

We propose a *new algorithm* that in each step decides whether to propagate data one-hop towards the final destination (the sink), or to send data directly to the sink. This randomized choice balances the (cheap) one-hop transimssions with the direct transimissions to the sink, which are more expensive but *"bypass" the sensors lying close to the sink*. Note that, in most protocols, these close to the sink sensors tend to be overused and die out early.

By a detailed analysis we *precisely estimate* the probabilities for each propagation choice in order to guarantee energy balance. The needed estimation can easily be performed by current sensors using simple to obtain information. Under some assumptions, we also derive a *closed form* for these probabilities.

The fact (shown by our analysis) that direct (expensive) transmissions to the sink are needed only rarely, shows that our protocol, besides energy-balanced, is *also energy efficient*.

1 Introduction

Wireless Sensor Networks are comprised of a very large number of micro-sensors, which form an ad-hoc network communicating with each other wirelessly. The micro-sensors are deployed in an area of interest, in order to e.g. monitor crucial

*This work has been partially supported by the IST/FET/GC Programme of the European Union under contract numbers IST-2001-33135 (CRESCCO) and 6FP 001907 (DELIS). A perliminary version of the work appeared in WMAN 2004 [11]. **Contact author**: Efthymiou Charilaos, Tel:+30 2610 960459, Fax:+30 2610 960 374

[†]Computer Technology Institute and Department of Computer Engineering & Informatics, University of Patras, Greece.E-mails: `euthimio@ceid.upatras.gr, nikole@cti.gr`.

[‡]University of Geneva, Switzerland, Email:`Jose.Rolim@cui.unige.ch`

events and propagate the collected results to a base station, which is usually called the *sink*. There are many important applications of such networks ranging from military applications to environmental studies, based on the existence of a wide variety of sensor types, such as for temperature, motion, noise, seismic activity etc (see [1, 2] for an excellent survey).

The dramatic progress in micro-electro-mechanical systems (MEMS) regarding the design and implementation of low-power electronics and low-power RF communication systems, has already led to small sized, low power and relatively inexpensive wireless sensors with important applications. However, wireless micro-sensors have to operate under severe limitations in their computational power, data storage, the quality of communication and, most crucially, their available amount of energy. Thus, the efficient distributed cooperation of sensor devices towards achieving a large computational and communication goal is a challenging task.

An important goal in the design and efficient implementation of wireless sensor networks is to save energy and keep the network functional for as long as possible. Towards this goal, various approaches including hop-by-hop transmission techniques ([13], [14], [7]), as well as clustering techniques ([12]) and alternating power-saving modes ([19]) have been proposed. Energy aware data gathering protocol have been introduced in [4].

All such techniques do not *explicitly* take care of possible overuse of certain sensors in the network. As an example, remark that in hop-by-hop transmissions towards the sink, the sensors lying closer to the sink tend to be utilized exhaustively (since all data passes through them). Thus, these sensors may die out very early, thus resulting to network collapse although there may be still significant amounts of energy in the other sensors of the network.

For a survey of data propagation protocols in wireless sensor networks, see [5]. For a focused discussion of energy efficiency issues, see [6]. For some recent research on energy efficiency and fault tolerance see [3], [16]. For a random graph theoretic model that captures proximity and seems approperiate for wireless networks see [17].

Our Contribution. In the light of the above, one should try to combine energy efficiency with a *balanced spread* of energy consumption among sensors.

In this paper, we propose and analyze *an energy efficient and energy balanced algorithm* for data propagation in sensor networks. The energy balance property of the algorithm ensures that the average energy dissipation per sensor is the same throughout the execution of the protocol in all sensors.

In particular, for the analysis we assume a random uniform placement of sensors in the network area and a random uniform generation of events. Once a data message reaches a sensor, it is either propagated one hop closer to the sink or it is sent directly to the sink. The choice of whether to forward one hop or send directly is done locally at each sensor, in a probabilistic manner.

By a detailed analysis, we accurately estimate these probabilities in a recurrent way, that can be supported by current sensors computing technology. We note that these probabilities are found to depend only on the network size and the distance of each particular sensor from the sink. Under certain assumptions, we also provide a simple closed form for the probability p_i to forward one-hop further. In particular, we show that $p_i = 1 - \frac{3x}{(i+1)(i-1)}$, where x is a free (e.g. controlling) parameter and i is related to the sensor's distance to the sink.

We note that, although we here present a specific model for energy dissipation in data transmission, our results can be easily extended to any energy cost model.

We also note that our protocol, besides being energy balanced is also energy efficient, in the sense that our analysis shows that, for guaranteeing energy balancedness, most transmissions must be one-hop transmissions, thus data is mostly transmitted along small distances and thus the overall energy consumption in the network is kept low.

Related Work. Our work has been inspired by the fundamental relevant work of [21] where the authors define the energy balancedness property and propose, analyze and evaluate an energy-optimal and energy balanced algorithm for sorting in wireless sensor networks. In particular, they consider a single-hop sensor network. Thus, our work extends their approach to the general case of a multi-hop network and for the (quite general) problem of propagating data to the sink.

Roadmap. In the next section (Section 2) we present the network model and define the problem of energy balanced data propagation. In Section 3 we present the distributed algorithm we propose. After some preliminary material in Section 4, we provide some simple examples of analysis in small networks, to get a feeling of the underlying ideas for the evaluation of the forwarding probability. In Section 6, we then carry out a detailed analysis of the general case leading to an accurate estimation of forwarding probabilities and discuss its behaviour for when we vary some parameters of the network. In Section 7 we provide a simple closed form for these probabilites under some assumptions that we discuss. Finally, in Section 8 we provide the analysis of the case where energy dissipation on the receipt of a data message is not negligible.

2 The Model and the Problem

We assume that crucial events, that should be reported to a control center, occur in the network area. Furthermore, we assume that these events are happening at random uniform positions. Let N be their total number in a certain period (i.e. during the execution of the protocol).

The sensors are spread in the network area randomly uniformly so their number in a certain area is proportional to the area's size. There is a single point in the network, which we call the "sink" S, that represents the fixed control center where data about event realization should be reported to. The sink is very powerful, in terms of both energy and computing power. Sensors can be aware of the direction (and position) of the sink, as well as of their distance to the sink. Such information can be easily obtained during a set-up phase, by having the (powerful) sink broadcast control messages to the entire network area. We assume the transmission range of sensors can vary with time (in fact

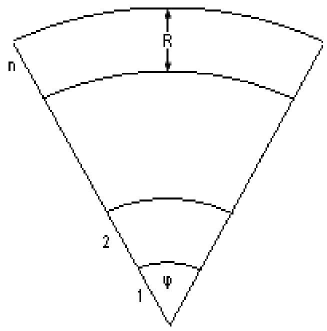

Figure 1: Sensor Network with n ring sectors, angle ϕ and ring "width" R

for each sensor our protocol may use only two different ranges: R and $i \cdot R$, where R is set by the network operator and i can be defined as $i = \lceil d/R \rceil$ with d to be the sensor's distance from the sink). We indeed assume a strong network model where all sensors are able to reach the sink in one hop. Similar modeling assumptions are made in the state-of-the-art (see e.g. the LEACH clustering protocol [12]). However, we note that this strong assumption does not trivialize the problem: former protocol with the same strong modeling assumptions are not energy balanced (see Section 2). The sensors do not move.

We virtually "cover" the network area by a disk sector of angle ϕ (see Fig. 1). The disk sector is divided into n ring sectors or "*slices*". The first slice has radius R (i.e. the sensors' transmission range). Slice i $(2 \leq i \leq n)$ is defined by two successive disk sectors, one of radius $i \cdot R$ and the other of radius $(i-1) \cdot R$. Taking a sufficiently large angle ϕ and/or by taking multiple sectors, we can cover the whole area.

As far as energy dissipation is concerned, we assume that the energy spent at a sensor when transmitting data messages is proportional to the square of the transmitting distance. Our protocol's performance analysis can be however extended to any energy cost model. The size of messages is considered to be constant. Note that the energy dissipation for receiving a data message is not always negligible with respect to the energy when sending such as in case when transmissions are too short and/or radio electronics energy is high (see [12]). In the analysis, initially, we only count (for simplicity) energy spent during transmissions. Since however in our protocol (see next section) there is one receipt for each transmission, it is clear that even when energy during receipt is more or less the same as energy during transmissions, the analysis can be extended easily to the full case (counting both transmissions and receipts). In fact, in Section 8 we explicitly provide the analysis of the full case.

Definition 1 The area between two consecutive cycle sectors is called *a ring sector* (or "slice"). Let T_i $(1 \leq i \leq n)$ be the i-th ring sector of the network.

T_1 stands for the ring sector with center the sink and radius equal to R.

Definition 2 Let S_i be the area size of the ring sector T_i of the network ($1 \leq i \leq n$).

We wish to solve the "*energy balanced data propagation problem*", i.e. to propagate data to the sink in such a way that the "average" energy dissipation in each sensor is at each time the same. The average energy dissipation per sensor is taken to be the fraction of the total energy spent by sensors in a ring sector over the number of sensors in that sector. Because of our assumption that the number of sensors in an area is proportional to the area size, the average energy dissipation per sensor is calculated by dividing the total energy spent in a sector by the sector size.

We do not study here medium access aspects, assuming the existence of underlying MAC protocol.

3 The Distributed Data Propagation Protocol

We assume that each event is sensed by only one sensor. This assumption is not restrictive since we may consider multiple sensing and propagation of an event by various sensors as sensing and propagation of many different events. A sensor sensing an event generates a data message which should be eventually delivered to the sink. On each ring sector, T_i, a number of events occur and a corresponding number of messages (one for each event) is generated.

Randomization is used to achieve some "load balancing" by evenly spreading the "load" (energy dissipation). In particular, on ring sector T_i each event is propagated to T_{i-1} (i.e. the "*next*" sector towards the sink) with probability p_i, while with probability $1-p_i$ it is propagated directly to the sink S. Each message in T_i is handled stochastically independently of the other events' messages.

The choice of probability p_i for T_i is made so as the average energy consumption per area unit (and thus per sensor) is the same for the whole network. There is a trade-off from choosing p_i: if p_i increases then transmissions tend to happen locally, thus energy consumption is low, however sensors closer to the sink tend to be overused since all data passes through them. On the other hand, if p_i decreases, there are distant transmissions (thus a lot of energy is consumed) however closer to sink particles are bypassed. Calculating the appropriate probability p_i for each T_i and solving the problem of energy balance is very important since it combines efficient data propagation with increased network's lifetime.

By using an underlying subprotocol we can guarantee that only one "next hop" sensor receives the transmitted message. Such a subprotocol can be any protocol that uses a (cheap) search phase to locate a "next" particle towards the sink and then directly transmits to that particle only. Such protocols are e.g. Directed Diffusion [13], the Local Target Protocol [7]. Note also that data messages are of fixed size i.e. no further information is added to a message on its route towards the sink.

Our protocol is (a) *distributed*, since each sensor chooses propagation probability independently of other sensors, (b) it uses only *local* information, in the sense that p_i depends only on i, i.e. a parameter related to the distance from the sink. Note that distance from the sink info for each sensor can be easily obtained i.e. during a set-up phase where the sink broadcasts control messages to the network. Several techniques (including signal attenuation evaluation) can be used to estimate each sensor's distance for the sink. (c) The protocol is very *simple*, since it just uses a random choice based only on parameter i.

4 Basic Definitions-Preliminaries

We aim at calculating probability p_i for each i in order to ensure the energy balance property.

Lemma 1 The area size, S_1, of the ring sector T_1 is

$$S_1 = \frac{\phi}{2} \cdot R^2$$

Proof: From geometry, it is

$$\begin{aligned} S_1 &= \tfrac{\phi}{2\pi} \cdot \pi \cdot R^2 \\ &= \tfrac{\phi}{2} \cdot R^2 \end{aligned}$$

◇

Lemma 2 The relation between the area size of the ring sector T_i and that of T_1 is

$$S_i = (2i - 1) \cdot S_1$$

Proof: The area size of S_i is equal to the area of the disk sector having radius $i \cdot R$ minus the area size of the disk sector of radius $(i-1) \cdot R$. Thus as in Lemma 1 it is:

$$\begin{aligned} S_i &= \tfrac{\phi}{2} \cdot (i \cdot R)^2 - \tfrac{\phi}{2} \cdot [(i-1) \cdot R]^2 \\ &= (2i - 1) \cdot \tfrac{\phi}{2} \cdot R^2 \end{aligned}$$

Using Lemma 1 we get

$$S_i = (2i - 1) \cdot S_1$$

◇

Definition 3 Let λ_i the probability that an event will occur on the ring sector T_i.

There are n ring sectors in the network. The value of λ_i is computed by the following lemma.

Lemma 3 Assuming a random uniform generation of events in the network area, the probability λ_i of an event occurring on the ring sector T_i $(1 \leq i \leq n)$, is:

$$\lambda_i = \frac{(2i-1)}{n^2}$$

Proof: Firstly, we will compute the probability λ_1. Because of the fact that events occur randomly uniformly in the network area we have that:

$$\begin{aligned} \lambda_1 &= \frac{S_1}{\sum_{i=1}^{n} S_i} \\ &= \frac{1}{\sum_{i=1}^{n} (2i-1)} \\ &= \frac{1}{n^2} \end{aligned}$$

Taking into consideration the random uniform distribution of the events in the network area and using Lemma 2, we have that $\lambda_i = (2i-1)\lambda_1$ and the result is proved. $\diamond$

Let us now consider sector T_i. We below provide the following basic definition:

Definition 4 An area T_i *"handles"* an event generated in ring sector j if either the message was generated in the area T_i (i.e. $j = i$) or the message was propagated to T_i from the ring sector T_{i+1}.

It is clear that, by the design of our algorithm, these options above are the only possible, since data is either propagated to the "*next*" sector or it is sent directly to the sink.

We now define energy ϵ_{ij} spent for message j when sector i handles it.

Definition 5 Let ϵ_{ij} a random variable which measures the energy that dissipates the sector T_i so as to handle the message j. For ϵ_{ij} we have that:

$$\epsilon_{ij} = \begin{cases} cR^2 & \text{with probability } p_i \\ c(iR)^2 & \text{with probability } 1 - p_i \end{cases}$$

where cR^2 is the energy dissipation for sending a message j from T_i to its adjacent ring sector T_{i-1} and c is a constant.

Thus, the expected energy dissipation in sector i for handling a message is

$$E[\epsilon_{i,j}] = cR^2 \cdot [i^2 - p_i(i^2 - 1)] \tag{1}$$

Note: The expected energy above is the same for all messages; we use j just for counting purposes.

Remark: We have used a special case for the energy cost model, for simplicity reasons. As will become evident in the analysis sections, our results can be easily extended to *any* cost model.

Definition 6 Let $\mathcal{E}_i$ the *total energy* spent by sensors in T_i. If h_i messages were handled in T_i then clearly:

$$\mathcal{E}_i = \sum_{j=1}^{h_i} \epsilon_{ij} \tag{2}$$

Energy balancedness is defined as follows:

Definition 7 The network is energy balanced if the average per sensor energy dissipation is the same for all sectors, i.e. when

$$\frac{E[\mathcal{E}_i]}{S_i} = \frac{E[\mathcal{E}_j]}{S_j} \quad i, j = 1, \ldots, n \tag{3}$$

5 Small Cases

As will become evident in the next sections, the calculation of probabilities p_i is very complicated. In order to first get a feeling and an intuition for the general solution, we first solve some small cases. In particular, before calculating the values of p_i for the general case of n sectors and N messages, we will give an example of this calculation for two special cases. In the first case we calculate the values of p_i's for 2 sectors and 2 messages, (i.e. $n = 2$ and $N = 2$), while in the next subsection we provide in the next subsection the calculation for $n = 3$ and $N = 2$.
From Lemma 3 we have that

$$\lambda_i = \frac{2i-1}{n^2}$$

In the specific case it is $n = 2$ thus $\lambda_1 = \frac{1}{4}$ and $\lambda_2 = \frac{3}{4}$. There are three different possibilities concerning the distribution of events into sectors and the probability of each possibility is:

$$\begin{array}{ll} Pr\{\text{2 events occur in sector 2}\} = (\lambda_2)^2 = \left(\frac{3}{4}\right)^2 = \frac{9}{16} & \text{(case a)} \\ Pr\{\text{1 event occurs in each sector}\} = 2\lambda_2 \cdot \lambda_1 = 2\frac{3}{4}\frac{1}{4} = \frac{6}{16} & \text{(case b)} \\ Pr\{\text{2 events occur in sector 1}\} = (\lambda_1)^2 = \left(\frac{1}{4}\right)^2 = \frac{1}{16} & \text{(case c)} \end{array}$$

For each one of the three cases above we calculate the average energy that each sector dissipates. Here, for simplification of the calculations we consider $R = 1$ and $c = 1$. Thus the energy dissipation for one hop transmission is 1 and for an i-hops transmission it is i^2. Thus we get:

$$\begin{array}{lll} \text{case a:} & \text{sector 2:} & 2 \cdot p_2^2 + 5 \cdot 2p_2(1-p_2) + 8(1-p_2)^2 \\ & \text{sector 1:} & 2p_2^2 + 1 \cdot 2p_2(1-p_2) \\ \text{case b:} & \text{sector 2:} & 1 \cdot p_2 + 4(1-p_2) \\ & \text{sector 1:} & 2 \cdot p_2 + 1(1-p_2) \\ \text{case c:} & \text{sector 2:} & 0 \\ & \text{sector 1:} & 2 \end{array}$$

To calculate the average energy dissipation per sensor in each ring sector we divide by sector area sizes:
Sector 2

$$\frac{E[\mathcal{E}_2]}{S_2} = \frac{\frac{9}{16}[2 \cdot p_2^2 + 5 \cdot 2p_2(1-p_2) + 8(1-p_2)^2] + \frac{6}{16}[p_2 + 4(1-p_2)]}{3\frac{\phi}{2}1^2}$$

Sector 1

$$\frac{E[\mathcal{E}_1]}{S_1} = \frac{\frac{9}{16}[2p_2^2 + 2p_2(1-p_2)] + \frac{6}{16}[2 \cdot p_2 + 1(1-p_2)] + 2\frac{1}{16}}{3\frac{\phi}{2}1^2}$$

By taking $\frac{E[\mathcal{E}_1]}{S_1} = \frac{E[\mathcal{E}_2]}{S_2}$ we can solve for p_2. The solution to this equation is

$$p_2 = \frac{1}{2}$$

Note. In this case (and also in the next) we remark that p_1 can be arbitrary. This is also verified by the general analysis (in the resulting recurrence relation the sequence term corresponding to p_1 is a free parameter).

The 2×3 case. An interesting special case is that in which there are 3 sectors and 2 messages. (i.e. $n = 3$ and $N = 2$). We take $R = 1$ (for calculation simplicity). By using lemma 3 we calculate the values of λ_i: $\lambda_1 = \frac{1}{9}$, $\lambda_2 = \frac{1}{3}$ and $\lambda_3 = \frac{5}{9}$. There are six different cases concerning the distribution of events in sectors and the probability of each one to occur is:

$Pr\{2 \text{ events in sector } 3\} = (\lambda_3)^2 = (\frac{5}{9})^2 = \frac{25}{81}$ (case a)
$Pr\{2 \text{ events in sector } 2\} = (\lambda_2)^2 = (\frac{1}{3})^2 = \frac{1}{9}$ (case b)
$Pr\{2 \text{ events in sector } 1\} = (\lambda_1)^2 = (\frac{1}{9})^2 = \frac{1}{81}$ (case c)
$Pr\{1 \text{ event in each of the sectors 3 and } 2\} = 2\lambda_3\lambda_2 = 2\frac{5}{9}\frac{1}{3} = \frac{10}{27}$ (case d)
$Pr\{1 \text{ event in each of the sectors 3 and } 1\} = 2\lambda_3\lambda_1 = 2\frac{5}{9}\frac{1}{9} = \frac{10}{81}$ (case e)
$Pr\{1 \text{ event in each of the sectors 2 and } 1\} = 2\lambda_2\lambda_1 = 2\frac{1}{9}\frac{1}{3} = \frac{2}{27}$ (case f)

The total energy dissipation in each sector for each case is:

case a	sector 3:	$2p_3^2 + 10 \cdot 2p_3(1-p_3) + 18(1-p_3)^2$
	sector 2:	$2p_3^2p_2^2 + 5p_3^2 2p_2(1-p_2) + 8p_3^2(1-p_2)^2 + 1 \cdot 2p_3(1-p_3)p_2 + 4 \cdot 2p_3(1-p_3)(1-p_2)$
	sector 1:	$2p_3^2p_2^2 + 1p_3^2 2p_2(1-p_2) + 1 \cdot 2p_3(1-p_3)p_2$
case b	sector 3:	0
	sector 2:	$2p_2^2 + 5 \cdot 2p_2(1-p_2) + 8(1-p_2)^2$
	sector 1:	$2p_2^2 + 1 \cdot 2p_2(1-p_2)$
case c	sector 3:	0
	sector 2:	0
	sector 1:	2
case d	sector 3:	$1p_3 + 9(1-p_3)$
	sector 2:	$2p_3p_2^2 + 5p_3 2p_2(1-p_2) + 8p_3(1-p_2)^2 + 1(1-p_3)p_2 + 4(1-p_3)(1-p_2)$
	sector 1:	$2p_3p_2^2 + 1p_3 2p_2(1-p_2) + 1(1-p_3)p_2$
case e	sector 3:	$1p_3 + 9(1-p_3)$
	sector 2:	$1p_3p_2 + 4p_3(1-p_2)$
	sector 1:	$2p_3p_2 + 1p_3(1-p_2) + 1(1-p_3)$
case f	sector 3:	0
	sector 2:	$1p_2 + 4(1-p_2)$
	sector 1:	$2p_2 + 1(1-p_2)$

To calculate the average energy dissipation per sensor in each sector, we divide by the sector sizes:

Sector 3: $$\frac{E[\mathcal{E}_3]}{S_3} = \frac{\frac{25}{81}(-16p_3+18)+\frac{10}{27}(9-8p_3)+\frac{10}{81}(9-8p_3)}{5\frac{\phi}{2}1^2}$$

Sector 2: $$\frac{E[\mathcal{E}_2]}{S_2} = \frac{\frac{25}{81}(-6p_3p_2+8p_3)+\frac{1}{9}(-6p_2+8-p_2^2)+\frac{10}{27}(-3p_3p_2+4p_3+3p_2+4)+\frac{10}{81}(-3p_3p_2+4p_3)+\frac{2}{27}(-3p_2+4)}{3\frac{\phi}{2}1^2}$$

Sector 1: $$\frac{E[\mathcal{E}_1]}{S_1} = \frac{\frac{25}{81}(2p_3p_2)+\frac{1}{9}(2p_2)+\frac{2}{81}+\frac{10}{27}(p_3p_2+p_2)+\frac{10}{81}(p_3p_2+1)+\frac{2}{27}(p_2+1)}{\frac{\phi}{2}1^2}$$

Taking $\frac{E[\mathcal{E}_1]}{S_1} = \frac{E[\mathcal{E}_2]}{S_2} = \frac{E[\mathcal{E}_3]}{S_3}$ we get a system of equations with solution:

$$p_3 = 0.5735 \qquad p_2 = 0.5815$$

Note that we have not included p_1 in the calculation above. This is because for sensors in T_1, both random choices have essentailly the same result. Thus, the value of p_1 does not affect the energy consumption in the network.

6 The General Solution

Now, we proceed to finding the solution for the general case of n sector and N events. We first define some useful random variables.

Definition 8 Let g_i be the number of the messages that are *generated* in the area T_i.

Note that messages are generated in an area only when events occur in this area.

Definition 9 Let f_i be the number of the messages that are *forwarded to* the area T_i.

Definition 10 Let h_i be the number of the messages that are "handled" by the area T_i.

We note that messages are forwarded to a ring sector (say i) only because of an event's generation at a sector $j > i$ and successive one-hop propagations from sector j to sector i.
We notice the following important relation:

$$h_i = g_i + f_i \tag{4}$$

which means that the number of messages that area T_i handles equals the number of the messages that are generated in T_i, plus the number of messages that are forwarded to it. Because of event generation according to a probability distribution and also because of the probabilistic nature of message propagation in, all three quantities above are random variables. By linearity of expectation, we get:

Lemma 4

$$E[h_i] = E[g_i] + E[f_i]$$

Furthermore, in the following lemma, we establish a relationship between $E[f_i]$, $E[g_{i+1}]$ and $E[f_{i+1}]$.

Lemma 5

$$E[f_i] = p_{i+1} \cdot (E[f_{i+1}] + E[g_{i+1}])$$

Proof: Let $\delta_{i,j}$ an indicator random variable that is equal to 1 if area T_i forwards the message j to the area T_{i-1} and 0. Thus:

$$\delta_{i,j} = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases}$$

Clearly, $\delta_{i,j}$ depends only on i, but we add j for counting purposes. Obviously, $E[\delta_{i,j}] = p_i$. It is:

$$f_i = \sum_{j=0}^{h_{i+1}} \delta_{i+1,j}$$

$$\begin{aligned} E[f_i] &= E\left[\textstyle\sum_{j=0}^{h_{i+1}} \delta_{i+1,j}\right] \\ &= \textstyle\sum_{n=0}^{N} E\left[\sum_{j=0}^{h_{i+1}} (\delta_{i+1,j} \cap h_{i+1} = n)\right] \\ &= \textstyle\sum_{n=0}^{N} E\left[\sum_{j=0}^{h_{i+1}} (\delta_{i+1,j} | h_{i+1} = n)\right] \cdot Pr\{h_{i+1} = n\} \end{aligned}$$

Furthermore

$$\begin{aligned} E\left[\textstyle\sum_{j=0}^{h_{i+1}} (\delta_{i+1,j} | h_{i+1} = n)\right] &= E\left[\textstyle\sum_{j=0}^{n} \delta_{i+1,j}\right] \\ &= \textstyle\sum_{j=0}^{n} E\left[\delta_{i+1,j}\right] \\ &= n \cdot E\left[\delta_{i+1,j}\right] \end{aligned}$$

Thus, we have from the above:

$$\begin{aligned} E[f_i] &= \textstyle\sum_{n=0}^{N} E\left[\sum_{j=0}^{h_{i+1}} (\delta_{i+1,j} | h_{i+1} = n)\right] \cdot Pr\{h_{i+1} = n\} \\ &= \textstyle\sum_{n=0}^{N} n \cdot E\left[\delta_{i+1,j}\right] \cdot Pr\{h_{i+1} = n\} \\ &= E\left[\delta_{i+1,j}\right] \cdot \textstyle\sum_{n=0}^{N} n \cdot Pr\{h_{i+1} = n\} \\ &= E\left[\delta_{i+1,j}\right] \cdot E[h_{i+1}] \\ &= p_i \cdot (E[f_{i+1}] + E[g_{i+1}]) \end{aligned}$$

The final equation follows by the fact that $E[\delta_{i+1,j}] = p_i$ and by Lemma 4. The proof is completed. $\diamond$

We can use Lemma 5 to calculate the value of p_i by computing the values of $E[g_i]$ and $E[f_i]$ for $i = 1, \ldots, n$, for the $E[g_i]$ we have,

Lemma 6 If N is the total number of events that are generated in the network, the mean value of g_i is given by the following relationship:

$$E[g_i] = N \cdot \lambda_i$$

Proof: Because the position of each event is independent of other events and because for each sector i, probability λ_i is the same, clearly g_i is binomial with parameters N, λ_i. $\diamond$

It now remains to compute the values of $E[f_i]$. This computation follows immediatelly. We provide a lemma useful in the estimation of the total energy dissipation in a sector.

Lemma 7 The expected total energy dissipation in sector i is:

$$E[\mathcal{E}_i] = E[h_i] \cdot E[\epsilon_{ik}]$$

Proof:

$$\begin{aligned} E[\mathcal{E}_i] &= E\left[\sum_{k=1}^{h_i} \epsilon_{ik}\right] \\ &= \sum_{n=0}^{N} E\left[\sum_{k=0}^{h_i} (\epsilon_{ik} \cap h_i = n)\right] \\ &= \sum_{n=0}^{N} E\left[\sum_{k=1}^{h_i} \epsilon_{ik} \middle| h_i = n\right] \cdot \Pr\{h_i = n\} \end{aligned}$$

Similarly to the proof of Lemma 5 we get

$$E[\mathcal{E}_i] = E[h_i] \cdot E[\epsilon_{ik}]$$

and the proof is completed. $\diamond$

Recall that, according to Definition 7, to achieve the same on the average energy dissipation *per area unit* (and thus per sensor) in the network area, the following equality should hold:

$$\frac{E\left[\sum_{k=1}^{h_i} \epsilon_{ik}\right]}{S_i} = \frac{E\left[\sum_{k=1}^{h_j} \epsilon_{jk}\right]}{S_j} \quad \forall i, j \in \{1, \ldots n\} \tag{5}$$

i.e. the average energy consumption per sensor should be equal in any two ring sectors. By induction, it suffices to guarantee this for any two adjacent sectors. In what follows, we guarantee the above balance property, requiring a certain recurrence relation to hold. This recurrence basically relates 3 successive terms of the $E[f_i]$ sequence (the $E[g_i]$ terms depend only on i and on input parameters).

Theorem 1 To achieve energy balance in the network, the following recurrency equation should hold:

$$a(i+1)E[f_{i+1}] - (d(i) + a(i))E[f_i] + d(i-1)E[f_{i-1}] = a(i)E[g_i] - a(i+1)E[g_{i+1}]$$

where

$$a(i) = \frac{i^2}{2i-1} \qquad d(i) = \frac{(i+1)^2 - 1}{2i+1}$$

Proof: For the case $j = i + 1$ of the equation 9 and using the lemmata 2 and 7 we have:

$$\begin{aligned}
\frac{E[h_i]E[\epsilon_{i,j}]}{S_i} &= \frac{E[h_{i+1}]E[\epsilon_{i+1,j}]}{S_{i+1}} \quad \Leftrightarrow \\
\frac{E[h_i]cR^2[i^2 - p_i(i^2-1)]}{(2i-1)S_1} &= \frac{E[h_{i+1}]cR^2\{(i+1)^2 - p_{i+1}[(i+1)^2-1]\}}{(2i+1)S_1} \quad \Leftrightarrow \\
\frac{E[h_i][i^2 - p_i(i^2-1)]}{(2i-1)} &= \frac{E[h_{i+1}]\{(i+1)^2 - p_{i+1}[(i+1)^2-1]\}}{(2i+1)} \quad \Leftrightarrow \\
\frac{i^2}{2i-1}E[h_i] - p_iE[h_i]\frac{i^2-1}{2i-1} &= \frac{(i+1)^2}{2i+1}E[h_{i+1}] - p_{i+1}E[h_{i+1}]\frac{(i+1)^2-1}{2i+1}
\end{aligned}$$

Let $a(i)$, $d(i)$ be

$$a(i) = \frac{i^2}{2i-1} \qquad d(i) = \frac{(i+1)^2-1}{2i+1}$$

By Lemma 4 we have that $E[h_i] = E[g_i] + E[f_i]$ and by Lemma 5 we have that $E[f_i] = p_{i+1}E[h_{i+1}]$ thus the last equation becomes:

$$a(i+1)E[f_{i+1}] - (d(i) + a(i))E[f_i] + d(i-1)E[f_{i-1}] = a(i)E[g_i] - a(i+1)E[g_{i+1}]$$

$\diamond$

Remark: In the above theorem we got a recurrence relation for a special case of energy dissipation in data transmission. This analysis can easily be extended to any cost model since only $a(i)$, $d(i)$ functions of i will be affected.

Structure of the rest of the recurrence solution evaluation.
The rest of the recurrence solution includes the following basic building blocks:

- First we transform the recurrency of Theorem 1 into a simpler recurrency (that of Lemma 8), since it includes only *two* successive terms (whose coefficient is 1) of the sequence.
- Having solved the latter recurrency, we solve (see Theorem 2) the initial one, i.e. that of Theorem 1.
- Having computed the values of $E[f_i]$ for $i = 1, \ldots, n$ we use Lemma 5 to compute the exact values of probabilites p_i.

In order to have a recurrence involving only two (succesive in fact) terms of the $E[f_i]$ sequence, we will transform the recurrency relation of Theorem 1. It is:

Definition 11 Let $t(i) = a(i+1)E[f_{i+1}] - d(i)E[f_i]$, for $i \in \{0, 1, \ldots, n-1\}$.

The recurrency relation of Theorem 1 can be expressed as:

$$t(i) - t(i-1) = a(i) \cdot E[f_i] - a(i+1) \cdot E[f_{i+1}] \quad i = 1, \dots n-1$$

$$\text{with} \quad t(0) = a(1) \cdot E[f_1]$$

Note that although we do not know yet the value of $E[f_1]$, we can compute the values of $t(i)$ as a function of $E[f_1]$. Furthermore, as we prove in the next lemma, this function is linear in $E[f_1]$.

Lemma 8 The recurrency relation:

$$t(i) - t(i-1) = a(i) \cdot E[g_i] - a(i+1) \cdot E[g_{i+1}] \quad i = 0, \dots n-1$$

$$\text{with} \quad t(0) = a(1) \cdot E[f_1]$$

has as a solution the function

$$t(i) = \sum_{j=1}^{i} (a(j)E[g_j] - a(j+1)E[g_{j+1}]) + a(1) \cdot E[f_1]$$

Proof: The proof is done by induction on i. For $i = 0$, it is obviously true. Let the relation be true for $i-1$, then for i we have:

$$\begin{aligned} t(i) - t(i-1) &= a(i) \cdot E[g_i] - a(i+1) \cdot E[g_{i+1}] \quad \Leftrightarrow \\ t(i) &= t(i-1) + a(i) \cdot E[g_i] - a(i+1) \cdot E[g_{i+1}] \end{aligned}$$

By the induction hypothesis we get the solution

$$t(i) = \sum_{j=1}^{i} (a(j)E[g_j] - a(j+1)E[g_{j+1}]) + a(1) \cdot E[f_1]$$

$\diamond$

Now the recurrency relation of Theorem 1 is simplified as follows:

$$a(i+1) \cdot E[f_{i+1}] - d(i) \cdot E[f_i] = t(i) \quad i = 1, \dots, n-1$$

Thus, we get a recurrence for sequence $E[f_i]$ involving only two successive terms of the sequence:

Theorem 2 The recurrency relation

$$a(i+1)E[f_{i+1}] - d(i)E[f_i] = t(i) \quad i = 1, \dots n-1$$

where $t(i)$ is defined in lemma 8, has the following solution

$$E[f_{n-i}] = -\sum_{k=1}^{i} \frac{\prod_{j=k}^{i-1} a(n-j)}{\prod_{j=k}^{i} d(n-j)} \cdot t(n-k)$$

Proof: To get an intuitive feeling of the solution, we first present a small case, i.e. $n = 4$. The recurrency relation can be written as a linear system of 3 variables, $E[f_1]$, $E[f_2]$, $E[f_3]$ and 3 equations:

$$\begin{array}{c} a(2)E[f_2] - d(1)E[f_1] = t(1) \\ a(3)E[f_3] - d(2)E[f_2] = t(2) \\ a(4)E[f_4] - d(3)E[f_3] = t(3) \end{array}$$

We note that $E[f_4] = 0$ because no messages are forwarded to the outmost sector. Thus we have:

$$\begin{array}{c} E[f_3] = -\frac{t(3)}{d(3)} \\ E[f_2] = -\frac{a(3)t(3)}{d(3)d(2)} - \frac{t(2)}{d(2)} \\ E[f_1] = -\frac{a(3)a(2)t(3)}{d(3)d(2)d(1)} - \frac{a(2)t(2)}{d(2)d(1)} - \frac{t(1)}{d(1)} \end{array}$$

We can express $t(i)$ as a linear function of $E[f_1]$, thus the last equation depends only on $E[f_1]$ while the others on $E[f_i]$ and $E[f_1]$. The last of the three equations can give the $E[f_1]$ and then we can calculate the rest $E[f_i]$.

We now proceed to the general n. The proof is done by induction on i. For $i = 1$, we compute the value of $E[f_{n-1}]$. From the recurrence relation we have for $E[f_{n-1}]$

$$a(n)E[f_n] - d(n-1)E[f_{n-1}] = t(n-1)$$

But, $E[f_n]$ is the number of events that are propagated to ring sector n, which is zero. Also notice that $d(i) = 0 \Leftrightarrow i = 0$. Thus:

$$E[f_{n-1}] = -\frac{t(n-1)}{d(n-1)}$$

which is equal to the solution. So we are done with the base of the induction. Assuming that the induction hypothesis is correct for $i - 1$, we will prove that it is also correct for i. From the recurrency relation we have that:

$$E[f_{n-i}] = \frac{a(n-i+1)}{d(n-i)} \cdot E[f_{n-i+1}] - \frac{t(n-i)}{d(n-i)}$$

We can safely divide with $d(i)$ because the recurrence is defined for $i = 1, \ldots, n-1$ and no value of these make $d(i)$ zero. But, from the induction hypothesis we know that

$$E[f_{n-i+1}] = -\sum_{k=1}^{i-1} \frac{\prod_{j=k}^{i-2} a(n-j)}{\prod_{j=k}^{i-1} d(n-j)} \cdot t(n-k)$$

Replacing $E[f_{n-i+1}]$ in the first relation we have

$$E[f_{n-i}] = \frac{a(n-i+1)}{d(n-i)} \cdot \left(-\sum_{k=1}^{i-1} \frac{\prod_{j=k}^{i-2} a(n-j)}{\prod_{j=k}^{i-1} d(n-j)} \cdot t(n-k) \right) - \frac{t(n-i)}{d(n-i)}$$

We finally get

$$E[f_{n-i}] = -\sum_{k=1}^{i} \frac{\prod_{j=k}^{i-1} a(n-j)}{\prod_{j=k}^{i} d(n-j)} \cdot t(n-k) \quad where \quad \prod_{i}^{i-1} a(i) = 1$$

$\diamond$

The full expression for $E[f_i]$ can be expressed by substituting i with $n-i$, thus

$$E[f_i] = -\sum_{k=1}^{n-i} \frac{\prod_{j=k}^{n-i+1} a(n-j)}{\prod_{j=k}^{n-i} d(n-j)} \cdot \left(\sum_{j=1}^{n-k} (a(j)E[g_j] - a(j+1)E[g_{j+1}]) + a(1) \cdot E[f_1] \right)$$

where $\prod_{i}^{i-1} a(i) = 1$.

We note that all the parameters of the recurrency solution above are expressed as a function of $E[f_1]$ and i. So as to compute them, we firstly compute the value of $E[f_1]$. Then we can compute all the other parameters by replacing the already computed $E[f_1]$.
Now, the calculation of the probabilities p_i is quite straightforward.

Theorem 3 The energy balance property is achieved if any ring sector (say T_i) propagates each message it handles with probability p_i to the next ring sector, T_{i-1}, and with probability $1-p_i$ it propagates the message directly to the sink. The value of each p_i is given by the following relation

$$p_i = \frac{E[f_{i-1}]}{E[g_i] + E[f_i]}$$

where the values of $E[f_i]$ and $E[g_i]$ are obtained from lemma 2 and lemma 6, respectively.

Proof: From Lemma 5 we know that $E[f_{i-1}] = p_i(E[f_i] + E[g_i])$ thus:

$$p_i = \frac{E[f_{i-1}]}{E[f_i] + E[g_i]}$$

Using Theorem 2 and Lemma 6 we compute the exact values of p_i $\diamond$
Remark. Note that, interestingly, p_i's are *independent* of the number N of the events that occur in the network, since p_i's depend only on i and the number of ring sectors n (which is broadcast to sectors by the sink). Thus the protocol assumes only local information.

We note that the analysis above allows the exact derivation of probabilities p_i's as a function of i and n which (although complicated and not obviously leading to a closed form) can be easily calcualted by the sensors in the network by carrying out very simple calculations.

Theorem 4 Given that the energy is each time on the average the same in all network sensors, each message will finally get to the sink.

Proof: Consider a message that is generated in a certain ring sector, say T_i. Because of the energy balance property, all sensors are operational, thus in each step of the protocol the message is either propagated to the "next" (i.e. immediately closer to the sink) ring sector, T_{i-1}, or is directly propagated to the sink. Consequently, after a maximum of i steps, the message will finally reach the sink. $\diamond$

We note that we guarantee that the energy per sensor is the same on the average; we plan to study energy balance properties guarantees with high probability as well, by taking into account the distribution (both spatial and in time) of event generation.

6.1 Some remarks on the result obtained in Theorem 3

In this subsection we discuss some graphical representation of the probabilities p_i as a function of the ring sector i for various network sizes and we study its behaviour. We also discuss a graphical representation of p_i as a function of the exponent that signal attenuates with distance (capturing the energy cost model) for a specific network size.

Figure 2 and Figure 3 represent p_i as a function of the sector number i for two different network sizes: 100 and 1000 ring sectors.

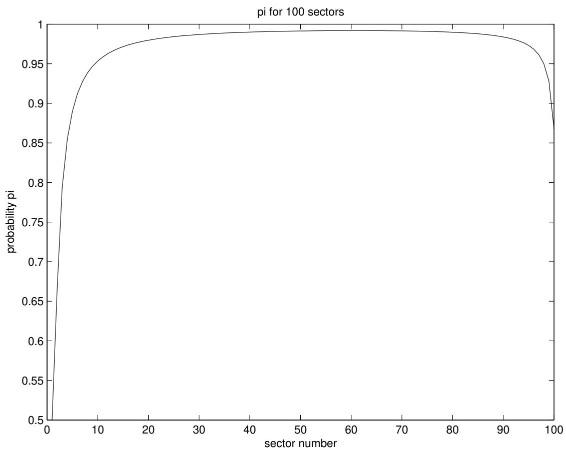

Figure 2: p_i when n=100

We remark that for all cases the probability initially increases, in fact very rapidly; then there is a territory of the area for which p_i is more or less constant and finally in the last 30-40 sectors the probability decreases slightly. This behaviour can be explained based on the fact that the energy spent in a transmission is proportional to the square of the transmission distance. Thus, only in close to the sink sectors direct transmissions to the sink tend to happen: this

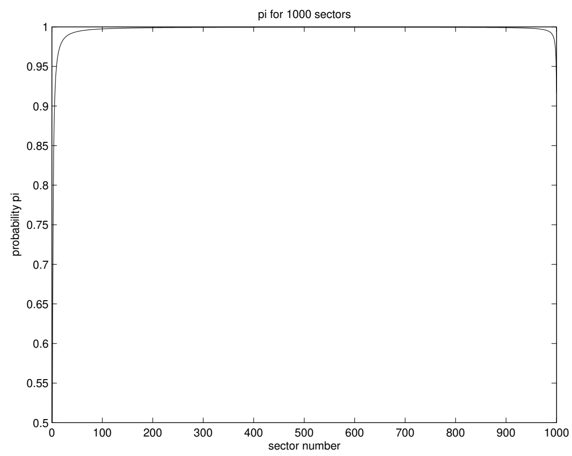

Figure 3: p_i when n=1000

is why p_i is pretty small for small i. In not very close to the sink sectors (i.e. when i becomes relatively large) then p_i approaches 1 i.e. mostly one-hop transmissions tend to occur. This also suggests that our approach is energy efficient in the sense that (expensive) direct transmissions to the sink are not favoured by our algorithm. This very important finding shows that the protocol is at the same time both energy balanced and energy efficient.

The decrease in the value of p_i when i approaches n is a consequence of the fact that there is an asymmetry between the furthermost ring sector and the other ring sectors. This asymmetry is due to the fact that, in contrast to other sectors, no sector propagates messages to the n-th ring sector. Thus the sectors close to n have a lighter load since they handle fewer events, thus, to achieve energy balance, they should spend more energy per message i.e. they should more frequently transmit directly to the sink. (Note that for large i the area size of adjacent sectors is approximately the same thus no significant increase in message generation happens when i is close to n).

In the following three figures we compare the behaviour of p_i in these three "territories" of p_i's behaviour change (initial increase, stability, slight decrease) for network sizes of 100, 1000 and 2000 sectors. Figure 4 represents p_i in the first territory where it increases.

In this figure we note that for the close to the sink sectors (i small) and independently of n, p_i increases in almost exactly the same manner. This is because of the following reason: when n increases then it is less probable that events make it to sector in the territory i.e. close to the sink; thus, these few events should spend more energy by jumping more frequently to the sink

In Figure 5, we compare p_i's for the middle territory of the network. We note that although in the case of network size 1000 and 2000 p_i's are almost the same, in the case of the 100 ring sector network the probability is relatively

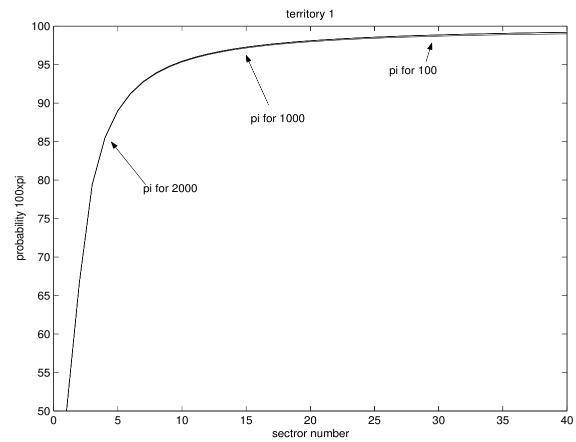

Figure 4: p_i on the first territory

small. This is because of the fact that the impact of the asymmetry effect of the last ring sector reaches even the middle territory of the network.

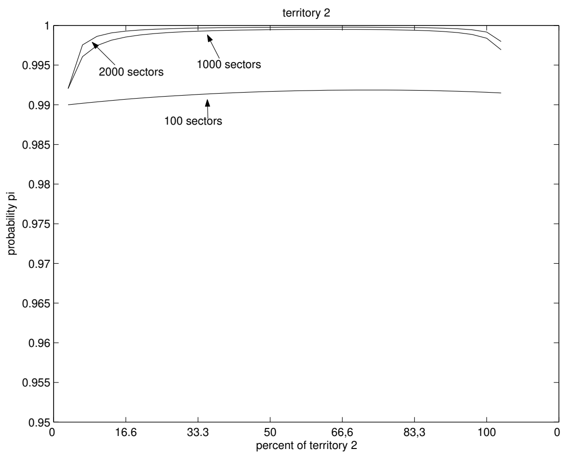

Figure 5: p_i on the second territory

In Figure 6 we give a detailed graphical representation of p_i in the last 50 sectors for all cases of network size. We notice that the probabilities p_i decrease with i, almost in the same manner for all network sizes.

In large networks (e.g. n=1000) the forwarding probability is indeed close to 1. However, the probability of directly transmitting to the sink at some stage is upper bounded by a geometric distribution with success probability close to 0.

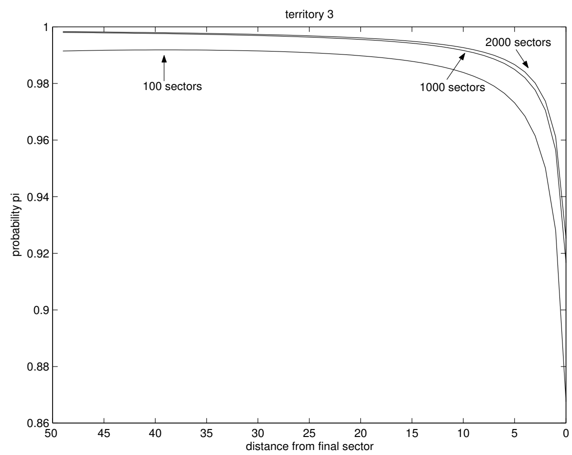

Figure 6: p_i on the third territory

Thus, direct transmissions to the sink occur after some hops and lead to energy balance.

Finally, in Figure 7 we study the effect of the energy cost model. In particular, we focus on the impact of the power of the signal power fading. The qualitative behaviour is basically the same i.e. p_i increases with i, close to the sink. However, when the fading factor increases (i.e. when transimssions are more expensive) direct transmissions to the sink are even less probable. This is why p_i increases when the fading factor becomes bigger.

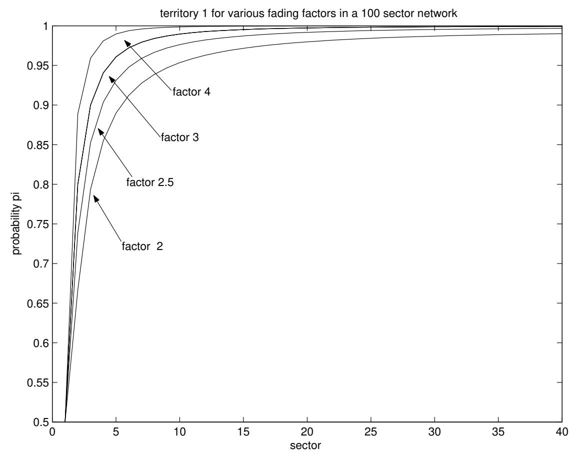

Figure 7: p_i increase for different fading factors

In the next section, we also provide a *closed form* for probabilities p_i under certain simplifying (however reasonable) assumptions.

7 A closed form under simplifying assumptions

Under specific assumptions (that we discuss and motivate) we can make the calculation of probabilities p_i simpler. Combining Lemma 4 and Lemma 6 and the assumption that we have N messages in the network we have from the energy balance property that

$$\frac{\frac{2i-1}{n^2}N + E[f_i]}{(2i-1)\frac{\phi}{2}R^2}[p_i+(1-p_i)i^2]R^2 = \frac{\frac{2i-3}{n^2}N + E[f_{i-1}]}{(2i-3)\frac{\phi}{2}R^2}[p_{i-1}+(1-p_{i-1})(i-1)^2]R^2$$

But $2i-1 \simeq 2i-3$ and $\frac{\phi}{2}$, R^2 cancel. Dividing by N we get:

$$\left[\frac{2i-1}{n^2} + \frac{E[f_i]}{N}\right]\left[p_i + (1-p_i)i^2\right] \simeq \left[\frac{2i-1}{n^2} + \frac{E[f_{i-1}]}{N}\right]\left[p_{i-1} + (1-p_{i-1})(i-1)^2\right]$$

If $E[f_i] \simeq E[f_{i-1}]$ then the previous relation becomes:

$$p_i + (1-p_i)i^2 = p_{i-1} + (1-p_{i-1})(i-1)^2 \tag{6}$$

We finally get:

Theorem 5 If $E[f_i] \simeq E[f_{i-1}]$, $3 \leq i \leq n$, then the one-hop forwarding probability, guaranteeing energy balance, is

$$p_i = 1 - \frac{3x}{(i+1)(i-1)}$$

where $p_2 = x \in (0,1)$ a free controlling parameter and $p_1 = 0$.

Proof: We can transform the recurrency in (6), i.e.

$$p_i + (1-p_i)i^2 = p_{i-1} + (1-p_{i-1})(i-1)^2$$

into

$$\frac{q_i}{q_{i-1}} = \frac{i(i-2)}{(i+1)(i-1)}$$

where $q_i = 1 - p_i$. ◇

Lemma 9 The solution of the recurrency relation

$$\frac{q_i}{q_{i-1}} = \frac{i(i-2)}{(i+1)(i-1)}$$

where $q_1 = 1$ and $q_2 = x$ (a free variable e.g. a controlling parameter) is $q_i = \frac{3x}{(i+1)(i-1)})$

Proof: By sequentially "unfolding" the recurrence we get:

$$\begin{aligned} q_i = q_{i-1}\tfrac{i(i-2)}{(i+1)(i-1)} &= q_{i-2}\tfrac{(i-1)(i-3)}{i(i-2)}\tfrac{i(i-2)}{(i+1)(i-1)} \\ &= q_{i-3}\tfrac{(i-2)(i-4)}{(i-1)(i-3)}\tfrac{(i-1)(i-3)}{(i+1)(i-1} \\ &\vdots \\ &= q_2\tfrac{3\cdot 1}{(i+1)(i-1)} \end{aligned}$$

Thus:

$$q_i = \frac{3x}{(i+1)(i-1)}$$

$\diamond$

We would like to remark that the assumption $E[f_i] \simeq E[f_{i-1}]$ is quite reasonable and well motivated. We provide the following intuitive explanation of why this happens. Remark indeed that the area sizes of adjacent sectors (and thus the number of events generated in such sections) are more or less the same, especially when i increases. Furthermore, the probability p_i of forwarding to the adjacent (towards the sink) sector increases very fast with i and becomes 1 in most sectors of the network (in the middle territory). Thus, the expected number of messages forwarded to adjacent sectors seems to be almost the same. In large networks (e.g. n=1000) the forwarding probability is indeed close to 1. However, the probability of directly transmitting to the sink at some stage is upper bounded by a geometric distribution with success probability close to 0. Thus, direct transmissions to the sink occur after some hops and lead to energy balance.

In Figure 8 we graphically display both functions of p_i (i.e. the general solution and that obtained by the simple closed form). We indeed remark that the curves are very similar in shape and very close to each other. In particular in the "middle" territory (i.e. for $50 \leq i \leq 950$ when n=1000) the two curves almost coincide. For $1 \leq i \leq 50$ (in the initial territory) the exact and approximation probability are in most sections less than 0.08 (Figure 9) apart and only in very few sectors the difference becomes 0.16. Finally, at the third territory ($950 \leq i \leq 1000$) the accurate solution and the approximation differ by less than 0.08 (Figure 10) This shows that our solution under the simplifying assumptions is very tight (Figure 11).

8 Analysis in cases where there is a receipt cost

The energy dissipation for receiving a message is not always negligible with respect to the energy when sending such as in case when transmissions are too short and/or radio electronics energy is high (see [12]). We define the total energy ϵ^s_{ij} spent for message j when sector i transmits it.

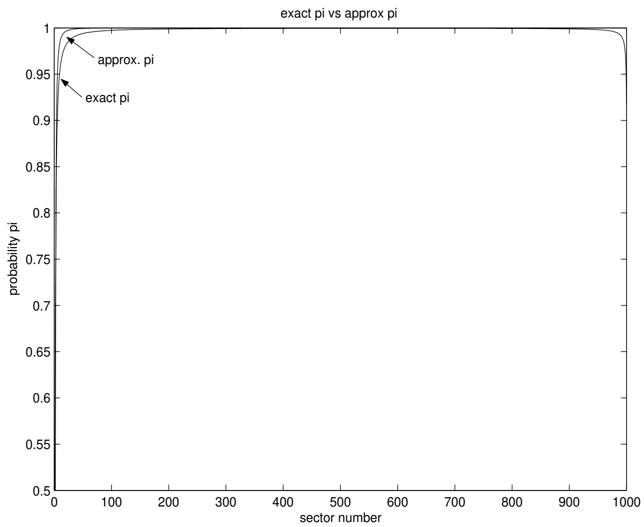

Figure 8: Representation of exact p_i and approximated p_i

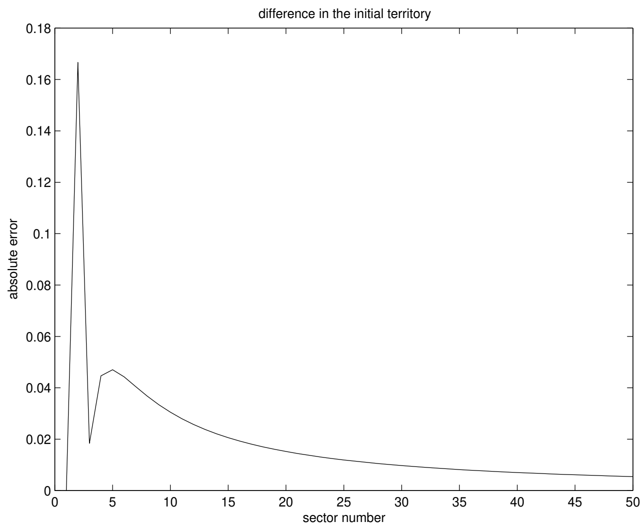

Figure 9: "Absolute Error" in the first territory

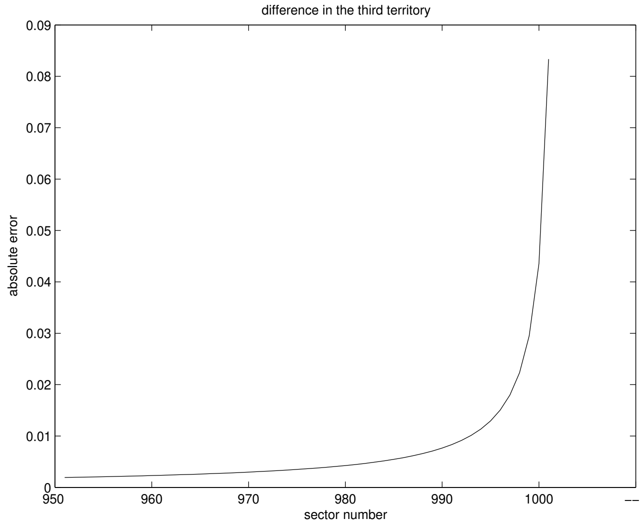

Figure 10: “Absolute Error” in the third territory

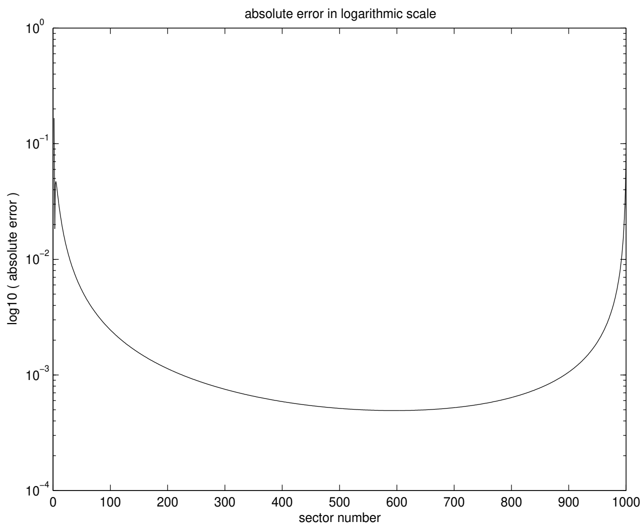

Figure 11: “Absolute Error” in logarithmic scale for the whole network

Definition 12 Let ϵ^s_{ij} a random variable which measures the energy that dissipates the sector T_i so as to *send* a message j. For ϵ^s_{ij} we have that:

$$\epsilon^s_{ij} = \begin{cases} c_1 R^2 + c_2 & \text{with probability } p_i \\ c_1 (iR)^2 + c_2 & \text{with probability } 1 - p_i \end{cases}$$

where $c_1 R^2 + c_2$ is the energy dissipation for sending a message j from T_i to its adjacent ring sector T_{i-1} and c_1, c_2 constants.

The expected energy dissipation in sector i for sending a message is

$$E[\epsilon^s_{i,j}] = c_1 R^2 \cdot [i^2 - p_i(i^2 - 1)] + c_2 \tag{7}$$

Also we have,

Definition 13 Let ϵ^r_{ij} a variable which measures the energy that dissipates sector T_i so as to *receive* the message j. For ϵ^r_{ij} we have that:

$$\epsilon^r_{i,j} = c_2$$

The total energy spent by sensor T_i is now extended as follows:

$$\mathcal{E}_i = \sum_{j=1}^{h_i} \epsilon^s_{ij} + \sum_{j=1}^{f_i} \epsilon^r_{ij} \tag{8}$$

Now, we proceed to finding the solution for the general case of n sector and N events. We next provide a lemma useful in the estimation of the total energy dissipation in a sector.

Lemma 10 The expected total energy dissipation in sector i so as to send the messages is:

$$E[\mathcal{E}^s_i] = E[h_i] \cdot E[\epsilon^s_{ik}]$$

Proof: Similarly as in Lemma 7. ◇

We also have that:

Lemma 11 The expected total energy dissipation in sector i so as to receive the messages is:

$$E[\mathcal{E}^r_i] = E[f_i] \cdot E[\epsilon^r_{ik}]$$

Proof: The variable ϵ^r_{ik} is constant for each area T_i. Thus:

$$E[\mathcal{E}^r_i] = E\left[\sum_{j=1}^{f_i} e^r_{ij}\right] = e^r_{ij} \cdot E[f_i]$$

◇

Thus the expected mean value of the energy dissipation in each area T_i is given by the following relation:

$$E[\mathcal{E}_i] = E[h_i] \cdot E[e^s_{ij}] + e^r_{ij} \cdot E[f_i]$$

Recall that, according to Definition 7, to achieve the same on the average energy dissipation *per area unit* (and thus per sensor) in the network area, the following equality should hold:

$$\frac{E\left[\mathcal{E}_i\right]}{S_i} = \frac{E\left[\mathcal{E}_j\right]}{S_j} \quad \forall i,j \in \{1,\dots n\} \tag{9}$$

i.e. the average energy consumption per sensor should be equal in any two ring sectors. By induction, it suffices to guarantee this for any two adjacent ring sectors.

In what follows, we manage to guarantee the above balance property, by requiring a certain recurrence relation to hold. We note that this recurrence basically relates 3 successive terms of the $E[f_i]$ sequence (the $E[g_i]$ terms depend only on i and on input parameters).

Theorem 6 To achieve energy balance in the network, the following recurrency equation should hold:

$$b(i+1)E[f_{i+1}] - (d(i)+b(i))E[f_i] + d(i-1)E[f_{i-1}] = a(i)E[g_i] - a(i+1)E[g_{i+1}]$$

where

$$a(i) = \frac{(iR)^2c_1 + c_2}{2i-1} \qquad d(i) = \frac{[(i+1)^2-1]c_1R^2}{2i+1} \qquad b(i) = a(i) + \tfrac{c_2}{2i-1}$$

Proof: For the case $j = i+1$ of the equation 9 and using the lemmata 2 and 7 we have:

$$\frac{E[h_i]E[\epsilon^s_{i,j}] + e^r_{ij}E[f_i]}{S_i} = \frac{E[h_{i+1}]E[\epsilon_{i+1,j}] + e^r_{i+1j}E[f_{i+1}]}{S_{i+1}} \quad \Leftrightarrow$$

$$\frac{(iR)^2c_1 + c_2}{2i-1}E[h_i] - p_iE[h_i]\frac{(i^2-1)c_1R^2}{2i-1} + \frac{c_2}{2i-1}E[f_i] =$$

$$= \frac{[(i+1)R]^2c_1 + c_2}{2i+1}E[h_{i+1}] - p_{i+1}E[h_{i+1}]\frac{[(i+1)^2-1]c_1R^2}{2i+1} + \frac{c_2}{2(i+1)-1}E[f_{i+1}]$$

Let

$$a(i) = \frac{(iR)^2c_1 + c_2}{2i-1} \qquad d(i) = \frac{[(i+1)^2-1]c_1R^2}{2i+1} \qquad b(i) = a(i) + \tfrac{c_2}{2i-1}$$

By Lemma 5 we know that $p_iE[h_i] = E[f_{i-1}]$ and by Lemma 4 it is $E[h_i] = E[g_i] + E[f_i]$ thus the last equation becomes:

$$b(i+1)E[f_{i+1}] - (d(i)+b(i))E[f_i] + d(i-1)E[f_{i-1}] = a(i)E[g_i] - a(i+1)E[g_{i+1}]$$

$\diamond$

In order to have a recurrence involving only two (succesive in fact) terms of the $E[f_i]$ sequence, we will transform the recurrency relation of Theorem 6 similarly to Theorem 1.

Definition 14 Let $t(i) = b(i+1)E[f_{i+1}] - d(i)Ef_i]$, for $i \in \{0, 1, \ldots, n-1\}$.

The recurrency relation of Theorem 6 can be expressed as:

$$t(i) - t(i-1) = a(i) \cdot E[g_i] - a(i+1) \cdot E[g_{i+1}] \quad i = 0, \ldots n-1$$

$$\text{with} \quad t(0) = b(1) \cdot E[f_1]$$

From Lemma 8 we know that

$$t(i) = \sum_{j=1}^{i} (a(j)E[g_j] - a(j+1)E[g_{j+1}]) + b(1) \cdot E[f_1]$$

Now the recurrency relation of Theorem 6 is simplified as follows:

$$b(i+1) \cdot E[f_{i+1}] - d(i) \cdot E[f_i] = t(i) \quad i = 1, \ldots, n-1$$

Thus, we get a recurrence for sequence $E[f_i]$ involving only two successive terms of the sequence:

Theorem 7 The recurrency relation

$$b(i+1)E[f_{i+1}] - d(i)E[f_i] = t(i) \quad i = 1, \ldots n-1$$

where $t(i)$ is defined in Definition 14, has the following solution

$$E[f_{n-i}] = -\sum_{k=1}^{i} \frac{\prod_{j=k}^{i-1} b(n-j)}{\prod_{j=k}^{i} d(n-j)} \cdot t(n-k)$$

Proof: Similarly as Theorem 2 having only $b(i)$ instead of $a(i)$. $\diamond$

Thus, the full expression for $E[f_i]$ can be expressed by substituting i with $n-i$, thus

$$E[f_i] = -\sum_{k=1}^{n-i} \frac{\prod_{j=k}^{n-i+1} b(n-j)}{\prod_{j=k}^{n-i} d(n-j)} \cdot \left(\sum_{j=1}^{n-k} (a(j)E[g_j] - a(j+1)E[g_{j+1}]) + b(1) \cdot E[f_1] \right)$$

where

$$\prod_{i}^{i-1} b(i) = 1$$

We, again, note that all the parameters of the recurrency solution above are expressed as a function of $E[f_1]$ and i. So as to compute them, we firstly compute the value of $E[f_1]$. Then we can compute all the other parameters by replacing the already computed $E[f_1]$.

Now, the calculation of the probabilities p_i is done as in Theorem 3.

9 Closing Remarks and Future Research

Our 2-level approach may indeed not be optimal with respect to some performance measures (e.g. energy dissipation). We, here, do not claim optimality; we claim energy balance and energy efficiency. Our work is a first step towards the optimality goal. We indeed intend to explore this issue in future research.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci: Wireless sensor networks: a survey. In the Journal of Computer Networks, Volume 38, pp. 393-422, 2002.

[2] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci: A Survey on Sensor Networks. In the IEEE Communications Magazine, pp. 102-114, August 2002.

[3] A. Boukerche, I. Chatzigiannakis and S. Nikoletseas: A New Energy Efficient and Fault-tolerant Protocol for Data Propagation in Smart Dust Networks using Varying Transmission Range, accepted in the Computer Communications Journal, Elsevier, 2004.

[4] A. Boukerche, X. Cheng and J. Linus: Energy-aware data-centric routing in microsensor networks, ACM Madeling Analysis and Simulation of Wireless and Mobile Systems (MSWIM2003) (Paris, France), 2003, pp. 42-49.

[5] A. Boukerche and S. Nikoletseas: Protocols for Data Propagation in Wireless Sensor Networks: A Survey, Chapter in the Book Wireless Communications Systems and Networks, Editor Mohsen Guizani, Kluwer Academic Publishers, Date Published: 06/2004, ISBN: 0306481901, 718 p.

[6] A. Boukerche and S. Nikoletseas: Energy Efficient Algorithms in Wireless Sensor Networks invited Book Chapter, Springer Verlag, to appear in 2004.

[7] I. Chatzigiannakis, S. Nikoletseas and P. Spirakis: Smart Dust Protocols for Local Detection and Propagation. In *Proc. 2nd ACM Workshop on Principles of Mobile Computing* – POMC'2002. Also accepted in the ACM Mobile Networks (MONET) Journal, Special Issue on Algorithmic Solutions for Wireless, Mobile, Ad-Hoc and Sensor Networks, to appear.

[8] I. Chatzigiannakis, T. Dimitriou, S. Nikoletseas and P. Spirakis: A Probabilistic Algorithm for Efficient and Robust Data Propagation in Smart Dust Networks. In Proc. 5th European Wireless Conference (EW'04), Barcelona, Spain, February 24-27, 2004. Also, invited in the Journal of Ad-Hoc Networks (under review)

[9] I. Chatzigiannakis, T. Dimitriou, M. Mavronicolas, S. Nikoletseas and P. Spirakis: A Comparative Study of Protocols for Efficient Data Propagation in Smart Dust Networks. In *Proc. International Conference on Parallel and Distributed Computing* – EUPOPAR 2003. Also Accepted in the Parallel Processing Letters Journal (PPL), Volume 13, pp. 615-627, 2003.

[10] I. Chatzigiannakis and S. Nikoletseas: A Sleep-Awake Protocol for Information Propagation in Smart Dust Networks. In *Proc. 3rd Workshop on Mobile and Ad-Hoc Networks (WMAN)*–IPDPS Workshops, IEEE Press, p. 225, 2003.

[11] C. Efthymiou, S. Nikoletseas and J. Rolim: Energy Balanced Data Propagation in Wireless Sensor Networks. In *Proc. 4th Workshop on Mobile, Ad-Hoc and Sensor Networks (WMAN)*–IPDPS Workshops, IEEE Press, 2004.

[12] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan: Energy-Efficient Communication Protocol for Wireless Microsensor Networks. In *Proc. 33rd Hawaii International Conference on System Sciences* – HICSS'2000.

[13] C. Intanagonwiwat, R. Govindan and D. Estrin: Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks. In *Proc. 6th ACM/IEEE International Conference on Mobile Computing* – MOBICOM'2000.

[14] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann and F. Silva: Directed Diffusion for Wireless Sensor Networking. Extended version of [13].

[15] J.M. Kahn, R.H. Katz and K.S.J. Pister: Next Century Challenges: Mobile Networking for Smart Dust. In *Proc. 5th ACM/IEEE International Conference on Mobile Computing*, pp. 271-278, September 1999.

[16] S. Nikoletseas, I. Chatzigiannakis, A. Antoniou, C. Efthymiou, A. Kinalis and G. Mylonas, Energy Efficient Protocols for Sensing Multiple Events in Smart Dust Networks. In Proc. 37th Annual ACM/IEEE Simulation Symposium (ANSS'04), IEEE Computer Society Press, pp. 15 24, 2004.

[17] S. Nikoletseas, C. Raptopoulos and P. Spirakis, The Existence and Efficient Construction of Large Independent Sets in General Random Intersection Graphs, in the Proceedings of the 31st International Colloquium on Automata, Languages and Programming (ICALP), Lecture Notes in Computer Science (Springer Verlag), 2004.

[18] S. M. Ross: Stochastic Processes, 2nd Edition. *John Wiley and Sons, Inc.*, 1995.

[19] C. Schurgers, V. Tsiatsis, S. Ganeriwal and M. Srivastava: Topology Management for Sensor Networks: Exploiting Latency and Density. In *Proc. MOBICOM 2002*.

[20] P. Triantafilloy, N. Ntarmos, S. Nikoletseas and P. Spirakis: NanoPeer Networks and P2P Worlds. In *Proc. 3rd IEEE International Conference on Peer-to-Peer Computing*, 2003.

[21] M. Singh, and V. Prasanna: Energy-Optimal and Energy-Balanced Sorting in a Single-Hop Wireless Sensor Network. In *Proc. First IEEE International Conference on Pervasive Computing and Comminications* – PERCOM, 2003.