An Adaptive Blind Algorithm for Energy Balanced Data Propagation in Wireless Sensors Networks*

Pierre Leone[1], Sotiris Nikoletseas[2], José Rolim[1]

[1] Computer Science Department, University of Geneva, 1211 Geneva 4, Switzerland
[2] Computer Technology Institute (CTI) and Patras University, P.O. Box 1122, 261 10, Patras, Greece

Abstract. In this paper, we consider the problem of energy balanced data propagation in wireless sensor networks and we generalise previous works by allowing realistic energy assignment. A new modelisation of the process of energy consumption as a random walk along with a new analysis are proposed. Two new algorithms are presented and analysed. The first one is easy to implement and fast to execute. However, it needs a priori assumptions on the process generating data to be propagated. The second algorithm overcomes this need by inferring information from the observation of the process. Furthermore, this algorithm is based on stochastic estimation methods and is adaptive to environmental changes. This represents an important contribution for propagating energy balanced data in wireless sensor netwoks due to their highly dynamic nature.

1 Introduction

Load balancing is a common important problem in many areas of distributed systems. A typical example is that of shared resources such as a set of processors, where it is of interest to assign tasks to resources without overusing any of them. A related but different aspect of load balancing appears in the context of sensor networks, where tiny smart sensors are usually battery powered: an important goal of data processing is to balance the total energy consumed among the entire set of sensors. However, limited local knowledge of the network, frequent changes in the topology of the network and the specifications of sensors, among others, make load balancing in sensors nets significantly different of classical load balancing in distributed systems.

To our knowledge, these considerations were first pointed out in the field of sensor networks in [9]. In this paper the authors deal with the problem of devising energy balanced sorting algorithms. In a subsequent paper [6] the authors deal with the problem of energy balanced data propagation in sensor networks. They propose a randomised data propagation protocol and provide recursive and closed form solutions for the appropriate data propagation probabilities.

* This work has been partially supported by the IST Programme of the European Union under contract numbers IST-2001-33135 (CRESCCO) and 001907 (DELIS).

Before describing our contributions we present the problem previously stated in [6]. Formal definitions are deferred to the next section. An important area of application of sensor networks is the monitoring of a given region. Tiny smart sensors are scattered in a given region in order to detect and monitor some phenomena. Once a sensor detects the occurrence of an event it is responsible to inform (through wireless transmissions) a particular station (representing the end users of the network), called sink, about the occurrence of this event. Since the energy necessary to transmit a data through radio waves is proportional to some power of the distance of transmission (usually square power), sensors located far away from the sink are prone (if they would transmit directly to the sink) to run out of their available energy before sensors located closer to the sink. This leads to the idea that the data traffic has to be handled by the network with multiple hops to the sink, allowing only short distance communication. However, this strategy tends to overuse sensors located close to the sink since these sensors have to handle the entire set of events. There is then a trade-off between long and short distance communication to forward a data to the sink in order to make the life time of the whole network longer. A possible probabilistic protocol divides the set of sensors into slices or ring sectors [6]. The first slice is composed of sensors at *unit* distance from the sink, the second slice of sensors at distance 2, and so on. Here the distance is the maximal number of hops necessary to send a data to the sink. Sensors may communicate directly to the sink with probability $(1 - p_i)$. In this case the consumed energy is proportional to i^2 with i is the slice number the sensor belongs to. In order to save energy, a sensor can probabilistically decide with probability p_i to transmit its data to a sensor belonging to the next slice to the sink (slice $i \rightarrow$ slice $i - 1$). In this case the amount of consumed energy is proportional to a constant which is assumed to be 1 for convenience. This is illustrated in Figure 1. The problem is to determine with which probability p_i a sensor located in slice i has to transmit to the next slice (or to directly transmit to the sink with probability $1 - p_i$) in order to balance the consumed energy among all the sensors.

In the literature, probabilistic data propagation protocols are proposed, in [5] (LTP, a local optimisation protocol) and [4] (PFR, a limited flooding protocol). Both protocols are energy efficient and fault-tolerant, but tend to strain close to the sink sensors. [1] proposes a variable transmission range protocol (VTRP) that tends to bypass strained sensors lying close to the sink.

In [2] the authors propose a novel and efficient energy-aware distributed heuristic (EAD), to build a special rooted tree with many leaves that is used to facilitate data-centric routing. In [3] a new protocol (PEQ) is proposed that uses a *driven delivery of events* mechanism that selects the fastest path and reduces latency. [7] inspired our work here since it proposes some stochastic models (Markov chains, dynamic systems) for dynamic sensor networks.

In this section, we generalise the energy balanced data propagation problem by allowing unrestricted realistic energy assignment and we propose two new probabilistic protocols, one of which is adaptive. Our analysis is based on the modelisation of the process of energy consumption as a random walk in $\mathcal{R}^n$.

The first algorithm we propose is relatively similar to the one suggested in [6] and corresponds to offline computation of the probabilities p_i of transmission to the next slice. Although very easy to implement and fast in execution it suffers from an important **weakness**; namely the probability of occurrence of the events per slice, i.e. the probability λ_i have to be known. This particularity allows very efficient computations of the probabilities p_i. However, this property is not realistic or at least we gain in flexibility and adaptability to devise an algorithm able to solve the problem without any assumption concerning these probabilities. The analysis of the problem is new and leads to a formal definition of the problem of energy balanced data propagation.

The second algorithm is adaptive and based on stochastic approximation methods. The algorithm does not assume that the probabilities of occurrence of the events are known and infers their values from the observation of the events. We refer to such an algorithm as **blind** algorithm for energy balanced data propagation to stress the fact that there is no *a priori* knowledge on the statistics concerning the localisation of the events. The algorithm can be accordingly implemented on any given network and run on the fly, allowing online adaptation of the parameters of the network. This characteristic is important if the parameters of the network are prone to change (this appears frequently in sensor networks). This algorithm is an important contribution of this section. Generally, adaptive algorithms, like the one proposed here, are most appropriate for wireless sensor networks because of their evolving nature due to dynamic properties of the networks such as sensors failures, obstacles, etc., leading to topology changes. We also formally define in a broader sense the problem of energy balanced data propagation and show formally under which conditions the problem is well formulated.

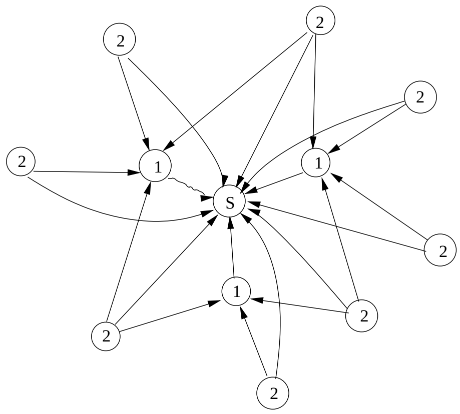

Fig. 1. The sink (S) with the first two slices of sensors

2 Framework and Formal Definition of the Problem

In this section we state formally the framework and notations and state the problem of energy balanced data propagation in wireless sensor networks. Notice that as a result of the analysis of the problem, that is presented in the next section, we show that the problem as stated in this section is well formulated.

The number of slices is denoted by n. The main assumption we need in this section is that the energy consumed per sensor to handle the data to the sink is the same among sensors belonging to a particular slice. This means that sensors belonging to the same slice exhaust their available energy more or less simultaneously. Both following assumptions give sufficient conditions validating this assumption. Notice that these assumptions are based on a probabilistic selection of a sensor belonging to a slice for data transmission. Different protocols can then be proposed.

We assume that the probability that an event is detected by a given sensor depends uniquely on the slice the sensor belongs to. This means that we can define and estimate $\lambda_1, \lambda_2, \ldots, \lambda_n$ ($\sum_i \lambda_i = 1$) where λ_i is the probability that an event occurs in slice number i. For example, this property is satisfied if the events are uniformly randomly distributed on the monitored region. Indeed, in this particular situation, the probabilities λ_i are proportional to the area covered by the i-th slice. Moreover, when a data is transmitted from slice i to slice $i-1$ the selected sensor belonging to the slice $i-1$ is uniformly selected among the whole set.

The probability p_i, $i = 1, \ldots, n$ denotes the probability that a sensor belonging to the slice i sends a data to a sensor belonging to the "next" slice $i-1$. The complementary probability $1-p_i$ denotes the probability that the sensor sends the data directly to the sink. Then, when a data is handled by a sensor belonging to the i-th slice the amount of consumed energy is a constant (assumed to be 1 for convenience) with probability p_i and i^2 with probability $1-p_i$. By definition we have $p_1 = 1$ because sensors belonging to the first slice can do nothing else than transmitting to the sink.

The number of sensors belonging to the i-th slice is denoted by S_i. It might be the case that there is a strong relationship between S_i, λ_i but this is not essential.

The total energy available at the i-th slice is denoted by E_i, thus $e_i = E_i/S_i$ is the available energy per sensor. The energy can be seen as a given amount of energy available at the start or as a rate of consumable energy.

An important aspect of our analysis is to model the energy consumption for handling a given event as a random walk in R^n. We group the available scaled energy of each slice as a vector

$$\left(E_n/S_n, E_{n-1}/S_{n-1}, \ldots, E_1/S_1\right)^T. \tag{1}$$

To start, consider an event generated in the $n-$th slice and consider the complete process of handling the data generated by the event to the sink. We have different possibilities for the scaled energy consumed in the different slices corresponding to the different paths of the data. The data is directly transmitted to the sink with probability $(1-p_n)$. The corresponding consumed energy vector per sensor is $\left(n^2/S_n, 0, \ldots, 0\right)^T$. The data can alternatively be transmitted to the next $(n-1)-$th slice from which it is directly transmitted to the sink with probability $p_n(1-p_{n-1})$. The corresponding consumed energy vector in this situation is $\left(1/S_n, (n-1)^2/S_{n-1}, 0, \ldots, 0\right)^T$. Repeating this enumerative process, we describe all possible events with their probability and the corresponding vector of energy consumed.

Formally, we denote the $U = \{U_1, U_2, \ldots\}$ the set of vectors describing the energy consumption for handling an event. Denoting by Ω the set of possible events we obtain a random variable $\Omega \rightarrow U$ which describes the energy consumed for handling an event. If we associate to each event its probability we have our probability space $(\Omega, \mathcal{P}(\Omega), P)$.

The process of energy consumption is described as a random walk in R^n with the energy consumed for handling m events in the form $X_1+X_2+X_3+\ldots+X_m$, where X_i are independent random realisations of the random variable $\Omega \rightarrow U$. The law of large numbers implies that $X_1+X_2+\ldots+X_m \rightarrow mE(X)$ thus, to ensure energy balanced data propagation we must have

$$E(X) = \lambda\big(E_n/S_n, E_{n-1}/S_{n-1}, \ldots, E_1/S_1)^T. \tag{2}$$

Intuitively, if the expected consumed energy does not satisfy (2) then there is a slice in which sensors will run out the available energy, described by (1), before the sensors belonging to others slices. The network stops working prematurely. Moreover, if (1) describes the rate of consumable energy requirement (2) amounts to preserving the ratio of consumed energy per slice. An energy assignment vector is a vector of the form (1) meaning that the ratio of energy consumed in slice i with respect to slice j should be E_i/E_j.

We later prove that that the set of admissible energy assignment vectors is $\{v \in R^n : v_i \geq 0, \| v \| = \text{constant}\}$ and to each such vector there is a unique assignment of the probabilities p_i. Besides this existential result, we propose two new protocols for calculating the optimal probabilities in an efficient manner. The first protocol assumes a certain amount of local knowledge, while the second one implicitly estimates the statistics of the events and is able to appropriately adapt to changes in the network parameters.

3 An aware strategy for balanced energy dissipation

To ensure energy balance we have to determine for each slice i the probability of transmitting a given data to the next slice p_i, the data being transmitted directly to the sink with probability $(1-p_i)$. This section deals with this problem assuming an a-priori knowledge of the probabilities λ_i of the distribution of occurrences of the events among different slices. The first slice, located just before the sink, has only to transmit the data to the sink directly $(p_1 = 1)$.

Hence, if n is the total number of slices we have $n-1$ unknown probabilities $p_2, \ldots, p_n$. The other free parameter to be determined is the factor λ appearing in (2).

Consider a node in the i-th slice which has to transmit a data. The data has to be transmitted because of an event occurring in the i-th slice with probability λ_i. The data can also be transmitted because it was previously generated by the preceding $(i+1)-$th slice. This occurs with probability $\lambda_{i+1} \cdot p_{i+1}$. The event can also be transmitted due to an event generated in the $(i+2)-$th slice, this occurs with probability $\lambda_{i+2} \cdot p_{i+2} \cdot p_{i+1}$ and so on up to the $n-$th slice. Then, a data is transmitted from the i-th slice with probability

$$\lambda_i + \lambda_{i+1} p_{i+1} + \lambda_{i+2} p_{i+2} p_{i+1} + \ldots + \lambda_n p_n p_{n-1} \cdot \ldots \cdot p_{i+1}. \tag{3}$$

The mean dissipated energy per sensor on the i-th slice is of the form

$$p_i \frac{1}{S_i} + (1-p_i) \frac{i^2}{S_i}. \tag{4}$$

Then the mean energy dissipated in the i-th slice is of the form

$$\big(\lambda_i + \lambda_{i+1} p_{i+1} + \ldots + \lambda_n p_n p_{n-1} \cdots p_{i+1}\big)\big(p_i \frac{1}{S_i} + (1-p_i)\frac{i^2}{S_i}\big) = \lambda e_i, \tag{5}$$

where the equality is imposed to ensure energy balanced data propagation through the network. With $p_{n+1} = \lambda_{n+1} = 0$ we define the x_i value as

$$x_i = \lambda_i + \lambda_{i+1} p_{i+1} + \ldots + \lambda_n p_n p_{n-1} \cdots p_{i+1}, \tag{6}$$

which satisfies the recurrence relation

$$x_i = p_{i+1} x_{i+1} + \lambda_i, \quad i = n, \ldots, 1, \tag{7}$$

with the convention $p_{n+1} = 0$. Solving (5) for p_i, $i = n, \ldots, 2$ we get

$$p_i = \frac{i^2 x_i - S_i e_i \lambda}{(i^2-1) x_i} = \frac{i^2}{i^2-1} - \frac{S_i e_i}{(i^2-1) x_i} \lambda, \quad i = n, \ldots, 2. \tag{8}$$

Since $p_1 = 1$ we solve (5) with $i = 1$ for λ and get

$$\lambda = \frac{x_1}{S_1 e_1}. \tag{9}$$

This last equation is actually a constraint on λ. To investigate properties of this constraint, we define a function $(\lambda, E_1, \ldots, E_n) \rightarrow \lambda' = f(\lambda)$ by substituting λ by λ' in (9). A fixed point of this function, i.e. $\lambda = f(\lambda)$, determines through (6) and (8) a solution of our problem. The recursive scheme described in Figure 2 calculates the value of the fixed point of this function. This algorithm is executed by the sink based on the knowledge of the energy assignment vector, the probabilities of occurrence of the events on the different slices and the number of sensors per slices. This information may come directly from the sensors (both at set-up or during protocol evolution) which know the slice they belong to.

Actually, there is a more efficient way of solving this problem, based on the result of Proposition 1, as illustrated in Figure 3.

Proposition 1. *The function $f(\lambda)$ defined through equations (6), (8) and (9) is linear. Then, we can write*

$$f(\lambda) = a + b\lambda, \tag{10}$$

with a and b real constants defined by

$$S_1 e_1 a = \lambda_1 + \lambda_2 C_2 + \lambda_3 C_3 C_2 + \ldots + \lambda_n C_n C_{n-1} \ldots C_2, \tag{11}$$

$$S_1 e_1 b = -D_2 - D_3 C_2 - D_4 C_3 C_2 - \ldots - D_n C_{n-1} \ldots C_2, \tag{12}$$

with

$$C_i = \frac{i^2}{i^2 - 1}, \quad D_i = \frac{S_i e_i}{i^2 - 1}. \tag{13}$$

Proof. Let us fix a value of λ. Computing $f(\lambda)$ amounts to recursively computing the parameters x_i for $i = n, n-1, \ldots, 1$ and then computing $f(\lambda)$ with (9). We prove by induction that the x_i values are all linear in λ. For $i = n$, we get $x_n = \lambda_n$. So, the assertion is true for $i = n$ Let us now assume it is true for $i = n, n-1, \ldots, k+1$. Using (7) and (8) we get $x_k = p_{k+1} x_{k+1} + \lambda_k = C_{k+1} x_{k+1} - D_{k+1}\lambda + \lambda_k$, which is linear in λ establishing the first part of the assertion. Since the x_i are linear functions in λ we introduce the notation

$$x_i = a_i + b_i \lambda, \quad i = n, \ldots, 1. \tag{14}$$

The coefficients a_i and b_i are recursively determined by

$$a_i = \lambda_i + a_{i+1} C_{i+1}, \tag{15}$$

$$b_i = b_{i+1} C_{i+1} - D_{i+1}, \tag{16}$$

with inital conditions $a_n = \lambda_n$ and $b_n = 0$ (i.e. $x_n = \lambda_n$). These equations prove (11) (12). To get (15) and (16), we write again (8) as $p_i = C_i - \frac{D_i}{x_i}\lambda$, which can be inserted into formula (7) to obtain $x_i = \lambda_i + x_{i+1} C_{i+1} - D_{i+1}\lambda$. Inserting (14) into this last equation leads to the recursive expression for a_i and b_i.

In the preceding computations, the dependence of the function $f(\lambda)$ on the total energy per slice (the E_i parameters) was not stated clearly, neither was the physical interpretation of this vector. The next results show that the entries of this vector are not important by themselves. What is important is the ratio E_i/E_j between the total mean energy available at the i-th slice with respect to the $j-$th slice.

Corollary 1. *Consider two total energy assignment vectors which are linearly dependent,*

$$\begin{pmatrix} E_n \\ E_{n-1} \\ \vdots \\ E_1 \end{pmatrix} = \mu \begin{pmatrix} E'_n \\ E'_{n-1} \\ \vdots \\ E'_1 \end{pmatrix},$$

```
Initialise p_2, ..., p_n and λ
Initialise NbrLoop=1
while not convergence
   x ← 0
   for counter = n to 2
      x ← x + λ_counter
      p_counter ← p_counter(λ, p_counter+1, ..., p_n) with (8)
      x ← x p_counter
   end for
   x ← x + λ_1
   Compute λ_inter with (9)
   λ ← λ + (λ_inter − λ)/nbrLoop
   nbrLoop ← NbrLoop + 1
end while
```

Fig. 2. Pseudo-code for iterative solution of (2)

1. The Sink compute the fixed point of $f(\lambda)$ defined in Proposition 1
2. The Sink sends to every sensor the relevant λ value
3. Each sensor computes its probability p_i

Fig. 3. High-level description of the energy balanced data propagation protocol

with μ a real non-zero constant. Then, the fixed points, λ and λ' of the functions

$$f(\lambda, E_1, \ldots, E_n) = \lambda, \ \textit{and}\ f(\lambda', E'_1, \ldots, e'_n) = \lambda'$$

are related by

$$\lambda' = \mu\lambda. \tag{17}$$

Moreover, the corresponding probabilities p_i and p'_i of transmitting directly to the next slice are equal.

Proof. This follows from the result of theorem 1. The value of the fixed point are given respectively by

$$\lambda = \frac{a}{1-b}, \quad \text{and} \quad \lambda' = \frac{a'}{1-b'}.$$

Direct inspection of the expressions for the coefficients of both functions, using $S_i e_i = E_i$ and $S'_i e'_i = E'_i$ shows that $\mu a = a'$ and $b = b'$ which implies (17). We know that the probabilities p_i of transmitting directly to the next slice depend only on the λ and λ' values which are fixed points of the functions $f(\lambda, E_1, \ldots, E_n) = \lambda$ and $f(\lambda', E'_1, \ldots, e'_n) = \lambda'$. By hypothesis, we know that $\mu\lambda = \lambda'$. However, in the equations determining the $p'_i s$ values only the products λE_i and $\lambda' E'_i$ are involved. Since these products are equal the result is proved. For the sake of completeness, we mention that we can also prove by induction that the values x_i and x'_i values are the same.

To conclude this section we finally state the result which shows how the energy balanced data propagation problem can be well formulated.

Proposition 2. *Given an energy assignment vector belonging to the set* $\{v \in R^n : v_i \geq 0, \| v \|= constant\}$ *then there exist unique probabilities* p_i, $i = 2, \ldots, p_n$ *to solve the energy balanced data propagation problem.*

Proof. Probabilities p_i are determined by the fixed point of the function defined by proposition 1. This fixed point is unique because the function is linear. Moreover, with our hypothesis the parameters a and b, see (11) and (12), satisfy $a > 0$ and $b < 0$ implying the existence of a fixed point.

This result shows formally that the problem is well formulated and possesses a unique solution if the energy assignment vector is restricted to belongs to the set $\{v \in R^n : v_i \geq 0, \| v \|= \text{constant}\}$.

4 A blind strategy for balanced energy dissipation

In this section, we deal with the problem of the estimation of the probabilities p_i of transmitting a data directly to a sensor which belongs to the next slice being blind to the probabilities λ_i of occurrence of events in a given slice. The blindness assumption is more general and realistic and allows the design of adaptive algorithms that appropriately adjust to the network parameters. However, we do not estimate directly the λ_i probability but directly the values of x_i (6). One reason for this is that the x_i values have probabilistic interpretation in terms of the path of the data through the different slices of the networks.

Proposition 3. *Consider an event occurring in slice* $i = 1, \ldots, n$ *with probability* λ_i. *The event is handled by the network which conveys it to the sink. Define* A_i *the event : "The data passes through slice number* i*", and* $\mathbf{1}_i$ *the indicator function of event* A_i. *Then*

$$Prob(\mathbf{1}_i = 1) = x_i. \tag{18}$$

Proof. To compute the probability that the data passes through slice i we can pass in review the different scenarios leading to the realisation of the event A_i. A necessary condition is that the event is generated in slice i, $i+1$, …, n. If we denote G_i the event : " The event is generated in slice i", we have

$$\text{Prob}(\mathbf{1}_i = 1) = \sum_{j=i}^{n} \text{Prob}(\mathbf{1}_i = 1 | G_j).$$

Because $\text{Prob}(\mathbf{1}_i = 1|G_j) = \lambda_j p_j p_{j-1} \ldots p_{i+1}$ if $j > i$ and $\text{Prob}(\mathbf{1}_i = 1|G_i) = \lambda_i$ the last equation leads to (18).

This result is useful for devising a blind strategy for balanced energy. Indeed, from the sink point of view the realisation of the events A_i can be observed if

```
Initialise x̃_0 = λ, ..., x̃_n
Initialise NbrLoop=1
repeat forever
  Send x̃_i values to the stations which compute their p_i probability
   wait for a data
   process the data
  for i=0 to n
    if the data passed through slice i then
       X ← 1
    else
       X ← 0
    end if
     Generate R a x̃_i-bernoulli random variable
     x̃_i ← x̃_i + 1/NbrLoop (X − R)
     Increment NbrLoop by one.
  end for
end repeat
```

Fig. 4. Pseudo-code for estimation of the x_i value by the sink

we assume that each sensor handling an event appends to the data associated to it the slice number the stations belongs to.

We describe the blind algorithm for energy data propagation. The algorithm does not know about the probability λ_i of occurrences of the events in the slices and indirectly estimates them. The algorithm is illustrated in Figure 4 in pseudo-code like form. The sink starts to assign values $\tilde{x}_i$ for the estimation of the x_i values and λ. For convenience, and since there are not intrinsic differences between λ and x_i we introduce the notation $x_0 = \lambda$. Each sensor is assigned a $\tilde{x}_i$ value depending on the slice number it belongs to and then computes the probability p_i of transmitting directly to the next slice using formula (8). As already mentioned, sensors add information to the propagated data to make possible for the sink to determine the slices a given data passed through. Based on these observations the sink recursively estimates the probability that the data passes through a given slice i. This probability is given by (see equation (6))

$$x_i = \lambda_i + \lambda_{i+1}p_{i+1} + \ldots + \lambda_n p_n p_{n-1} \cdots p_{i+1}.$$

Here, we used x_i without tilde to refer to the real probability of the event A_i which is the observable event. Moreover, we have seen that they can be written as (see equations (10)(14))

$$x_i = a_i + b_i\lambda, \quad i = 1, \ldots, n, \text{ and, } \lambda = x_0 = a + b\lambda.$$

This means that from the point of view of the sink an event A_i occurs with probability x_i given above.

Proposition 4. *The algorithm illustrated in Figure 4 converges in probability to the solution of the energy balanced data propagation. Precisely this means that*

$$Prob\big((\lambda_n - \frac{a}{1-b})^2 > \epsilon\big) \to 0, \quad \textit{when} \quad n \to \infty,$$

with a and b defined in Theorem 1 and λ_n defined recursively by

$$\lambda_{n+1} = \lambda_n + \frac{1}{n}(X_n - R_n),$$

with X_n a Bernoulli random variable with parameter $a + b\lambda_n$ (the observable event A_1) and R_n a Bernoulli random variable with parameter λ_n (generated internally by the sink).

Proof. Since the values $\tilde{x}_i$ for $i = 1, \ldots, n$ depend on the $\tilde{x}_0 = \lambda$ value it is enough that the algorithm converges for λ_n. First notice that

$$E(X_n - R_n|\lambda_n) = a + (b-1)\lambda_n.$$

So, we get

$$E\big((\lambda_{n+1} - \frac{a}{1-b})^2|\lambda_n\big) = (\lambda_n - \frac{a}{1-b})^2 + 2\frac{b-1}{n}(\lambda_n - \frac{a}{1-b})^2 + \frac{1}{n^2}E(X_n - R_n|\lambda_n).$$

Taking expectation in both sides of this equality we get

$$\begin{aligned} E\big((\lambda_{n+1} - \frac{a}{1-b})^2\big) =& E\big((\lambda_n - \frac{a}{1-b})^2\big) + 2\frac{b-1}{n}E\big((\lambda_n - \frac{a}{1-b})^2\big) \\ &+ \frac{1}{n^2}E(X_n - R_n|\lambda_n) \\ &= \ldots \\ &= E\big((\lambda_1 - \frac{a}{1-b})^2\big) + 2(b-1)\sum_{j=1}^{n}\frac{1}{j}E\big((\lambda_j - \frac{a}{1-b})^2\big) \\ &+ \underbrace{\sum_{j=1}^{n}\frac{1}{j^2}\underbrace{E(X_n - R_n|\lambda_j)}_{\text{bounded}}}_{\text{convergent}}. \end{aligned}$$

Since $b < 0$ and $E\big((\lambda_{n+1} - \frac{a}{1-b})^2\big) > 0$ the first sum on the right side of this equation converges. Hence,

$$E\big((\lambda_j - \frac{a}{1-b})^2\big) \to 0 \quad \text{when} \quad j \to \infty,$$

which implies the convergence in probability.

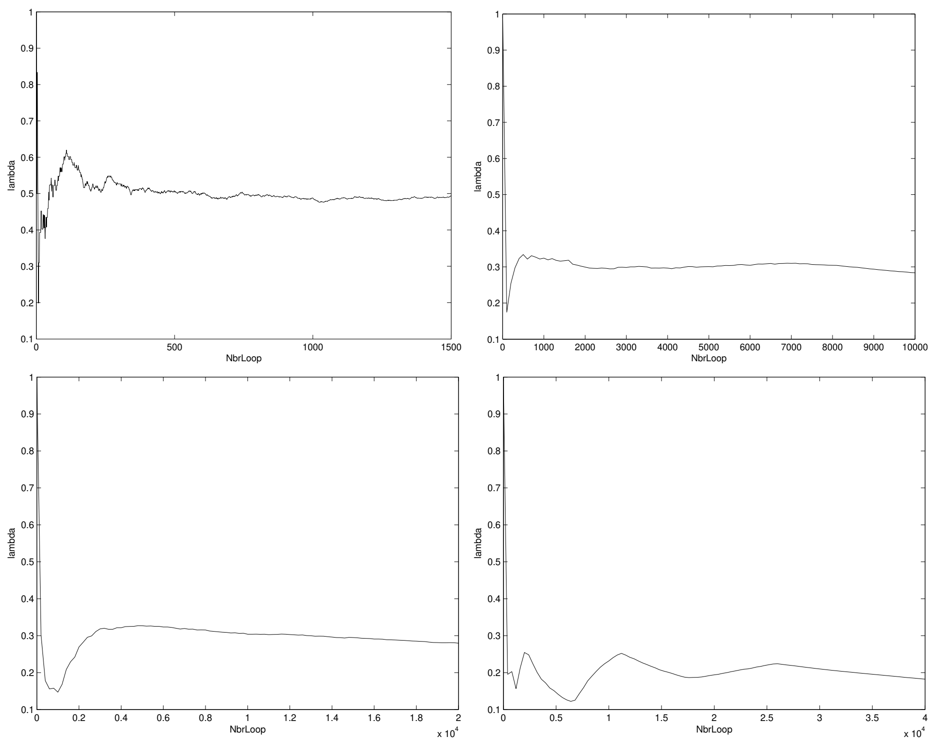

Fig. 5. Numerical experiment of algorithms (λ in terms of number of loops) illustrated in Figure 4 with $3, 10, 20, 30$ slices from left to right and top to bottom

5 Numerical experiments and conclusion

Recursive stochastic estimation procedures are very useful for solving practical problems and leads to adaptive protocols such that the one presented in Figure 4. The main drawback of these methods is the slow rate of convergence, typically of order $\mathcal{O}(1/\sqrt{n})$ and the lack of a robust stopping criterion. Intuitively, this is due to the fact that the procedure tracks the true value of the parameter to be estimated with correction of order $\mathcal{O}(\frac{1}{n})$. This implies that the procedure is robust in the sense that every possible value of the parameter is reached, but makes the estimate of the number of steps necessary very difficult.

Numerical experiments are then presented in order to validate the efficiency of the blind protocol introduced in this section. The framework we choose for our experiment is the same as the particular one described in [6]. This framework is realistic and allows us to compare our numerical results. The probability that an event occurs in slice number $i = 1, \ldots, n$ is proportional to the area of this slice and is given by $\lambda_i = \frac{2i-1}{n^2}$. We choose to deal with energy balanced data propagation and we have for the energy assignment vector

$$E(X) = \lambda\big(E_n/S_n, E_{n-1}/S_{n-1}, \ldots, E_1/S_1\big)^T = \lambda\big(1, 1, \ldots, 1\big).$$

We simulate the algorithm executed by the sink and illustrated in Figure 4. We start by arbitrarily fixing $\tilde{x}_1 = 0.5$ for $i = 2, \ldots, n$ and $\lambda = 1$. This last choice corresponding to the worst a priori estimation possible. We simulate the occurrence of the events with respect to the known probability λ_i. Notice that these probabilities are known from the simulation but are undirectly estimated by the algorithm (the sink). The path of the data generated by the event are simulated using the successive values of the probabilities p_i for $i = 2, \ldots, n$ which are computed on the basis of the $\tilde{x}_i$ values using formula (8). Once the path is simulated the sink updates the values of $\tilde{x}_i$ and a new event is generated. We proceed the simulations for 3, 10, 20 and 30 slices. These experiments are reported in figure 5. We observe from the experiments that, as expected, we quickly get a good estimation of the value of λ but need many more iterations to get high precision estimate due to the convergence of order $\mathcal{O}(1/\sqrt{n})$.

We discussed in this section the problem of balancing the consumed energy per sensors for the process of data propagation thourgh wireless sensor networks. The main novelty contained in this section is related to the very realistic hypothesis that no a priori knowledge on the probability of the region in which the events occur is known. We show that stochastic approximation methods can be applied and lead to protocols able to estimate these probabilities. Although high precision estimate needs many iterations of the estimation process, good estimations are provided by the algorithm prior to convergence.

An other important point to develop is to estimate the energy consumed before convergence of the algorithm and whether energy are wasted during this period of time or not.

A possible extension of this problem is not to consider slices of sensors but to consider all the sensors individually. In this situation, all the sensors have their own energy restriction and their own probability of observing an event. In this situation, it seems important that the algorithm can be devised in a distributed way. Indeed, if sensors are considered individually the process of broadcasting the x_i values from the sink to the sensors leads to an important traffic of data and it is likely that in this situation the impact of collisions cannot be longer ignored. Moreover, in the situation described in this section we ignore problems related to the size of the data sent to the sink. Indeed, when sensors add some information to the propagated data, such as the slice number, the size of the data can become prohibitive with respect to small memory capacity of smart sensors. In our situation it seems likely that the number of slice is of reasonable order, but if the sensors are individualised, the situation can prove to be very different

References

1. T. Antoniou, A. Boukerche, I. Chatzigiannakis, S. Nikoletseas and G. Mylonas, *A New Energy Efficient and Fault-tolerant Protocol for Data Propagation in Smart Dust Networks*, in Proc. 37th Annual ACM/IEEE Simulation Symposium (ANSS'04), IEEE Computer Society Press, pp. 43 52, 2004.
2. A. Boukerche, X. Cheng, and J. Linus, *Energy-Aware Data-Centric Routing in Microsensor Networks*, in Proc. of ACM Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), pp. 42-49, Sept 2003.
3. A. Boukerche, R. Werner, N. Pazzi and R.B. Araujo, *A Novel Fault Tolerant and Energy-Aware Based Algorithm for Wireless Sensor Networks*, First International Workshop on Algorithmic Aspects of Wireless Sensor Networks (ALGOSENSORS), Turku, Finland, Lecture Notes in Computer Sciences 3121, Springer-Verlag, July (2004).
4. I. Chatzigiannakis, T. Dimitriou, S. Nikoletseas, and P. Spirakis, *A Probabilistic Algorithm for Efficient and Robust Data Propagation in Smart Dust Networks*,in the Proceedings of the 5th European Wireless Conference on Mobile and Wireless Systems beyond 3G (EW 2004), pp. 344-350, 2004.
5. I. Chatzigiannakis, S. Nikoletseas and P. Spirakis, *Smart Dust Protocols for Local Detection and Propagation*, in the Proceedings of the 2nd ACM Workshop on Principles of Mobile Computing (POMC), ACM Press, pp. 9-16, 2002.
6. C. Efthymiou, S. Nikoletseas and J. Rolim, *Energy Balanced Data Propagation in Wireless Sensor Networks*, invited paper in the Wireless Networks (WINET, Kluwer Academic Publishers) Journal, Special Issue on "Best papers of the 4th Workshop on Algorithms for Wireless, Mobile, Ad Hoc and Sensor Networks (WMAN 2004)", to appear in 2005.
7. P. Leone, J. Rolim, *Towards a Dynamical Model for Wireless sensor Network*, First International Workshop on Algorithmic Aspects of Wireless Sensor Networks (ALGOSENSORS), Turku, Finland, Lecture Notes in Computer Sciences 3121, Springer-Verlag, July (2004).
8. H. Robbins, S. Monro, *A Stochastic Approximation Method*, Ann. Math. Stat. 22, 400-407, 1951.
9. M. Singh, 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 Communications - PERCOM, 2003.