

# Analysis and Simulation for Parameterizing the Energy-Latency Trade-off for Routing in Sensor Networks

Dionysios Efstathiou  
Research Academic Computer  
Technology Institute (CTI) and  
University of Patras, Greece  
eustathi@ceid.upatras.gr

Andreas Koutsopoulos  
Department of Computer  
Science, University of  
Paderborn, Germany  
koutsopo@mail.upb.de

Sotiris Nikolettseas  
Research Academic Computer  
Technology Institute (CTI) and  
University of Patras, Greece  
nikole@cti.gr

## ABSTRACT

We study the problem of greedy, single path data propagation in wireless sensor networks, aiming mainly to minimize the energy dissipation. In particular, we first mathematically analyze and experimentally evaluate the energy efficiency and latency of three characteristic protocols, each one selecting the next hop node with respect to a different criterion (minimum projection, minimum angle and minimum distance to the destination). Our analytic and simulation findings suggest that any single criterion does not simultaneously satisfy both energy efficiency and low latency. Towards parameterized energy-latency trade-offs we provide as well hybrid combinations of the two criteria (direction and proximity to the sink). Our hybrid protocols achieve significant performance gains and allow fine-tuning of desired performance. Also, they have nice energy balance properties, and can prolong the network lifetime.

## Keywords

Wireless sensor Networks, Algorithms, Data Propagation, Energy Efficiency, Routing, Performance Evaluation

## 1. INTRODUCTION AND CONTRIBUTION SUMMARY

A wireless sensor network (WSN) is an adhoc collection of large numbers of geographically distributed autonomous nodes that communicate over a wireless link. Each node can directly communicate with other nodes lying within its transmission range. In greedy data propagation, for a packet to reach the destination (e.g. sink, or base station), a node forwards the packet to a suitably chosen neighbor, which in turn forwards data to one of its neighbors, and so on, until the data reaches the final destination.

We study the problem of routing to a static sink by performing greedy, local next-node choices, when all the sensors in the network are also static. Our goal is, first, to mathematically analyse for the first time, experimentally evaluate

and compare the strengths and weaknesses of characteristic greedy protocols such as the Location-Aware Routing (LAR), the Direction-Aware Routing (DAR) and the Nearest with Forward Progress (NFP) Protocol. Motivated by this study, we propose two hybrid protocols (hybridization of location-aware and direction-aware) towards a satisfactory performance trade-off: a) an energy aware hybrid protocol and b) a threshold-based hybrid protocol. We note that greedy protocols must be localized and light-weight in order to fit to the resource-constrained nature of sensor networks. We do not consider protocols with guaranteed delivery.

**Related Work:** In greedy algorithms [3, 5, 6, 7, 10], a node possessing data under propagation can make a locally optimal, greedy choice (with respect to a specific local criterion) according to the information it has about its one-hop neighbors in order to forward data along a path towards the sink.

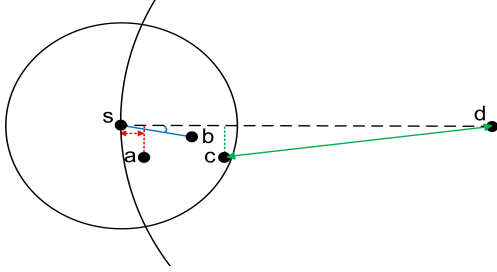
A widely used approach of greedy routing is Maximum Horizontal Progress (MHP) [9]. In MHP each node search for the neighbour who corresponds to the maximum horizontal progress, on the line between the current node and the sink. Our analysis and experiments suggest that MHP and LAR are similar and behave almost identically for medium or large densities of sensor deployment.

In the Local Target Protocol (LTP) [1], the model is slightly different than our model and that used in LAR and MHP. The sink is not a single point but a receiving “wall”  $W$  (a line segment), each node is aware of the direction towards  $W$  and no geolocation abilities are assumed. Just as normal gradient routing, the search phase of LTP finds out the direct neighbour closer to the destination (in this case  $W$ ), and the direct transmission phase sends the data out. LTP performs very good for dense networks, but its performance drops in sparse or faulty networks.

When enhancing the communication model with varying transmission range capabilities, local probabilistic algorithmic design choices can simultaneously satisfy fast forwarding and desired global network properties, such as energy balance towards prolonging the network lifetime ([2, 8]).

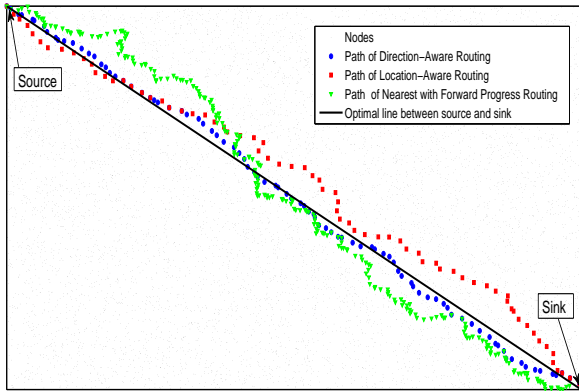
**Our approach:** We mathematically analyze and experimentally evaluate the energy efficiency of three characteristic protocols: a) the Nearest with Forward Progress (NFP) in which the node with the minimum projection is chosen

as the next hop node (node  $s$  selects node  $a$  in Fig. 1), b) the Location-Aware Routing (LAR) where the node with the minimum distance to the destination is chosen as the next hop node (node  $s$  selects node  $c$  in Fig. 1) and c) the Direction-Aware Routing (DAR) where the next hop node is the neighboring node with the minimum angle (node  $s$  selects node  $b$  in Fig. 1).



**Figure 1: Illustration of the three greedy routing protocols**

In Fig. 2 we present an example of the routing paths that are generated using Nearest with Forward Progress, Location-Aware Routing and Direction-Aware Routing in a square ( $200 \times 200$ ) network area, where the nodes are placed randomly and uniformly. The source of the data is located at  $(0, 200)$  (top left corner) and the sink is placed at  $(200, 0)$  (bottom right corner). We observe that the path generated using NFP consists of a large number of small hops (low energy consumption and very high data latency), the path of LAR consists of a small number of large hops (high energy consumption and low data latency) and the path of DAR contains less but larger hops than the NFP's path and more but shorter hops than the LAR's path.



**Figure 2: Comparing the routing paths of DAR, LAR and NFP with the optimum path**

The main advantage of LAR is that it minimizes data delivery latency but it increases a lot the energy consumption, whilst DAR reduces energy consumption over LAR's; however it increases data delivery latency.

Our analytic and simulation findings suggest that any single criterion does not simultaneously satisfy both energy

efficiency and low latency. Towards a satisfactory energy-latency trade-off, we propose two hybrid protocols (that assume different model strength) in order to use the advantages of both LAR and DAR. In the first protocol before each transmission we perform a probabilistic choice according to probability  $p_i$  (see Section 2.5). More concretely, we use location greedy routing with probability  $p_i$  and directional greedy routing with probability  $1 - p_i$ , where  $p_i$  depends on the residual energy and the distance to the sink; for large distance and high energy it tends to perform location aware choices to accelerate data propagation while for small distance and low energy it rather employs direction awareness to save energy. In the second hybrid protocol, we decide to transmit data using LAR or DAR by comparing the latency ratio between the current and the optimal data delivery latency with a predetermined threshold value  $t_{value}$ . If the current ratio is lower than the  $t_{value}$  we use location greedy routing, otherwise we use directional greedy routing, i.e. the rationale is to guarantee a certain level for the one metric while optimizing the other. Our findings demonstrate in detail the performance characteristics of each protocol and show that significant performance gains are achieved by the hybrid design. We note that the energy-aware heuristic needs stronger modelling assumptions (e.g. distance evaluation), and has better energy balance properties. Still, the threshold-based heuristic (which does not use distance knowledge) achieves similar performance.

## 2. THE ROUTING PROTOCOLS

### 2.1 The Model

We consider a two-dimensional (plane) sensor network, in which the sensors and the single sink node are static. We abstract the network by a graph  $G(V, E)$ , where  $V$  denotes the set of nodes (sensors), while  $E \subseteq V^2$  represents the set of edges (wireless links). The deployment of the sensors is random uniform and various densities (low, medium, high) are considered. An edge between two nodes in the graph exists iff the distance between the corresponding sensors in the network is below a certain limit, capturing the wireless transmission range  $R$ . The distance between nodes is the Euclidean distance, and the path length is the sum of the distances of the intermediate pairs of subsequent nodes (hops).

Nodes are localized; localization can be achieved by either GPS technology (based on localization of a few reference nodes only) or by a system of virtual coordinates. Nodes are aware only of their one-hop away (immediate) neighbours, as well as their locations. Localization allows some direction sensing capabilities e.g. nodes know the direction towards the sink, and can estimate angles around a certain direction. Finally, we assume a set-up phase initiated by the sink during which some (limited) global network information is diffused.

Our protocols operate at the network layer, so we are assuming appropriate underlying data-link, MAC and physical layers. The nodes' memory is assumed limited e.g. we allow messages to piggy-back a constant number of bits of information only, encoding the position of the last node visited and the position of the sink.

Sensor networks are characterized by high network dynam-

ics such as frequent, dense failures. So we investigate the detailed impact of failures on protocol performance. In particular, for each unit of time, failures occur at randomly chosen nodes, instantly, and no further computation and/or communication can be performed by these failed nodes. We examine a broad set of failure probabilities, including low, medium and high ones.

## 2.2 The Location-Aware Routing Protocol

A commonly used method of location-aided (or position-based) greedy routing, is Location-Aware Routing (LAR) [12] and its variations. In LAR each node basically forwards a packet to the neighbour that is closest to the destination. LAR provides a way to deliver a packet to a destination location (e.g. to a static sink), based only on local information without the need for any extra infrastructure or information. For this reason, the LAR protocol is very suitable and efficient in resource-constrained wireless networks, such as WSNs. In this single-path multi-hop routing protocol, each node needs to know only the location information of its direct neighbours in order to forward data packets. This approach basically attempts to find the shortest path to the destination, in terms of either distance or number of wireless hops towards minimizing data propagation latency. As LAR tries to find the next node that is the closest to the sink in order to transmit the message to, a node considers only those neighbours that are closer to the destination than itself. Sensor nodes can calculate their position in some common co-ordinate system (e.g., by using navigational equipment or running a virtual coordinates algorithm) and are aware of the sink's position. The above assumptions can be relaxed in many ways. We note that LAR has not been analyzed mathematically as far as energy efficiency in sensor context is concerned.

## 2.3 The Direction-Aware Routing Protocol

The main goal of DAR is to minimize energy consumption following an efficient path, which is as close as possible to the optimal (wrt both energy and latency) direct line that connects the current node to the sink. The basic idea in the protocol is doing shorter transmissions of “nice” direction to neighbours that are closer to the current node, in order to keep the energy consumption at low levels (the energy consumed to transmit a message is assumed proportional to some power of the distance between sender and receiver). As mentioned above, the main idea is to transmit the data to the node that has the minimum divergence (in terms of angle of direction) from the optimal transmission line which is the line that connects the current node to the sink. More precisely, the neighbouring node that forms the minimum angle  $a_i$  between the line that connects itself with the current node and the line that connects the current node with the sink is chosen, see Fig. 3. We note that although the selection criterion is not based on progress to the sink, the fact that data propagation is kept within the optimal transmission zone leads indirectly to not increasing the path length and number of hops too much. DAR is in fact very similar to the Compass Routing protocol [7]; however [7] analyzes the correctness and time efficiency of the protocol while the energy efficiency is not studied.

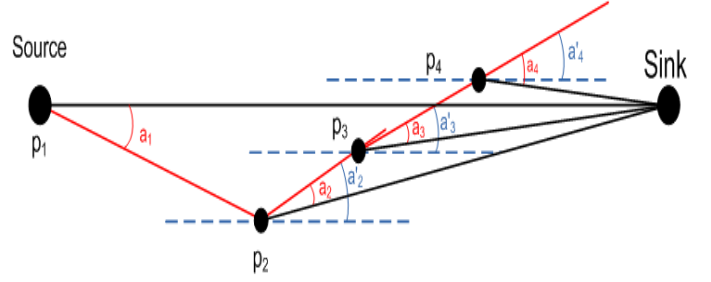


Figure 3: Direction-aware routing example.

## 2.4 The Nearest with Forward Progress Protocol

Nearest with Forward Progress (NFP) [4] is an energy-aware protocol which tries to minimize the energy consumption by sending the message to the closest node in the direction of the sink. The main advantage of NFP is that it makes collisions less likely as a node will adjust its transmission power to be just strong enough to reach the nearest neighbor which will result in forward progress. This leads to a succession of a large number of small hops (thus high data delivery latency) and less energy consumption than long hops. We note that NFP has not been analyzed mathematically.

## 2.5 The Hybrid Routing Protocols

To handle the arising performance trade-off between energy and latency, we propose two hybrid heuristic routing protocols. In the first protocol each node decides whether to forward data using LAR or DAR by taking into account the distance of the node from the sink and the residual energy of the node. The second protocol takes its decision based on a threshold choice criterion, i.e. it tries to optimize one metric as long as the performance of the other one is kept above a certain threshold that can be set by the network implementor.

### 2.5.1 The Energy-Aware Hybrid Routing Protocol

Each node forwards data with a probability  $p_i$  to that neighbour which is closest to the sink (using LAR) and with probability  $1 - p_i$  to that neighbour which forms the minimum angle between the line that connects itself with the current node and the line that connects the current node with the sink (using DAR).

$$\text{Forwarding decision} = \begin{cases} \text{LAR,} & \text{with probability } p_i \\ \text{DAR,} & \text{with probability } 1 - p_i \end{cases}$$

We take  $p_i = \sqrt{\frac{D_i}{D_{max}}} \cdot \frac{E_{current}}{E_{initial}}$  where  $D_i$  is the distance of the node from the sink,  $D$  is the “dimension” of the network (e.g. the “width” of the rectangular network region),  $D_{max} = \sqrt{2} \cdot D$  is a node’s maximum distance from the sink for the rectangular network region  $D \times D$ ,  $E_{current}$  is the residual energy of the node and  $E_{initial}$  is the initial energy of the node. The rationale of *Hybrid forwarding* is to forward data using LAR in order to accelerate data transmission when data is far away from the sink and has much residual energy. The opposite happens when data is close to the sink or has little residual energy, since then data

tends to be forwarded using DAR so as to reduce energy consumption in the critical, bottleneck region close to the sink. This bottleneck is created because the sensors close to the sink always relay the data for other sensors, resulting in an unequal distribution of network residual energy.

### 2.5.2 The threshold-based Hybrid Routing Protocol

In this protocol we wish to guarantee the minimum energy consumption for a predetermined latency ratio between the current and the optimal data delivery latency.

$$\text{Forwarding decision} = \begin{cases} \text{LAR,} & \text{if } \text{Latency}_{\text{ratio}} < t_{\text{value}} \\ \text{DAR,} & \text{if } \text{Latency}_{\text{ratio}} \geq t_{\text{value}} \end{cases}$$

We define  $\text{Latency}_{\text{ratio}}$  as follows:  $\text{Latency}_{\text{ratio}} = \frac{\text{Latency}_{\text{current}}}{\text{Latency}_{\text{optimum}}}$  where  $\text{Latency}_{\text{current}}$  is the total packet travel time from source to the current node,  $\text{Latency}_{\text{optimum}}$  is the optimum packet travel time from source to the current node and  $t_{\text{value}}$  is the predefined threshold value which can take values between 0 and 1, e.g. by choosing  $t_{\text{value}} = 0.75$  we declare that  $\text{Latency}_{\text{current}}$  at every routing step must be at least 75% of the  $\text{Latency}_{\text{optimum}}$ .

The rationale of this threshold-based protocol is to forward data using cheap but slow transmissions (DAR) when the  $\text{Latency}_{\text{ratio}}$  is higher than the  $t_{\text{value}}$ . This means that we are able to conserve energy by transmitting to a short-hop neighbor without increasing data latency. On the other hand, we forward data using fast but expensive transmissions (LAR) when the  $\text{Latency}_{\text{ratio}}$  is lower than the  $t_{\text{value}}$ , in order to improve the latency and approach the desirable threshold value  $t_{\text{value}}$ .

## 3. RIGOROUS PERFORMANCE ANALYSIS

We below analyse some important performance properties (energy dissipation, data propagation latency) of each of the three basic routing methods.

### 3.1 Analysis of the Direction-Aware Routing Protocol

Let  $S_1$  be the current node possessing data and  $S_2$  the next node.

*Definition 1.* Let  $a_i$  or  $a_{\min}$  (as depicted in Fig. 3) be the angle between the line that connects the current node to the sink and the line that connects the current node with the next node of the propagation path. It is equal to the minimum angle of all angles that are formed between the line that connects each neighbour with the current node  $i$  and the line that connects the current node with the sink.

This angle corresponds to the direction of the node that is chosen for the data propagation.  $a_i$  is a random variable since the neighbouring nodes are positioned at random, so the angles are also random.

Angle  $a'_i$ , as depicted in Fig. 3 is the angle between the line that connects the current node with the next node of the propagation path, with the line that passes through the current node and is parallel to the source-destination line.

To simplify our analysis, we assume that the line  $x_2$  (see Fig. 4) is parallel to the line  $x_1$  as:

$$\sin(\theta) = \frac{p}{x_1} \leq \frac{R}{k \cdot R} = \frac{1}{k} \rightarrow_{k \rightarrow \infty} 0.$$

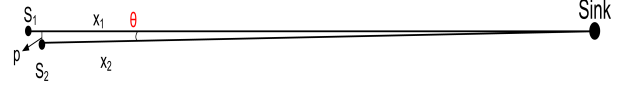


Figure 4:  $x_2$  is parallel to the line  $x_1$  example.

where  $p$  is the vertical distance from the line  $x_1$  to the line  $x_2$ ,  $p$ 's length is  $O(R)$  and  $x_1$ 's length is  $O(k \cdot R)$ , where  $k$  indicates how many times longer is the source-destination line than radius  $R$ , where  $R$  is the maximum transmission range. So it is  $a'_i \simeq a_i$ .

*Definition 2.* Let  $n$  be the number of neighbours a node has, that is the number of nodes are positioned on the interior of the circular disc with the node as its center and  $R$  as its radius. Let  $\text{den}$  be the density of sensors in the random uniform placement. So:

$$n = \pi \cdot R^2 \cdot \text{den}.$$

*Definition 3.* Let  $a_{ij}$  be the random variable that is equal to the angle corresponding to the neighbour  $j$  of a node  $i$ .

In order to evaluate the performance of the direction-aware protocol we first must calculate the expected value of this minimum angle.

$$F_{a_i}(x) = P(a_i \leq x) = 1 - P(a_i \geq x) = 1 - P(a_{i_1} \geq x, a_{i_2} \geq x, \dots, a_{i_n} \geq x)$$

The random variables  $a_{ij}$  are independent, since the nodes are positioned independently of each other and uniformly at random. As we assume that we investigate nodes only in the positive halfcircle of the neighbourhood region, the random variables are uniformly distributed in  $[0, \frac{\pi}{2}]$ . So:

$$F_{a_i}(x) = 1 - P(a_{i_1} \geq x) \cdot P(a_{i_2} \geq x) \cdots P(a_{i_n} \geq x) = 1 - \left(1 - \frac{x}{\frac{\pi}{2}}\right)^n = 1 - \left(1 - \frac{2 \cdot x}{\pi}\right)^n.$$

The probability density function of random variable  $a_i$  is:

$$f_{a_i}(x) = \frac{dF_{a_i}(x)}{dx} = \frac{2 \cdot n}{\pi} \cdot \left(1 - \frac{2 \cdot x}{\pi}\right)^{n-1}.$$

We now calculate the expected value of random variable  $a_i$ :

$$\begin{aligned}
E(a_i) &= \int_0^{\frac{\pi}{2}} x \cdot \frac{2 \cdot n}{\pi} \left(1 - \frac{2 \cdot x}{\pi}\right)^{n-1} dx = \\
&= \left[ -x \cdot \left(1 - \frac{2 \cdot x}{\pi}\right)^n \right]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} \left(1 - \frac{2 \cdot x}{\pi}\right)^n \cdot dx \\
&= 0 - \left[ \frac{\left(1 - \frac{2 \cdot x}{\pi}\right)^{n+1}}{2 \cdot (n+1)} \right]_0^{\frac{\pi}{2}} = \frac{\pi}{2 \cdot (n+1)}.
\end{aligned}$$

Thus we get the following lemma:

LEMMA 1. *The expectation of angle  $a_i$  is*

$$E(a_i) = \frac{\pi}{2 \cdot (n+1)}.$$

*Definition 4.* Let  $d_i$  be the random variable that is equal to the euclidean distance between the current node  $i$  and the node (from the set of neighbours of the current node) the protocol chooses to be the one the data is propagated to.

We calculate the distribution function of random variable  $d_i$ :

$$F_{d_i}(x) = P(d_i \leq x) = \frac{\pi \cdot x^2}{\pi \cdot R^2} = \frac{x^2}{R^2}.$$

The probability density function of random variable  $d_i$  is:

$$f_{d_i}(x) = \frac{dF_{d_i}(x)}{dx} = \frac{2 \cdot x}{R^2}.$$

We now calculate the expected value of random variable  $d_i$ :

$$E(d_i) = \int_0^R \frac{2 \cdot x^2}{R^2} dx = \left[ \frac{2 \cdot x^3}{3 \cdot R^2} \right]_0^R = \frac{2}{3} \cdot R \quad (1)$$

Also, we have:

$$E(d_i^2) = \int_0^R \frac{2 \cdot x^3}{R^2} dx = \left[ \frac{2 \cdot x^4}{4 \cdot R^2} \right]_0^R = \frac{1}{2} \cdot R^2.$$

Thus we get the following:

LEMMA 2. *The expectation of distance  $d_i$  is:*

$$E(d_i) = \frac{2}{3} \cdot R$$

while the expectation of  $d_i^2$  is

$$E(d_i^2) = \frac{1}{2} \cdot R^2$$

THEOREM 1. *Let the source-destination distance be  $dis = k \cdot R$ . The mean number of hops required for direction-aware routing is upper bounded by the quantity*

$$\frac{3 \cdot k}{2 \cdot \left(1 - 2 \cdot \sin \frac{\pi}{4 \cdot (n+1)}\right)}$$

where  $n$  is the expected number of sensors within each transmission range of radius  $R$ .

PROOF. First we must calculate the expected value of random variable  $x_i$ , where  $x_i$  is the projection of the line segment  $(p_i, p_{i+1})$  on the line which is parallel to the line from the source  $S$  to the sink. We use the expectation of  $a_i$  and the expectation of  $d_i$ , as derived in Lemmata 1 and 2. Note that  $d_i$  and  $a_i$  are independent since the distance  $d_i$  of a neighbour to the current node does not give us any information about the angle. So, we have:

$$E(x_i) = E(d_i \cdot \cos a_i) = E(d_i) \cdot E(\cos a_i). \quad (2)$$

$$\begin{aligned}
E(\cos a_i) &= E\left(1 - 2 \cdot \sin^2 \frac{a_i}{2}\right) = 1 + E\left(-2 \cdot \sin^2 \frac{a_i}{2}\right) \geq \\
&1 + E\left(-2 \cdot \sin \frac{a_i}{2}\right).
\end{aligned}$$

Because  $-2 \cdot \sin \frac{a_i}{2}$  is convex as  $f''(-\sin \frac{a_i}{2}) = \frac{\sin \frac{a_i}{2}}{4} > 0$  for  $a_i \in [0, \frac{\pi}{2}]$ ,

we get from Jensen's inequality:

$$1 + E\left(-2 \cdot \sin \frac{a_i}{2}\right) \geq 1 - 2 \cdot \sin \frac{E(a_i)}{2} = 1 - 2 \cdot \sin \frac{\pi}{4 \cdot (n+1)}. \quad (3)$$

So (1) through (2) and (3)  $E(x_i) = \frac{2}{3} \cdot R \cdot \left(1 - 2 \cdot \sin \frac{\pi}{4 \cdot (n+1)}\right)$

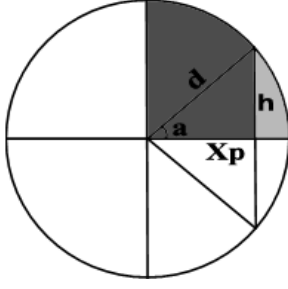
Let  $h$  be the number of hops needed to get to the destination. Clearly:

$$\sum_{i=1}^h x_i \leq k \cdot R < \sum_{i=1}^{h+1} x_i$$

so  $h$  is a stopping time for the sequence of random variables  $x_i$ . So from Wald's equation on the expectation of a random number of random variables ([11]) we get:

$$E(h)E(x_i) \leq k \cdot R < E(h+1)E(x_i)$$

so



**Figure 5: Distribution of  $X_p$  in Location Aware Protocol**

$$E(h) \leq \frac{k \cdot R}{E(x_i)}$$

and this concludes the proof.  $\square$

Now we calculate the expected value of energy consumption of the protocol. Let  $E_n$  be the r.v. that is equal to the energy consumed at a data propagation. Since we assumed that the energy consumed for a hop is proportional to the second power of the distance traversed, we have:

$$E[E_n] = O\left(E\left[\sum_{i=1}^h d_i^2\right]\right)$$

We also get (via Wald's equation, similarly to Theorem 1) that:

$$E[E_n] = O(E[h]E[d^2]) \leq O\left(\frac{3 \cdot k}{2 \cdot \left(1 - 2 \cdot \sin \frac{\pi}{4 \cdot (n+1)}\right)} \frac{R^2}{2}\right)$$

### 3.2 Analysis of the Location-Aware Routing Protocol

Let us now analyse the location aware algorithm in terms of hop count and energy consumption. Let  $X_p$  be the random variable that is equal to the length of the projection of the distance traversed by a hop to the line that connects the source to the sink. We show the following.

**LEMMA 3.** *If the source-destination distance is  $dis = k \cdot R$  the mean number of hops  $h$  required to reach the destination is equal to the quantity*

$$\int_0^R 2n \left(1 - \frac{2 \arccos\left(\frac{x}{R}\right) - \frac{2x\sqrt{1-\frac{x^2}{R^2}}}{R}}{\pi}\right)^{n-1} \left(\frac{1}{R\sqrt{1-\frac{x^2}{R^2}}} + \frac{\sqrt{1-\frac{x^2}{R^2}}}{R} - \frac{x^2}{R^3\sqrt{1-\frac{x^2}{R^2}}}\right) dx$$

where  $n$  is the number of sensors in each transmission range of radius  $R$ .

**PROOF.** First we must compute the distribution function of r.v.  $X_p$ .

Let AreaA be the area that is seen as dark-grey at the Fig. 5 and AreaB the light-grey area. The distribution function of  $X_p$  is the following:

$$F(x) = P(X_p \leq x) = \frac{AreaA}{AreaA + AreaB}.$$

Obviously, dark-grey and light-grey area form a quadrant so:

$$AreaA + AreaB = \frac{\pi R^2}{4}.$$

$$\begin{aligned} AreaB &= \frac{R^2}{4} (2a - \sin 2a) = \\ &= \frac{R^2}{4} \left(2 \arccos \frac{X_p}{R} - 2 \sin a \cos a\right) = \\ &= \frac{R^2}{4} \left(2 \arccos \frac{X_p}{R} - 2 \sin \left(\arccos \frac{X_p}{R}\right) \cos \left(\arccos \frac{X_p}{R}\right)\right) \\ &= \frac{R^2}{4} \left(2 \arccos \frac{X_p}{R} - 2 \sin \left(\arcsin \sqrt{1 - \left(\frac{X_p}{R}\right)^2}\right) \frac{X_p}{R}\right) \\ &= \frac{R^2}{4} \left(2 \arccos \frac{X_p}{R} - 2 \left(\sqrt{1 - \left(\frac{X_p}{R}\right)^2}\right) \frac{X_p}{R}\right). \end{aligned}$$

$$\begin{aligned} F(x) &= P(X_p \leq x) = 1 - P(X_p \geq x) = 1 - \frac{AreaB}{AreaA + AreaB} \\ &= 1 - \frac{2}{\pi} \left(\arccos \frac{X_p}{R} - \left(\sqrt{1 - \left(\frac{X_p}{R}\right)^2}\right) \frac{X_p}{R}\right). \quad (4) \end{aligned}$$

Then, we compute the distribution function of the maximum projection from all the projections that are formed with every neighbour of a current node, in other words, the length of the hop in a location-aware transmission:

$$\begin{aligned} F_{X_{pmax}}(x) &= P(X_{pmax} \leq x) = (P(X_p \leq x))^n \\ &= \left(1 - \frac{2}{\pi} \left(\arccos \frac{X_p}{R} - \left(\sqrt{1 - \left(\frac{X_p}{R}\right)^2}\right) \frac{X_p}{R}\right)\right)^n \end{aligned}$$

So,  $f_{X_{pmax}}(x) = F'_{X_{pmax}}(x) =$

$$2n \left(1 - \frac{2 \arccos\left(\frac{x}{R}\right) - \frac{2x\sqrt{1-\frac{x^2}{R^2}}}{R}}{\pi}\right)^{n-1} \left(\frac{1}{R\sqrt{1-\frac{x^2}{R^2}}} + \frac{\sqrt{1-\frac{x^2}{R^2}}}{R} - \frac{x^2}{R^3\sqrt{1-\frac{x^2}{R^2}}}\right)$$

The expected value is

$$E[X_{pmax}] = \int_0^R x f_{X_{pmax}}(x) dx.$$

$$\text{and } E[X_{pmax}^2] = \int_0^R x^2 f_{X_{pmax}}(x) dx.$$

If now we take into account that  $E[h] = \frac{dis}{E[Xp]}$  we have completed our proof.  $\square$

At the end of this section we numerically estimate the integral. Finally, the energy consumed by the location aware protocol can be computed as follows:

$$E[En] = O(E[h]E[d^2]) \geq O(E[h]E[Xp_{max}^2]).$$

### 3.3 Analysis of the Nearest with Forward Progress Routing

Let us now focus at the analysis of the Nearest with Forward Progress Routing. Let again  $Xp$  be the random variable that is equal to the length of the projection of the distance traversed by a hop to the line that connects the source to the sink. In the case of computing the expected hop count, we will conclude to a similar Lemma to Lemma 3 in the previous section.

LEMMA 4. *If the source-destination distance is  $dis = k \cdot R$  the mean number of hops  $h$  required to reach the destination in the case of the Nearest with Forward Progress Routing is equal to the quantity*

$$\int_0^R x^n \left( \frac{2 \arccos(\frac{x}{R})}{\pi} - \frac{2x \sqrt{1 - \frac{x^2}{R^2}}}{\pi R} \right)^{n-1} \left( \frac{2}{\pi R \sqrt{1 - \frac{x^2}{R^2}}} + \frac{2 \sqrt{1 - \frac{x^2}{R^2}}}{\pi R} - \frac{2x^2}{\pi R^3 \sqrt{1 - \frac{x^2}{R^2}}} \right) dx$$

where  $n$  is the number of sensors in each transmission range of radius  $R$ .

PROOF. We have shown in the previous section that the distribution function of  $Xp$  is the following:

$$F(x) = 1 - \frac{2}{\pi} \left( \arccos \frac{Xp}{R} - \left( \sqrt{1 - \left( \frac{Xp}{R} \right)^2} \right) \frac{Xp}{R} \right). \quad (5)$$

We now compute the distribution function of the minimum projection from all the projections that are formed with every neighbour of a current node, in other words, the length of the hop in transmission of the Nearest with Forward Progress Routing Protocol:

$$F_{Xp_{min}}(x) = P(Xp_{min} \leq x) = 1 - P(Xp_{min} \geq x) =$$

$$1 - (P(Xp \geq x))^n = 1 - (1 - P(Xp \leq x))^n$$

$$= 1 - \left( \frac{2}{\pi} \left( \arccos \frac{Xp}{R} - \left( \sqrt{1 - \left( \frac{Xp}{R} \right)^2} \right) \frac{Xp}{R} \right) \right)^n$$

It follows that the probability distribution function of  $Xp$  is

$$f_{Xp_{min}}(x) = F'_{Xp_{min}}(x) =$$

$$n \left( \frac{2 \arccos(\frac{x}{R})}{\pi} - \frac{2x \sqrt{1 - \frac{x^2}{R^2}}}{\pi R} \right)^{n-1} \left( \frac{2}{\pi R \sqrt{1 - \frac{x^2}{R^2}}} + \frac{2 \sqrt{1 - \frac{x^2}{R^2}}}{\pi R} - \frac{2x^2}{\pi R^3 \sqrt{1 - \frac{x^2}{R^2}}} \right)$$

The expected value is

$$E[Xp_{min}] = \int_0^R x f_{Xp_{min}}(x) dx \quad (6)$$

The expected number of hops are:  $E[h] = \frac{dis}{E[Xp_{min}]}$  and so the proof is completed.  $\square$

Let us now focus at the expected energy consumption:

$$E[En] = O(E[h]E[d^2])$$

To compute the energy, we first define the random variable  $Z = d^2 = Xp_{min}^2 + H^2$  (where  $H$  and  $Xp_{min}$  are both random variables, and  $H$  is equal to the height of the node, as seen in Fig. 5). We now compute  $E[Z]$ .

From probability theory we know that if we have a random variable  $Z$ , that is a function of two other random variables  $Xp_{min}$  and  $H$ , so that  $Z = g(Xp_{min}, H)$  the expectation of  $Z$  is given by the formula

$$E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(xp_{min}, h) f_{Xp_{min}, H}(xp_{min}, h) dxp_{min} dh \quad (7)$$

So, we have  $E[Z] =$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(xp_{min}, h) f_H(h|xp_{min}) f_{Xp_{min}}(xp_{min}) dxp_{min} dh \quad (8)$$

We already have computed  $f_{Xp_{min}}(x)$  above. From Fig. 5 we can see that  $h$  reaches its maximum value when  $d=1$ .

When  $d = 1$ , we have  $1 = Xp_{min}^2 + h^2$ , it follows that  $h = \sqrt{1 - Xp_{min}^2}$ .

So  $h$  is uniformly distributed in  $[0, \sqrt{1 - Xp_{min}^2}]$  and

$$f_H(h|xp_{min}) = \frac{1}{\sqrt{1 - Xp_{min}^2}}.$$

### 3.4 Numerical Evaluation

Now we have all the information needed to compute the double integral and get the expected value. In fact, the double integral, as well as the integral in the Lemma and the integrals in the previous section, with are used to compute the expected values, are precisely estimated with an numerical integration technique, in particular the adaptive Simpson's method via Matlab, having an error less than  $10^{-6}$ . W.l.o.g. we set  $R = 1$  and we get the upper bound on the average number of hops in the case of the direction aware protocol and the estimation on the other values discussed above, when the distance is equal to 100. We do this for different values of  $n$ , as seen in the Fig. 6 and Fig. 7.



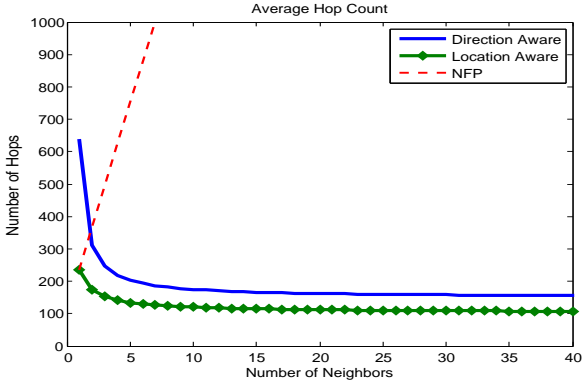


Figure 6: Average Hop Count Analysis.

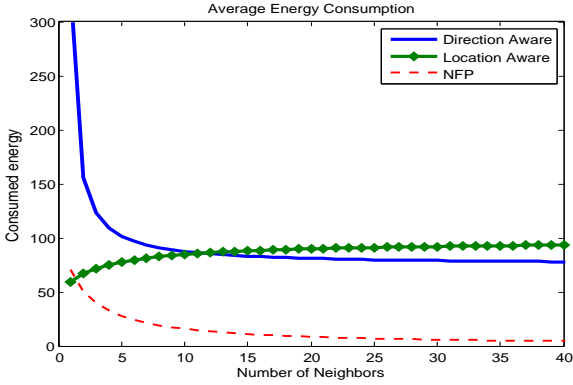


Figure 7: Average Energy Consumption Analysis.

In particular, in the Fig. 6 and Fig. 7 we can see two upper bounds (on the average energy consumption and the average hop count) of the direction aware protocol, a lower bound on the average energy consumption and an accurate estimation of the average hop count of the location aware protocol. We can see also an estimation of the average hop count and average energy consumption of the NFP. We see that the upper bound of the average number of hops in the direction aware protocol is in the  $[160, 180]$  regime, i.e. 1.6 - 1.8 times the network dimension ( $k = 100$ ). We note that in routing with fixed transmission range a number of hops 100 to get to the sink represents a minimum needed. This optimal value can only be achieved if we can always get a next hop sensor on the direct source-destination line at pairwise distances  $R = 1$ . However by examining the energy consumption we observe that the energy consumption upper bound in the direction aware protocol is low, especially as the network density increases. In contrast to that, the location aware protocol energy consumption approaches a high value, as the density gets higher. Overall, the analysis validates the properties of our protocol, namely that it saves a lot of energy while increasing latency, compared to the location aware protocol.

The results also show us that the NFP is an extreme case. As the number of neighbours increases, the hop count increases extremely as the energy consumption falls at an extremely low point. One can explain this, by considering the the NFP

is always choosing the neighbour with the smallest horizontal projection, so when are many neighbours, there will be extremely many small hops made, which would also result to an extremely low energy consumption.

## 4. EXPERIMENTAL EVALUATION

### 4.1 The simulation environment

Our simulation environment for making the experiments is the environment of Matlab 7.9.0. We deploy uniformly at random nodes in the network area. We choose as a communication model the unit disc graph. This means that each node is able to send a message to another iff the distance between them is at most a given threshold (in particular, the wireless transmission range  $R$  is taken 5). Using the unit disc graph means that the expected number of neighbours per node is close to  $d\pi$ , where  $d$  is the global density of sensors in the network.

In detail, the network area is rectangular, with length and width equal to 40 units. We apply several times the deployment of nodes in the network and repeat the experiments, for statistical smoothness. We consider different numbers of nodes (with a range from 5000 up to 30000 nodes), forming each time a network of different density  $d$  (diversely ranging from 10 to 60). For statistical reasons, we take 50 random uniform deployments and for each deployment 1000 data propagations are simulated and the average value is taken. The statistical analysis of the findings (the median, lower and upper quartiles, outliers of the samples) demonstrates very high concentration around the mean, so in the figures 8 and 9 we only depict average values.

For each deployment the source is chosen from the nodes of the network randomly uniformly and the sink is placed at the center  $(100, 100)$  of the  $200 \times 200$  deployment area. We measure the average number of hops needed to reach the sink, the average energy consumed in the network and the average success rate of each algorithm. The success rate is taken as the percentage of generated events that are reported to the sink. We assume an energy model, in which the energy consumed by a message transmission between two nodes is considered the square of the distance between the nodes. For our three metrics the average is taken over all sensor deployments and algorithm's repetitions. Also, the initial energy available at the sensor devices was set to 1000 energy units at the start of the simulation.

We investigate the performance of the protocols in the presence of permanent node failures during protocol evolution. We study the characteristic cases where 10%, 25% and 50% of the network nodes fail during the simulation time. For each unit of time, failures occur at randomly chosen nodes, instantly, and no further computation and/or communication can be performed by these failed nodes.

### 4.2 Findings

We compare our Hybrid protocols with the Nearest with Forward Progress, Direction-Aware and baseline Location-Aware protocol.



#### 4.2.1 No failures

We examine the mean number of hops the algorithms need to reach the sink (only the successful trials count here) and Fig. 8 depicts this metric. We note that, in terms of data packet latency NFP has the highest latency, DAR has the second higher latency and LAR is the fastest way to deliver the data to the sink. Our Hybrid protocols, are in-between the latency values of the DAR and the LAR. That is an expected result, since in the hybrid protocols, a fraction of the hops are made through the LAR and the remaining hops through the DAR. More specific, the Threshold Hybrid protocol is a little bit slower, in terms of latency, than the Energy Hybrid protocol.

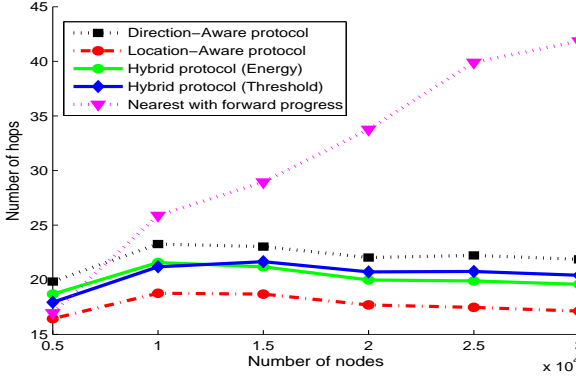


Figure 8: Data latency - no failures.

We now compare the protocols' performance regarding the energy consumption, Fig. 9. In this figure we can see that NFP protocol is the most energy conservative in contrast to the LAR which is the most energy consuming. This result is also expected, since the NFP, as mentioned before, chooses as next-hop node the node with the minimum projection. Short hops means less energy, since the energy consumed in a hop is considered to be proportional to the second power of the wireless transmission distance. DAR is the second energy conservative protocol as it makes short hops with a path close to the optimal line that connects the current node with the sink. Regarding the hybrid protocols, the values of the findings are again in-between that of the DAR and the LAR for the same reasons mentioned in the paragraph above. However, the Threshold Hybrid protocol is cheaper than the Energy Hybrid protocol and consumes almost as much energy as DAR.

In general, the NFP protocol increases dramatically the data latency while it is very cheap in terms of energy consumption due to the short-range transmissions of data messages. The DAR protocol increases the latency, but it has better results regarding the energy consumption. More specifically, we can see from the figures that the gain we have (in percentage) in energy is the same as the loss we have in latency and this is the basic trade-off. So we can conclude that DAR could be used as a routing strategy in applications where we can afford a small increase in latency, but the energy consumption is the most important factor. NFP could be used in applications where latency is not critical factor and our main goal is to conserve energy.

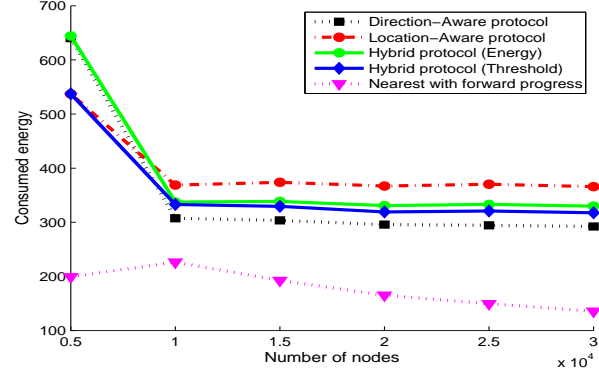


Figure 9: Energy consumption - no failures.

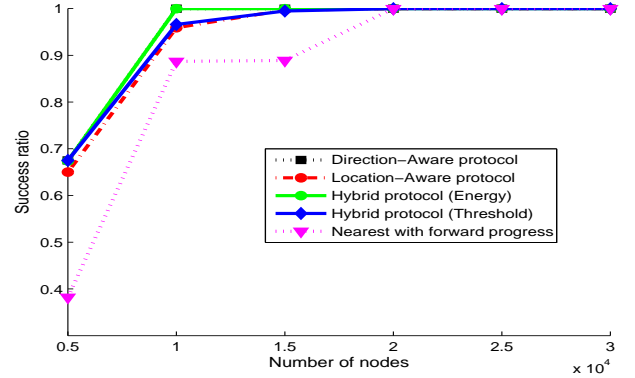


Figure 10: Success ratio - no failures.

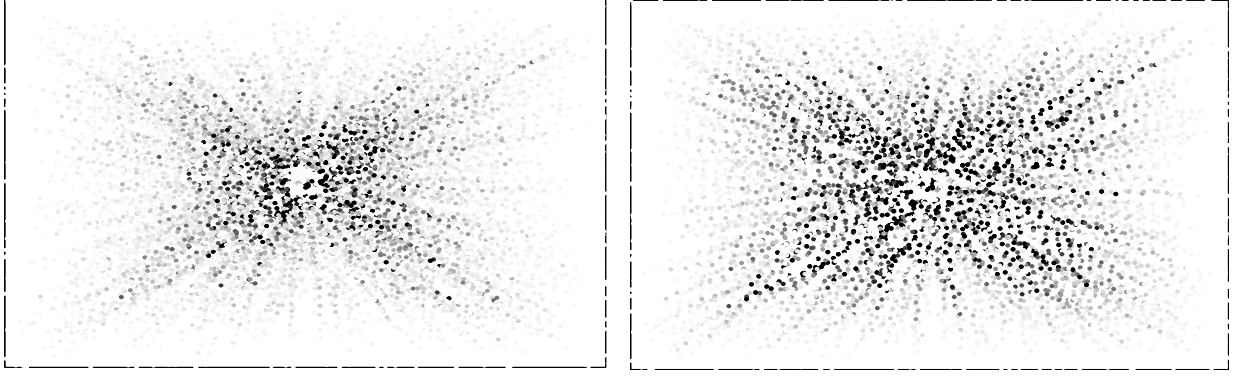
In Fig. 10 we observe that the success delivery ratio is almost identical for DAR, LAR and Hybrid protocols. For sparse networks NFP has the worse behavior. In contrast with the DAR protocol where the next hop node is the node with the minimal directional divergence from the optimal source-destination line, in NFP there is no guarantee that the chosen next-hop node will be close to the optimal source-destination line, so the followed path can deviate a lot from reaching the sink and the data can be trapped in a routing hole. For high densities (over 20.000 nodes) the success ratio is in fact 1. This is because the network is dense enough, so there is a path from every possible source towards the sink.

#### 4.2.2 The impact of failures

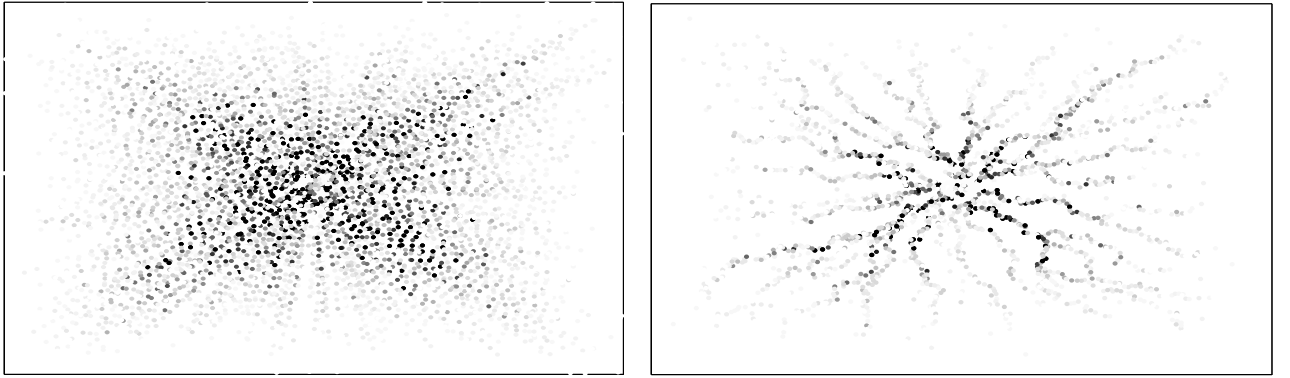
Secondly, we examine the cases where the 10%, 25% and 50% of the network nodes fail during the simulation time. In Figures 11 - 19 we depict the average number of hops needed to reach the sink, the average energy consumed in the network and the average success rate of each algorithm. We observe that the failure rate mainly affects the success rate of the protocols, as expected. While the absolute values of the energy consumption and the delivery delay are increased, we notice a consistent behaviour of all five protocols when the failure rate increases.

#### 4.2.3 Energy Balance Aspects

Third, in Fig. 20 and Fig. 21 we present the spatial evolution of energy dissipation in a network of 30.000 nodes



**Figure 20:** Energy dissipation in nodes using Energy-based Hybrid protocol (left) and Location-Aware Protocol (right).



**Figure 21:** Energy dissipation in nodes using Direction-Aware Protocol (left) and Nearest with Forward Progress (right).

Energy Units	0-40	40-80	80-120	120-160	160-200	200-240	240-280	280-320	320-360	360-400	400-440	440-1000
No. of nodes	21442	1716	1755	892	536	316	239	151	109	81	72	234

**Table 1:** Distribution of node's energy dissipation using the energy-based Hybrid Protocol.

Energy Units	0-40	40-80	80-120	120-160	160-200	200-240	240-280	280-320	320-360	360-400	400-440	440-1000
No. of nodes	23733	1236	886	543	390	252	249	196	175	141	116	554

**Table 2:** Distribution of node's energy dissipation using the LAR Protocol.

Energy Units	0-40	40-80	80-120	120-160	160-200	200-240	240-280	280-320	320-360	360-400	400-440	440-1000
No. of nodes	24728	897	889	545	369	277	186	153	92	86	58	388

**Table 3:** Distribution of node's energy dissipation using the DAR Protocol.

Energy Units	0-40	40-80	80-120	120-160	160-200	200-240	240-280	280-320	320-360	360-400	400-440	440-1000
No. of nodes	27477	518	439	280	174	115	83	38	38	26	19	94

**Table 4:** Distribution of node's energy dissipation using the NFP Protocol.

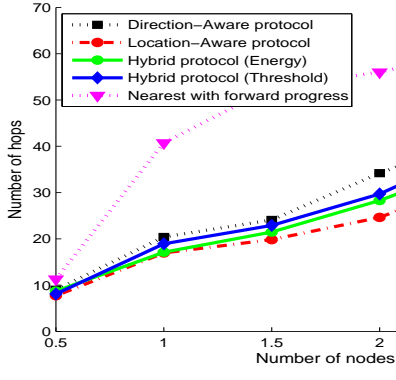


Figure 11: Data latency - 10% Failed Particles.

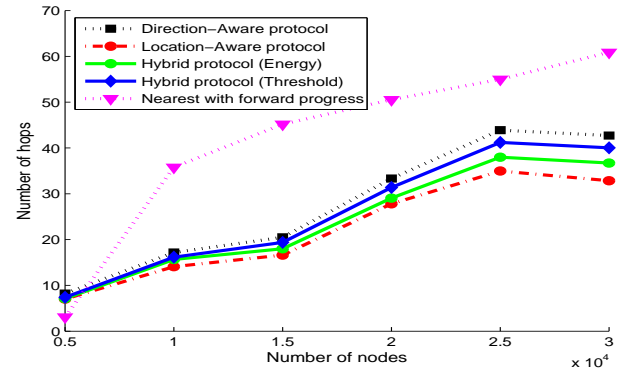


Figure 14: Data latency - 25% Failed Particles.

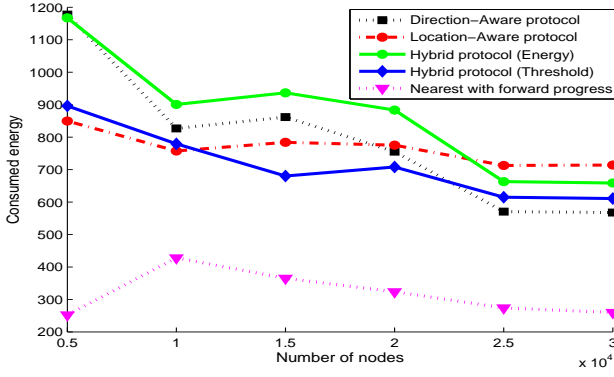


Figure 12: Energy consumption - 10% Failed Particles.

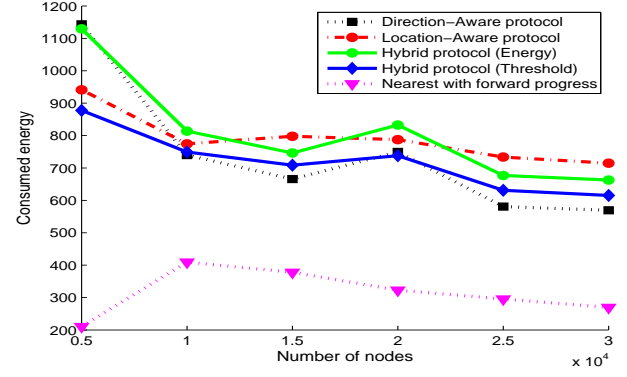


Figure 15: Energy consumption - 25% Failed Particles.

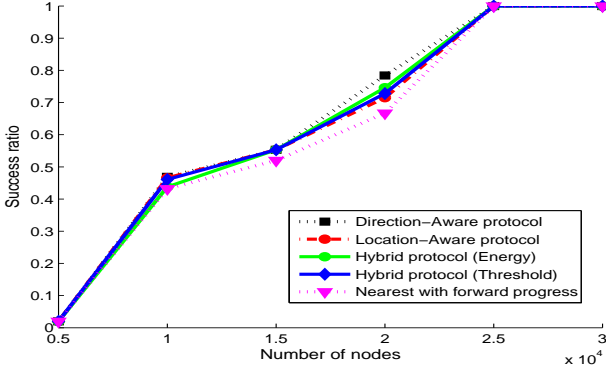


Figure 13: Success ratio - 10% Failed Particles.

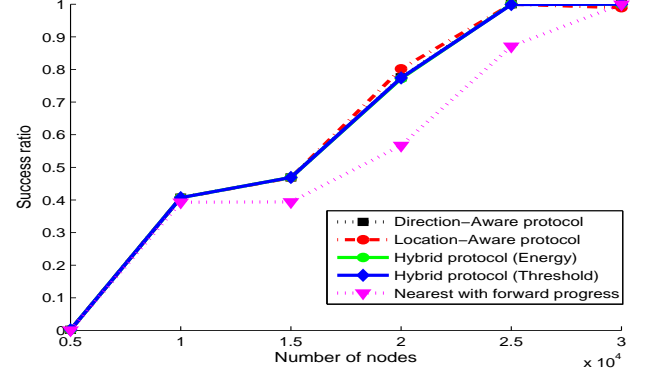


Figure 16: Success ratio - 25% Failed Particles.

after 4.000 data propagations. The initial energy available to the sensor devices was set to 1000 energy units at the start of the simulation. Nodes with high energy dissipation are depicted with dark colours. In contrast, nodes with high residual energy are depicted with bright colours.

We observe in Fig. 20 (left), that our energy-based Hybrid protocol achieves better energy balancing than the LAR protocol Fig. 20 (right) and DAR protocol Fig. 21 (left), however in Fig. 21 (right) it is obvious that the NFP protocol is the most energy conservative. We obtained two interest-

ing observations from the Tables 1 - 4: a) the number of the nodes that have used little or no energy is smaller in the Hybrid (21442 nodes) than the LAR (23733 nodes), the DAR (24728 nodes) and the NFP (27477 nodes) protocol, and b) the number of the nodes that have used most of their initial energy is smaller in the Hybrid (234 nodes) than the LAR (554 nodes) and the DAR protocol (388 nodes) but is larger than the slow NFP protocol (94 nodes).

By using the energy based Hybrid protocol instead of the LAR or the DAR protocol, we can achieve more balanced

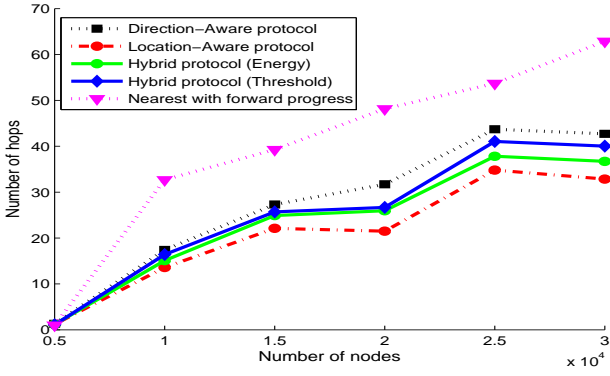


Figure 17: Data latency - 50% Failed Particles.

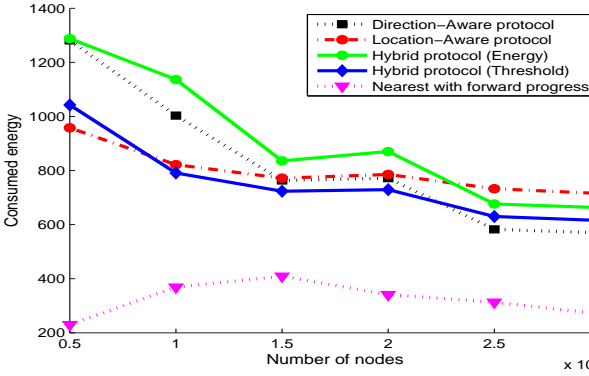


Figure 18: Energy consumption - 50% Failed Particles.

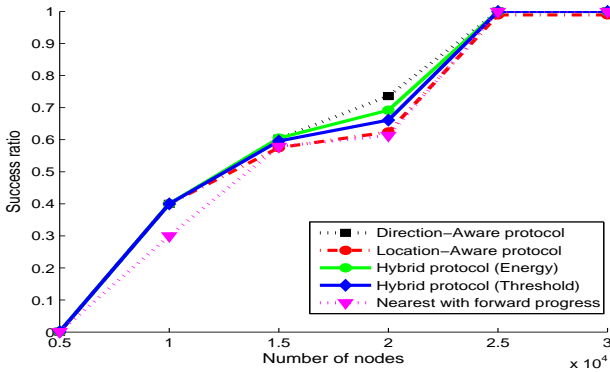


Figure 19: Success ratio - 50% Failed Particles.

energy dissipation per node, as in the operation of the protocol participate more nodes than in the LAR or the DAR protocol, in order to prolong the network lifetime by avoiding early network disconnection. Moreover, LAR and DAR tends to overuse the nodes in the critical region around the sink. Finally, NFP is the most energy conservative protocol, but it is the slowest protocol among all.

## 5. CONCLUSIONS

We study the problem of greedy data propagation, aiming mainly to reduce the energy dissipation of the routing al-

gorithm. Towards parameterized energy-latency trade-offs we provide hybrid combinations of two greedy optimization criteria (direction and proximity to the sink), as any single criterion does not simultaneously satisfy both energy efficiency and low latency.

We rigorously analyzed the direction-aware DAR protocol, the location-aware LAR protocol and the nearest with forward progress NFP protocol. Also, we compared experimentally the above protocols and hybrid combinations of location-aware and direction-aware protocols towards a satisfactory performance trade-off. Although the DAR protocol increases latency, on the other hand it achieves much better results regarding the energy consumption. On the contrary, LAR is the fastest way to transmit data towards the sink, whilst it is the most energy consuming. NFP is very cheap in terms of energy consumption, but it increases data delivery latency dramatically. The Hybrid protocols' performance is in-between the latency values of the DAR and the LAR and allow a fine-tuning of performance. Interestingly, the hybrid protocols beat both LAR and DAR as far as the energy balance is concerned.

## 6. REFERENCES

- [1] I. Chatzigiannakis, S. Nikolettseas, and P. Spirakis. Efficient and robust protocols for local detection and propagation in smart dust networks. *ACM Mobile Networks and Applications (MONET) Journal*, 10(1):133–149, 2005.
- [2] C. Efthymiou, S. Nikolettseas, and J. Rolim. Energy balanced data propagation in wireless sensor networks. *Wireless Networks (WINET)*, 12:691–707, 2006.
- [3] R. Haider and M. Y. Javed. *Energy Efficient Greedy Approach for Sensor Networks*. Published in Chapter 8, *Greedy Algorithms ISBN 978-953-7619-27-5*, 2008.
- [4] T.-C. Hou and V. Li. Transmission range control in multihop packet radio networks. *IEEE Transactions on Communications*, 34:3844, 1986.
- [5] R. Jain, A. Puri, and R. Sengupta. Geographical routing using partial information for wireless ad hoc networks. *Personal Communication*, 2001.
- [6] B. Karp and H. T. Kung. Greedy perimeter stateless routing for wireless networks. *MOBICOM*, 2000.
- [7] E. Kranakis, H. Singh, and J. Urrutia. Compass routing on geometric networks. *Proc. 11th Canadian Conference on Computational Geometry*, 1999.
- [8] P. Leone, S. Nikolettseas, and J. Rolim. Stochastic models and adaptive algorithms for energy balance in sensor networks. *Theory of Computing (TOCS) Journal*, to appear, 2010.
- [9] S. Nikolettseas and J. Rolim. *Theoretical Aspects of Distributed Computing in Sensor Networks*. Springer Verlag, to appear in 2010.
- [10] A. Papadopoulos and J. McCann. Towards the design of an energy-efficient, location-aware routing protocol for mobile, ad-hoc sensor networks. *DEXA*, 2004.
- [11] S. Ross. *Stochastic Processes*. John Wiley and Sons Inc, 1996.
- [12] H. Takagi and L. Kleinrock. Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Transactions on Communications*, 1984.