Energy-aware Geographic Forwarding of Prioritized Messages in Wireless Sensor Networks

Rocío Arroyo-Valles(1), Antonio G. Marqués(2), Jesús Cid-Sueiro(1)
(1)Department of Signal Theory and Communications. Universidad Carlos III de Madrid.
marrval@tsc.uc3m.es, jcid@tsc.uc3m.es
(2)Department of Signal Theory and Communications. Universidad Rey Juan Carlos de Madrid.
antonio.garcia.marques@urjc.es

***Abstract*— Energy is a valuable resource in wireless sensor networks since it constitutes a limiting factor for the network lifetime. In order to make an efficient use of its own energy resources, each node in the network should be aware of the energy resources at other nodes, which can be relevant to the success of their routing decisions. The proposal of this paper is twofold: (i) to design a routing algorithm based on learning patterns using geographic information and (ii) to focus on the cut down in energy consumption. We show that by exploiting local information from the signals detected at each node, sensor nodes can learn to route messages in order to improve the communication performance of the overall network and minimize the need of coordination or signalling protocols among nodes. Moreover, if messages are prioritized by some importance parameter, the overall importance of the successfully transmitted messages can be drastically improved. Experimental results highlight that our algorithm achieves a good performance in terms of successful delivery rate and maximizes the importance of the received messages.**

I. Introduction

In the last years, the application of wireless sensor networks (WSN) has been spread to different scenarios, such us environment monitoring, smart spaces, medical systems or robotic explorations, to name a few. WSN typically consist of hundreds to thousands of self-organized low-cost nodes whose batteries can not be (easily) refilled [1]. Sensor nodes are constrained by resources such as storage capacity and data processing capability [2], but the main drawback of these networks is often due to energy consumption. Although different architectures have been considered for WSN, the potential of distributed cooperation among nodes to perform advanced signal processing tasks with unprecedented robustness and versatility makes decentralized multi-hop WSN one of the most appealing architectures. One of the multiple roles sensors can play in such WSN consists of forwarding information originated by other nodes. As a consequence, sensor failures (thus battery consumption) originate changes in re-routing and network re-organization. Energy conservation and power management are then crucial issues to prolong network lifetime as much as possible while preserving network connectivity and data delivery at the same time.

In order to save energy, considerable research has been focused on the design of power-aware techniques for WSN [1]. This research includes: (i) sleep modes, where nodes turn off their radio, that are often used to reduce the energy consumption (this needs to be done without compromising connectivity so that a path between a source and a destination can be ensured [3]); (ii) power scheduling, used to reduce the energy consumption in the physical layer [4], [5]; (iii) or different energy-efficient algorithms, for network coverage, medium access control protocols, and routing (see e.g., [6], [7], and [8]).

1-4244-1455-5/07/$25.00 ©2007 IEEE

Many routing protocols and algorithms have been proposed in the literature to transmit data efficiently on multi-routes in multi-hop WSN (see e.g., [2],[1]). However, route discovery still remains a challenging issue given the difficulty of designing a routing algorithm which shows good performance under all scenarios and for all applications [2]. As topological changes require updating the node distributions periodically, location techniques are reasonable useful, specially when forwarding data. Several techniques have been proposed on the topic of geographical forwarding and greedy forwarding. Different alternatives for routing protocols are GFG [9], GPSR [10], GEAR [11] [12] and GeRaF [13].

Motivated by the aforementioned challenges, in this paper we propose LPGR (Learning-based Prioritized Geographical Routing), a novel energy-efficient routing algorithm based on learning models, which uses geographical information and considers messages with different importance initially introduced in [14]. The importance can be interpreted as the priority of the transmitted message and is related to parameters such as the quality of the estimation, the relevance of the information, or to quality of service (QoS) requirements (e.g., delay constraints). Applying the benefits from learning patterns to routing, sensor nodes may learn from the success or failure of past routing decisions to make intelligent decisions according to future conditions. Each node observes if neighboring nodes forward their messages and based on these observations, the estimation of available energy in the neighborhood, geographical information, and importance (priority) of the message to be transmitted, improves its routing performance in later chances by doing probabilistic routing [14], [15], [16], [17]. Unlike most of the routing algorithms, LPGR considers a realistic physical layer model where the reception of the message depends on the distance between the sender and the receiver.

Using an algorithm that implements learning mechanisms and probabilistic routing entails different advantages. First, when analytical information relevant to routing is not known (e.g., the statistical distribution of variables or the model describing the dependence among parameters) it can be estimated using "a priori" information from available data [18]. Second, probabilistic estimation of parameters that are distributed among nodes will reduce the amount of information stored, transmitted and updated through the network. Third, even if the network infrastructure let all nodes have full knowledge about the network state, typical operating conditions in WSN make not pragmatically possible to acquire this knowledge in real time (failure in batteries, broken links, nodes change their place). This way, to appropriately address the problem of finding (sub)optimal routing solutions, it is necessary to consider a level of uncertainty in the network state, a complex aspect to take into account in deterministic algorithms, but considered in a natural way in the probabilistic approximation. As a result, we end-up with a lower complexity approach which obtains results even better than usual routing techniques.

The rest of the paper is organized as follows. The decision model as well as our LPGR algorithm is detailed in Section II; experiments are implemented and analyzed in Section III; conclusions, future research and extensions of this work are exposed in Section IV.

II. Decision Model

A. Scenario

Consider a static network of N sensor nodes $\{i|i = 1, \ldots, N\}$. All nodes are assumed to be homogeneous and non hierarchically organized, have similar resources, and behave according to the same rules. These nodes are spread along a geographical area, and can send information packets among themselves. Due to power limitations, each node can only transmit messages to nodes inside its coverage area, which is modelled as a Unit Disk Graph (UDG) (i.e., a node communicates with other node if it is in the transmission radius, which is the same for all nodes). Therefore, network connectivity can be described in terms of the edge set given by

$$\mathcal{E} = \{(u, v) \in \mathcal{V}^2 | u \neq v, d(u, v) \leq R\} \tag{1}$$

where $\mathcal{V}$ is the set of vertices (nodes), R is the communication radius and $d(u, v)$ is the Euclidean distance between nodes u and v.

However, ignoring the impact on the performance of power fluctuations introduced by the channel is not realistic (in fact, this constitutes the main drawback of UDG). Therefore, to represent a more realistic physical layer we consider a *log-normal* shadow fading model (LNS), where the probability for receiving a message successfully (reception probability) is represented as a function of the distance between nodes [19].

Despite the fact that we do not explicitly include collisions in our model, the reception probability captures the uncertainty associated to the fact that a message may not arrive to its next relay in its route to its final destination. Nodes can estimate this reception probability based on the distance between nodes, signal strength or statistics from the packets sent and received between nodes. This means that messages have now a probability to be received without error that decreases with the distance. The log-normal shadowing model takes into account these remarks. As the exact computation is complex and a time consuming process in nodes restricted by energy, the use of an approximated function for simplifying the analysis and the calculations in nodes while simulating is justified, and the expression is equal to [20]:

$$P_{Rx}(d) = \begin{cases} 1 - \frac{1}{2}(\frac{d}{R})^{2\beta} & \text{if } 0 \leq d < R \\ \frac{1}{2}(2 - \frac{d}{R})^{2\beta} & \text{if } R \leq d < 2R \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

being β the power attenuation factor, $2 \leqslant \beta \leqslant 6$. This formula holds for the set-up analyzed in [21] where packets have length 120 bits .

Following the lines in [19], we assume that all nodes consume the same power for each transmission and the packet length is constant for all messages. Wireless networks usually use a single-frequency for the communication purpose in which a message sent to other node is also heard by others belonging to the sender's transmission range [22]. We will denote $\phi(i)$ as the set of all nodes in the coverage area of node i, so that any message sent by node i will be received by all nodes in $\phi(i)$. Reciprocity between coverage areas is assumed (i.e. if $j \in \phi(i)$ then $i \in \phi(j)$), which is a natural assumption if they have a single omnidirectional antenna.

It is widely recognized that the routing performance can be improved if nodes have (possibly local) information about their own geographical position and that of other nodes. In this paper, we assume that node i is aware of its own location (with $\mathbf{z}_i$ denoting its own geographical coordinates), the location of its neighbors in $\phi(i)$ (cf. [23], [24]) and the location of the sink (a.k.a. access point), but global knowledge of the network topology is not needed. Although preliminary works considered that using GPS was impractical due to energy and size restrictions, the recent availability of small, inexpensive, low-power GPS receivers as well as other novel techniques for finding relative coordinates based on signal strength [22], are making feasible to provide sensors nodes with services such as Ad-Hoc positioning Systems (APS) [23] or the GPS-less-low cost outdoor localization for very small devices, as proposed in [24].

All messages must be addressed to a special node called sink. This node is in charge of merging all the received information given that it is connected to an external structure to further processing. Without loosing generality, N will be used as the index of the sink node, so that $\mathbf{z}_N$ is its geographical position. Typically, the sink node is unique, and so we have considered it in the model. Packet time-life is not limited by the number of hops.

Regarding to the MAC layer, we will focus on contention protocols since they show clear advantages with respect to scheduled protocols. Contention protocols can scale more easily across changes in node density or traffic load, can be more flexible as topologies change and do not require fine-grained time synchronization [25]. Nevertheless, the main drawback is its inefficient usage of energy. Several techniques have been proposed to reduce energy consumption in contention-based protocols for sensor networks. The simplest approach is to put the radio into sleep state when it is not needed, turning off the radio interface [7]. Since nodes change their activity between sleep and active modes (idle, transmitting and receiving) in order to economize in energy, the activity status is also required. Our approach here is to model this status in a probabilistic manner. This way we avoid the need of holding neighboring tables and schedules with this sort of information, which in turn will entail energy expense due to storage and exchanges among nodes.

As it is also widely considered that nodes turn off the radio interface periodically to reduce energy consumption, we suppose that nodes maintain their radio interface active for a time interval of average duration T_{ON} and inactive for a time interval of average duration T_{OFF}. Unless nodes synchronize among themselves and coordinate, messages will be lost. Instead of using beacons each time a node wakes up, we have established low duty-cycle operation on nodes in the multi-hop network. According to [25] duty cycles are on the range of 1-10%, and we assume an operation cycle of 10%. Moreover, in order to simplify the analysis, we will assume that nodes are coordinated, neglecting synchronization errors or possible energy costs due to coordination issues.

As contention protocols do, active nodes listen to the channel before transmitting to detect if the channel is busy and thus a node transmits if the medium is idle. Since most sensor networks are designed to operate over long time, nodes will be in idle state for a long time, so that idle state predominates among the others (transmitting and receiving) [25]. The probability of a node being idle, P_{Idle}, can be adjusted based on the specific problem requirements, in our case, since we consider a dense WSN we can set the percentage of time sensors are in idle state to a high value (e.g. $P_{Idle} = 98\%$).

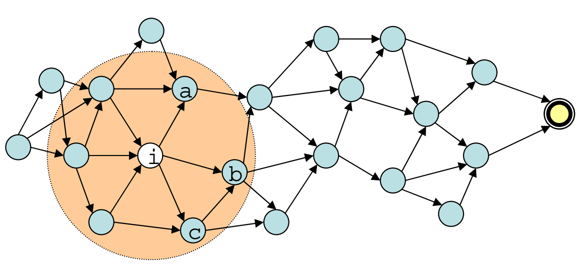

Fig. 1. Network of sensor nodes. The shaded circle shows the coverage area of node i (if no obstacles are considered and sensor antennas are omnidirectional the coverage area of i will be represented by a circle whose center is the node i). All nodes inside the circle (unless i) is the set of neighbors, $\phi(i)$. Each time node i receives a message, it should make a decision about forwarding it to other node or not. To do so, it ranks all its neighbors according to a metric that tries to maximize the expectation of the probability of transmission and the progress towards the sink.

B. Selection of the candidate node

Each time a node generates or receives a message it must make a decision about sending it to other node, or not (see Fig. 1).

The selection of the subset of neighboring nodes that can forward messages towards the sink node is a key issue. Close to the criterion for selecting the candidate node proposed in [20] but to avoid flooding (as it may stem from the seminal model presented in [14]) as well as exhaustion of the batteries due to multiple receptions, the node currently holding the message will select a unique candidate among all its neighbors so that it maximizes the expected forward progress

$$\begin{aligned} &\Pr\{T_j = 1\}(\| \mathbf{z}_i - \mathbf{z}_N \| - \| \mathbf{z}_j - \mathbf{z}_N \|) = \\ &\mathbb{E}\left[\Pr\{T_j = 1|\mathbf{x}\}\right](\| \mathbf{z}_i - \mathbf{z}_N \| - \| \mathbf{z}_j - \mathbf{z}_N \|) \end{aligned} \tag{3}$$

where node i is currently holding the message, node j is the candidate, $\Pr\{T_j = 1|\mathbf{x}\}$ is the probability that node j forwards the message, i.e, the probability that candidate node j transmits the message successfully, $\| \cdot \|$ denotes the Euclidean distance, and $\mathbb{E}$ stands for expectation. This metric ranks neighboring nodes with higher progress towards the sink and higher probability of transmission to the first positions in the prioritized list of candidates while, at the same time, drastically reduces the probability of spreading messages in the opposite direction to the sink.

As in other routing algorithms, each node makes a decision about forwarding the message to the selected candidate. If the node holding the message decides not to transmit it to a given candidate, the following candidate is chosen from the list and a new decision is made. If all candidates reject the re-transmission, the message is discarded. On the other hand, when the node that currently holds the message decides to transmit, it keeps listening to the channel to check whether the candidate node re-transmits the message. If it never hears it or the message is lost during the transmission, the node will select the following candidate from its list.

At this point, it is worth clarifying that LPGR routing algorithm is not a flooding scheme since messages are not sent to all the nodes, but just to the selected candidate from a prioritized list (and only after checking a decision rule). In order to LPGR send a message to each and every neighboring node, it would be required that all transmissions fail and all decisions rule are solved favorably, which is clearly unusual.

Decisions at node i will be based on the following variables:

- An estimation of the available energy at neighboring nodes, $\{\hat{E}_{ij}, j \in \phi(i)\}$.
- The importance of the message to be transmitted, I. Note that the evaluation of the importance of a message is a responsibility of the source node (different importance values are selected when going beyond different thresholds), and it should be transmitted along with the message.
- The distances to its neighboring nodes, $\{d_{ij}, j \in \phi(i)\}$.

- The transmission energy, E_T, which is assumed to be constant for all messages.
- The idle/listening energy, E_I.
- The reception energy, E_R.

C. Energy estimation

At the beginning of this section we highlighted that energy efficiency is one of the most important issues in wireless sensor networks. In addition to collisions, which involve a great waste of energy, idle state and message receptions also consume a non-negligible amount of energy. The idle listening happens when the sensor's radio is listening to the channel to receive possible data. The exact cost of the idle listening depends on radio hardware and the operation mode. It is known that for long-distance radios (0.5 km or more) transmission power dominates receiving and listening costs, however several generations of short-range radios show listening costs of the same order of magnitude for transmission and reception. Detailed measurements of power and energy consumption have been carried out to determine the power/energy drain of devices and some results are included, e.g., in [25].

Based on these measurements, we have considered the power consumption ratios obtained from the Digital 2 Mbps Wireless LAN Module (IEEE 802.11/2Mbps) specification, that shows 1:2:2.5 ratios (idle:receive:send). Note that the cost of sending packets is approximately double that of idling for the same amount of time. A natural reaction is trying to minimize the number of sent packets. However, despite the fact that the cost of sending packets is bigger than the cost of being idle, the application and transport level considerations make the idle time the dominant [26], as it has already been mentioned.

Even though it is possible to include the energy reserve value of a sensor in the periodical 'keep alive' beacons, the worst case has been considered given that nodes consume energy while doing the estimation of the available energy at neighboring nodes. Assuming that energy consumption is caused by transmissions, receptions and listening, the estimation (by node i) of the available energy at neighboring node j is given by

$$\hat{E}_{ij}(k+1) = \hat{E}_{ij}(k) - t(k)E_T - r(k)E_R - \gamma E_I \quad (4)$$

where $\hat{x}$ denotes the estimated value for variable x, $t(k)$ is the number of messages transmitted by node j at time k, E_T is the energy consumed per transmission, $r(k)$ is the number of messages received by node j at time k, E_R is the energy consumed per reception, E_I is the energy consumed due to idle state and γ is a factor that represents how long node j remains listening. Note that our model assumes that energy consumptions are constant for any transmission (which is a reasonable approximation if information is sent in packets of equal size and the transmission power remains unchanged). Node i "listens" to all transmissions carried out by its neighbor, j, so that $t(k)$ can be directly obtained by listening to the channel (remember that that we assume omnidirectional antenna). The remaining variables in (4) are known (by node i) except for the number of receptions, $r(k)$, and constant γ.

The value of $r(k)$ can be approximated by the number of transmissions, since a reception is usually followed by the transmission of the node. Note that the message is addressed to a unique candidate. There are, however, messages received by j that node i is not aware of. Experimental results show that this slight energy underestimation results not significant for the routing decision making since: it seems to affect in a similar way neighboring nodes concentrated in a local region and nodes give up sending periodical beacons at nodes' death so that its neighbors realize that no energy remains at the node. Parameter γ is obtained from heuristics.

D. Decision rules

At node i, energy estimation $\{\hat{E}_{ij}, j \in \phi(i)\}$ and message importance I are grouped into observation vector $\mathbf{x}$. Each node with a message to transmit states the decision as a result of solving a hypothesis testing problem with two hypotheses, $T = 0$ or $T = 1$, where:

- $T = 1$ if at least one neighbor will forward the message.
- $T = 0$ if no neighbor will forward the message in order to save energy for future high important messages.

Depending on its belief about the value of T, node i will make decision $\mathcal{D}_1$ (the message is transmitted) or $\mathcal{D}_0$ (the message is not transmitted).

To do so, we define cost $C(\mathcal{D}_i, z) = c_{iz}$ as the cost of deciding D_i when the true hypothesis is $T = z$ (where $i, z \in \{0, 1\}$) so that:

- $c_{00} = E_I$.
- $c_{10} = E_T$.
- $c_{01} = E_I$.
- $c_{11} = E_T - \alpha I$.

It is worth mentioning that the cost of refusing a transmission is the energy spent while keeping listening. We consider the same cost when a node decides not to transmit a message, even if any of the candidates would have chosen to forward the message or not to do it. The cost for node i when deciding not to transmit is independent of the decisions taken by candidate nodes. The importance of the messages contributes to the reduction of the cost only if the message is forwarded by a neighboring node. As the message importance I is not scaled as regards the energy consumption, parameter α modulates the trade-off between the transmission energy (cost) and the message importance (benefit). However, the cost model proposed is, to some extent, arbitrary.

According to this, the mean cost of deciding $\mathcal{D}_0$ and $\mathcal{D}_1$ are

$$C(\mathcal{D}_0|\mathbf{x}) = E_I, \quad (5)$$

$$C(\mathcal{D}_1|\mathbf{x}) = E_T \Pr\{T = 0|\mathbf{x}\} + (E_T - \alpha I) \Pr\{T = 1|\mathbf{x}\}$$

$$= E_T - \alpha I \Pr\{T = 1|\mathbf{x}\}, \quad (6)$$

respectively. As the objective is to minimize the cost, the final decision is given by D_1 if

$$\Pr\{T=1|\mathbf{x}\} > \frac{E_T - E_I}{\alpha I} \tag{7}$$

and D_0 otherwise.

In order to estimate the posterior probability of each hypothesis, node i makes two simplifying assumptions:

a1) The probability of node j forwarding a message is independent of the forward decision made by any other nodes.

a2) The probability of node j forwarding a message is independent on the state of any other nodes.

Defining the random variable T_j equal to 1 if node j will forward the message and 0 otherwise, we can rely on a1) to write

$$\begin{aligned} y_{ij} &= \Pr\{T=1|\mathbf{x}\} = 1 - \Pr\{T=0|\mathbf{x}\} \\ &= 1 - \prod_{\substack{j=1 \\ j\in\phi(i)}}^{L} (1 - \Pr\{T_j = 1|\mathbf{x}\}). \end{aligned} \tag{8}$$

As we make the decision of transmitting based on a candidate each time, L is set to 1. In order to apply all the previous ideas, let define the random variable O_j that is equal to 1 if node j is *ON* and 0 otherwise; L_j that is equal to 1 if node j is *idle* and 0 otherwise; and R_j as the reception probability. Then, we can write

$$\begin{aligned} &\Pr\{T_j=1|\mathbf{x}\} = \Pr\{O_j=1|\mathbf{x}\} \\ &\quad\times \Pr\{L_j=1|O_j=1,\mathbf{x}\} \\ &\quad\times \Pr\{R_j=1|O_j=1,L_j=1,\mathbf{x}\} \\ &\quad\times \Pr\{T_j=1|O_j=1,L_j=1,R_j=1,\mathbf{x}\}. \end{aligned} \tag{9}$$

Since we assume that nodes are synchronized to be in state *ON* at the same time (as we have already exposed in Section II-A), the probability that node j is *ON* when node i is planning to transmit is equal to 1, so that $\Pr\{O_j|\mathbf{x}\}$. The second factor, $\Pr\{L_j|O_j=1,\mathbf{x}\}$ is the probability that node j is able to receive the message properly due to the fact that it is listening, so we can approximate it as the probability of being idle, $\hat{P}_{Idle}$. The third factor is the reception probability, which can be computed according to (2). Therefore, we can write (9) as

$$\Pr\{T_j=1|\mathbf{x}\} = \hat{P}_{Idle}\hat{P}_{Rx}\Pr\{T_j=1|O_j,L_j,R_j,\mathbf{x}\}, \tag{10}$$

where to simplify the notation, we have avoided the fact that conditioned variables are equal to 1. In this paper we assume a perfect channel knowledge, i.e. $\hat{P}_{Rx} = P_{Rx}$, therefore parameters such as β and R from (2) are known. As a direct consequence of a2), the posterior probability of transmitting only depends on local information of node j, and then we can write

$$\Pr\{T_j=1|O_j,L_j,R_j,\mathbf{x}\} = \Pr\{T_j=1|O_j,L_j,R_j,\mathbf{x}_j\} \tag{11}$$

where $\mathbf{x}_j = (\hat{E}_{ij} \quad I \quad 1)^T$. Notice that the last component (equal to unity) represents a bias constant that has been included for mathematical convenience.

Since a closed-form expression for (11) is unknown, to model it we will assume a truncated logistic model [27]

$$\begin{aligned} \Pr\{T_j &= 1|O_j,L_j,R_j,\mathbf{x}_j\} = \\ &= \frac{1}{1+\exp\left(-\mathbf{w}_j^T\mathbf{x}_j\right)} u(\hat{E}_{ij} - E_T) \end{aligned} \tag{12}$$

where u is the Heaviside step function. A direct consequence of (12) is that node i assigns a zero probability of retransmission to any node that (according to its estimates) does not have energy to transmit the message.

For the remaining factors in (10) (namely, $\hat{P}_{Idle}$ and $\hat{P}_{Rx}$), we will use the same approach to end-up with estimations based on logistics models depending only on local information.

The probabilistic dependencies which define the decision process at each node are illustrated in Fig. 2 for the case of 3 nodes. Each transmitting node "builds" a graphical model including the most relevant variables in the node decision: namely, the importance of the message and the energy of the neighboring nodes. Though each node makes the simplifying assumption that the neighbor decision will not depend on the energy at other nodes, it learns existing probabilistic dependencies through the logistic model.

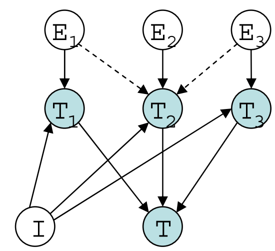

Fig. 2. The graphical model built by a transmitting node including the importance of the message and the energy of the neighboring nodes. Each node makes the simplifying assumption that the neighbor decision will not depend on the energy at other nodes (thus omitting dependencies given by the dashed arrows, that could appear if neighbor are neighbors themselves.

E. Learning

When node i sends an information packet, it keeps 'listening' to the channel. Due to the reciprocity between the coverage areas of neighboring nodes, if an element of $\phi(i)$ forwards the message, node i can detect the retransmission, and use this feedback information to update its profile of the neighboring nodes. Let d_j a binary variable equal to 1 if node j forwards a message received from i (or, more precisely, if node i listens node j forwarding its message) and 0 otherwise. Parameters $\mathbf{w}_j$ are estimated in order to minimize cross entropy loss function [27], a loss function commonly used in neural network training algorithms and which is given by

$$L(y_{ij}, d_j) = -d_j \ln y_{ij} - (1-d_j)\ln(1-y_{ij}) \tag{13}$$

where y_{ij} represents the estimated probability that node i transmits the message through node j. This cost function is widely used for learning problems.

For computational simplicity, we use stochastic gradient learning rules, so that, after transmitting any message, node i updates all parameters as

$$\begin{aligned}\mathbf{w}_j(k+1) &= \mathbf{w}_j(k) + \mu(d_j(k) - y_{ij}(k)) \\ &\times \frac{1 - \frac{y_{ij}(k)}{\hat{P}_{Idle}\hat{P}_{Rx}}}{1 - y_{ij}(k)} u(\hat{E}_{ij} - E_T)\mathbf{x}(k) \qquad (14)\end{aligned}$$

where μ is the updating stepsize and we recall that k represents the discrete time index.

It is important to emphasize that sensor nodes do not need to exchange any specific information among nodes to carry out the learning phase, since they just use the information associated to forwarded transmissions. Proceeding in this way, we enable a fully decentralized routing design which at the same time takes into account non-local information.

III. Experiments and evaluation

In this section, simulated results are presented and analyzed. Performance of our routing protocol LPGR is evaluated and compared to the performance of other geographic forwarding protocol, as the widely well-known GPSR (Greedy Perimeter Stateless Routing) protocol [10]. To fairly compare both algorithms, it is worth noticing that while the GPSR routing algorithm solves the problem of having no neighbors closer to the destination than itself by means of perimeter forwarding on the planarized network graph; in our case, the prioritized list of candidates built from (16) automatically solves that problem by itself. Bare in mind that prioritized list does not exclude nodes that are initially further from the sink than the holder node, but it considers all neighboring nodes and classifies them according to the criterion presented in II-B.

As in [14], nodes are uniformly random placed in a 10×10 area square region and all of them are provided with the same resources except for the sink, which has no restriction on energy reserve since it is the most critical node to collect data in a sensor network. Notice that the choice of a square field is made in order to simplify the experiments. Origin nodes are randomly chosen and the sink is placed at the east border of the field. Our simulations are for networks of 100 nodes. We mainly focus on the role that the importance of the messages plays in routing, what is rarely studied in the literature to our knowledge.

Considering the Log-Normal Shadowing Model and according to [20], two nodes are considered neighbors if the distance between them is at most hR, where $p(R) = 0.5$ and hR is the distance such that $p(hR) = w$. The selected values corresponding to w, h and R are 0.05, 1.4377 and 1.74, respectively, so that the distance to consider two nodes as neighbors is at most 2.5. Note that link losses are included when using this model. According to the set-up in [20], parameter β in (2) has been set β = 2. Parameter α in (7) was fixed to 1.

In order to simulate both routing protocols under the same conditions, we have made GPSR more robust establishing a maximum number of transmission retries before discarding the message, as it was considered in [10] for cases of mobile networks. This value has been fixed to 10 in the simulated networks.

In test *scenario A*, sensors (elected at random) keep transmitting to the sink messages of importances I ($I \in [1, 10]$) until network lifetime expires. Network lifetime is defined as the moment at which the sink is isolated from its neighboring nodes, i.e., there is no route available to reach the destination node. For the same set-up, *scenario B* considers high important traffic (I with values between 5 and 10). Finally, *scenario C* shows permanent failures in a local region of the network topology. This region is geographically at the same location, at the right extreme of the sensor network fields, and only a 10% of the nodes that are inside a region of radius $radius = 3$ from the center of the failure can be affected by the breakdown.

Parameters such as μ (the adaptive step of the weight vector in (14)) and the ratio between the initial and the consumed energy to perform a transmission are reasonably chosen based on preliminary simulations. Presented results are averaged over 50 different topologies.

Test case 1: Performance is assessed in terms of the metrics shown in Table I. This table which corresponds to *Scenario A* evidences that LPGR achieves a higher successful delivery rate than the GPSR algorithm: 81.55% against 46.73%. Note that our algorithm allows to transmit to the sink around a 40% more than GPSR does (267 against 155 on average). The high rate of link losses contributes to obtain this value. Since the power loss dictated by the physical layer depends on the distance, nodes that are close to neighborhood's boundaries have lower probability of successful reception (see graphic in [20]). GPSR forwards the message to the neighbor geographically closest to the sink (greedy forwarding) when it is possible, and so the failure rate increases, even if GPSR possesses retransmission capabilities. Only messages originated in nodes close to the sink are capable of delivering to it. Thus, the average number of hops to reach the destination is lower in GPSR than in our model (3.33 against 7.68). In LPGR, although a node may decide not to transmit to the first candidates in the prioritized list due to the low probability of reception, it makes a decision on another node with an optimal trade-off between progress to the sink and probability of transmission, and so the messages need more hops to reach its final destination but in most of cases, they arrive successfully. As far as transmitted messages are concerned, both algorithms perform similarly. Furthermore, LPGR not only achieves better results in terms of number of received messages, but also in terms of the sum of the importance of the received messages.

It is also remarkable that in order to transmit messages of higher importance, LPGR is able to save energy and so contribute to maximize the sum of the importance of received messages. So that there are some non-sent messages by source nodes, whose importance average depends on the α parameter in the decision rule. In summary, the results for the first test case show that LPGR loses less messages while maximiz their importance.

Test case 2: Simulation results from *scenario B* listed in Tab II show that for LPGR algorithm, the higher the importanc of the message is, the lower the possibility of discard message in nodes This fact contributes to increase the numb of received messages and thus, successful delivery rate. No that the number of hops to reach the sink is slightly low compared to the general scenario, and also the increase of t average importance of the received messages and thus the su of the received importance.

Test case 3: Results corresponding to *scenario C* are listed Table III. From the inspection of the table, we can conclu that when the focus of the failure is close to the sink, t delivery rate decreases in both algorithms. And the same effe is appreciated in the other metrics shown in Table III, suc as the number of transmitted and received messages and the sum of the importances corresponding to those messages that arrived to the sink successfully. A consequence of the failure of some nodes is the increase of the number of hops and the number of retransmissions. To provide understanding of how LPGR behaves when failures happen, Fig. 3 illustrates how a message generated at given node (node 45) reaches successfully the sink (node 4), finding a transmission path around the void which is produced when some nodes failed, according to *Scenario C* in LPGR.

Test case 4: In this case we focus on the method that LPGR algorithm implements to select the neighboring candidates and compare it with other alternatives. To do so, we keep the construction of a prioritized list but we change the selecting rule. Two of the routing algorithms presented in [28] are considered as alternatives: nEPR (non-acknowledge Expected Progress Routing) and PP (Projection Progress). nEPR ranks neighbors according to a metric that maximizes the expected progress, which is the product of the probability of successful delivery between the holder node and the candidate and the progress made to the sink by forwarding to that candidate

$$\hat{P}_{Rx}(\parallel \mathbf{z}_i - \mathbf{z}_N \parallel - \parallel \mathbf{z}_j - \mathbf{z}_N \parallel). \qquad (15)$$

On the other hand, PP ranks candidate nodes according to a metric that maximizes the product

$$\hat{P}_{Rx}(\parallel \mathbf{z}_i - \mathbf{z}_N \parallel \cdot \parallel \mathbf{z}_i - \mathbf{z}_j \parallel) \qquad (16)$$

where $\cdot$ stands for the dot product of two vectors.

From the analysis of the results showed in Table IV, LPGR and LPGR-nEPR roughly have similar outcomes for all the metrics considered (slightly better for the nEPR version) while both of them clearly outperform LPGR-PP. The small gap between LPGR and LPGR-nEPR may be due to the fact that the LPGR-nEPR algorithm has complete information about the channel, useful when computing the probability of reception compared with LPGR that relies on past routing decisions. More importantly, numerical results in all the cases emphasize the utility of I when making routing decisions as well as the effectiveness of prioritized lists.

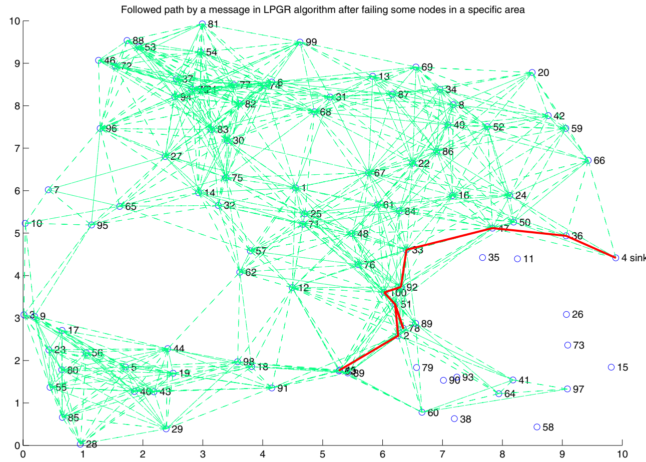

Fig. 3. Illustration showing the path followed by a message generated at node 45 until it is delivered to the sink (node 4) according to *Scenario C* in LPGR. Nodes without a line to their neighbors have used up their batteries or have suffered from failures.

IV. Conclusions and Future Work

In this paper, we have presented a geographical routing protocol for wireless sensor networks. The LPGR algorithm determines to which node from a prioritized list of candidates built from the progress to the sink and the probability of transmission retransmit a generated message. Routing decisions made by nodes that successfully receive the messages are based on Bayesian decision rules, which learn from previous experiences. Thus, by extracting information from the received signals, nodes can estimate the available energy at neighbors and carry out intelligent forwarding decisions. As a consequence, high importance messages are prioritized in forwarding and the average level of importance is higher than the one achieved by other routing algorithms that do not bare it in mind. Experimental results show that LPGR takes advance of the importance of the message, ensures successful transmissions of the messages entailing a low loss rate, and maximizes the importance of the received messages. For the physical layer model implemented, this is accomplished with an acceptable trade off in the number of hops to the sink.

Future work includes to explicitly consider collisions in the network; a more refined definition of the costs, especially a better adjustment between the transmission energy and the importance of the messages (adjustment of α parameter in Eq. (7)); to consider the energy of transmission as a variable and not as a constant (since it can be adjustable depending on the distance between transmitting-receiving nodes and the effect of spreading new sensor nodes after a first deployment); or to merge the GPSR algorithm with the idea of prioritized messages, in order to increase the importance average of the received messages. Comparison with other routing algorithms that address realistic physical layer would be more suitable and are also suggested as future work to be done.

TABLE I

PERFORMANCE EVALUATION AND COMPARISON BETWEEN LPGR AND GPSR ALGORITHM, SCENARIO A

	Successful delivery rate % (mean value ± std)	Transmitted messages	Received messages	Received avg. importance	Sum of received importance	Avg. Hops to the sink	Avg. Number of Retx
LPGR	81.55 ± 8.24	322 ± 87	267 ± 89	6.52 ± 0.19	1730.96 ± 562.02	7.68 ± 1.38	10.53 ± 2.12
GPSR	46.73 ± 5.48	330 ± 31	155 ± 29	5.53 ± 0.22	858.26 ± 170.66	3.33 ± 1.72	9.82 ± 2.17

TABLE II

PERFORMANCE EVALUATION AND COMPARISON BETWEEN LPGR AND GPSR ALGORITHM, SCENARIO B

	Successful delivery rate % (mean value ± std)	Transmitted messages	Received messages	Received avg. importance	Sum of received importance	Avg. Hops to the sink	Avg. Number of Retx
LPGR	90.96 ± 7.91	323 ± 79	298 ± 86	7.51 ± 0.14	2234.36 ± 641.95	7.41 ± 1.49	10.42 ± 2.35
GPSR	46.73 ± 5.48	330 ± 31	155 ± 29	7.50 ± 0.15	1162.04 ± 217.686	3.33 ± 1.72	9.82 ± 2.17

TABLE III

PERFORMANCE EVALUATION AND COMPARISON BETWEEN LPGR AND GPSR ALGORITHM, SCENARIO C

	Successful delivery rate % (mean value ± std)	Transmitted messages	Received messages	Received avg. importance	Sum of received importance	Avg. Hops to the sink	Avg. Number of Retx
LPGR	76.08 ± 8	274 ± 70	212 ± 71	6.56 ± 0.19	1382.3 ± 444.71	7.81 ± 1.25	10.70 ± 1.88
GPSR	46.55 ± 5.45	323 ± 32	152 ± 29	5.49 ± 0.23	832.52 ± 164.24	3.59 ± 1.78	10.12 ± 2.22

TABLE IV

PERFORMANCE EVALUATION AND COMPARISON AMONG LPGR, LPGR-NEPR AND LPGR-PP ALGORITHMS, SCENARIO A

	Successful delivery rate % (mean value ± std)	Transmitted messages	Received messages	Received avg. importance	Sum of received importance	Avg. Hops to the sink	Avg. Number of Retx
LPGR	81.55 ± 8.24	322 ± 87	267 ± 89	6.52 ± 0.19	1730.96 ± 562.02	7.68 ± 1.38	10.53 ± 2.12
LPGR-nEPR	83.16 ± 9.12	342 ± 78	289 ± 85	6.52 ± 0.17	1878.54 ± 530.50	7.43 ± 1.23	10.10 ± 1.88
LPGR-PP	67.70 ± 16.44	178 ± 44	124 ± 51	6.73 ± 0.30	824.18 ± 322.55	13.53 ± 2.80	17.54 ± 3.62

ACKNOWLEDGMENT

This paper has been partially supported by Spanish MEC grant TEC2005-06766-C03-01/TCM, and -2/TCM, CAM-UC3M project S-MORSE CCG06-UC3M/TIC-0805, and CAM project P-TIC-000223-0505.

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[2] Q. Jiang and D. Manivannan, “Routing protocols for sensor networks,” 2004.

[3] M. Zorzi, “A new contention-based mac protocol for geographic forwarding in ad hoc and sensor networks,” *Proceedings of the IEEE International Conference on Communications*, vol. 6, no. 3481- 3485, 2004.

[4] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, “Power scheduling of universal decentralized estimation in sensor networks,” *IEEE Trans. on Signal Processing*, vol. 54, no. 2, p. 413422, 2006.

[5] A. G. Marques, X. Wang, and G. B. Giannakis, “Minimizing transmit-power for coherent communications in wireless sensor networks using quantized channel state information,” *IEEE Proc. of Intl. Conf. on Acoustics, Speech and Signal Processing*, 2007.

[6] M. Bhardwaj, T. Garnett, and A. P. Chandrakasan, “Upper bounds on the lifetime of sensor networks,” *Proc. of Intl. Conf. on Commun.*, vol. 3, pp. 785–790, 2001.

[7] W. Ye, J. Heidemann, and D. Estrin, “An energy-efficient mac protocol for wireless sensor networks,” *Proc. of 21st Annual Joint Conf. IEEE Computer Commun. Societies*, vol. 3, pp. 1567–1576, 2002.

[8] J. N. Al-Karaki and A. E. Kamal, “Routing techniques in wireless sensor networks: A survey,” *IEEE Wireless Commun. Mag.*, vol. 11, no. 6, pp. 6–28, 2004.

[9] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, “Routing with guaranteed delivery in ad hoc wireless networks,” *Proc. of 3rd ACM Int. Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications DIAL M99*, pp. 48–55, 2001.

[10] B. Karp and H. Kung, “Gpsr: Greedy perimeter stateless routing for wireless networks,” *Proceedings of ACM Mobicom*, 2000.

[11] I. Stojmenovic and X. Lin, “Power aware localized routing in wireless networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 12, no. 11, pp. 1122–1133, 2001.

[12] Y. Yu, D. Estrin, and R. Govindan, “Geographical and energy-aware

routing: A recursive data dissemination protocol for wireless sensor networks," *UCLA Computer Science Department Technical Report*, 2001.

[13] M. Zorzi and R. Rao, "Geographic random forwarding (geraf) for ad-hoc and sensor networks: Multihop performance," *IEEE Transactions on Mobile Computing*, vol. 2, no. 4, 2003.

[14] R. Arroyo-Valles, A. G. Marques, J. J. Vinagre-Diaz, and J. Cid-Sueiro, "A bayesian decision model for intelligent routing in sensor networks," *Proceedings of the 3rd International Symposium on Wireless Communication Systems (ISWCS06)*, 2006.

[15] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE Mobile Computing and Communications Review, Springer*, 2003.

[16] A. A. Papadopoulos, J. A. McCann, and A. Navarra, "Connectionless probabilistic (cop) routing: an efficient protocol for mobile wireless ad-hoc sensor networks," in *Proceedings of the 24th IEEE International Conference on Performance, Computing, and Communications,(IPCCC 2005)*, 2005.

[17] F. Stann and J. Heidemann, "Bard: Bayesian-assisted resource discovery in sensor networks," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communication Societies (INFOCOM 2005)*, 2005.

[18] E. Castillo, J. Gutierrez, and A. Hadi, *Expert systems and probabilistic network models*. Springer-Verlag, 1997.

[19] I. Stojmenovic, A. Nayak, and J. Kuruvila, "Design guideliness for routing protocols in ad hoc and sensor networks with a realistic physical layer," *IEEE Communications Magazine*, pp. 101–106, 2005.

[20] J. Kuruvila, A. Nayak, and I. Stojmenovic, "Hop count optimal position based packet routing algorithms for ad hoc wireless networks with a realistic physical layer," *First IEEE International Conference on mobile Ad Hoc and Sensor Systems*, pp. 398–405, 2004.

[21] M. Stojmenovic and A. Nayak, "Localized routing with guaranteed delivery and a realistic physical layer in wireless sensor networks," *Computer Communications (Elsevier)*, vol. 29, no. 13-14, pp. 2550–2555, 2006.

[22] I. Stojmenovic, "Position-based routing in ad hoc networks," *IEEE Communications Magazine*, pp. 128–134, 2002.

[23] D. Niculescu and B. Nath, "Ad hoc positioning system (aps)," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, 2001.

[24] N. Bulusu, J. Heidemann, and D. Estrin, "Gps-less-low-cost outdoor localization for very small devices," *IEEE Personal Communications (see also IEEE Wireless Communications)*, vol. 7, no. 5, pp. 28–34, 2000.

[25] W. Ye and J. Heidemann, "Medium access control in wireless sensor networks," *Technical Report*, 2003.

[26] M. Stemm and R. H. Katz, "Measuring and reducing energy consumption of network interfaces in hand-held devices," *IEICE Transactions on Communications*, vol. E80-B, no. 8, pp. 1125–1131, 1997.

[27] J. W. Miller, R. Goodman, and P. Smyth, "On loss functions which minimii to conditional expected values and posterior probabilities," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1404–1408, 1993.

[28] I. Stojmenovic, A. Nayak, J. Kuruvila, F. Ovalle-Martinez, and E. Villanueva-Pena, "Physical layer impact on the design and performance of routing and broadcasting protocols in ad hoc and sensor networks," *Computer Communications (Elsevier)*, vol. 28, no. 10, pp. 1138–1151, 2005.