UNIVERSITE PARIS DAUPHINE

Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision

CNRS UMR 7024

CAHIER DU LAMSADE

273

Avril 2008

A message-optimal sink mobility model for Wireless Sensor Networks

Aristotelis Giannakos, Gregory Karagiorgos, Ioannis Stavrakakis

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

A message-optimal sink mobility model for Wireless Sensor Networks

Aristotelis Giannakos
LAMSADE, CNRS UMR 7024, Université de Paris-Dauphine, Paris, France
aristotelis.giannakos@lamsade.dauphine.fr,
Gregory Karagiorgos
and
Ioannis Stavrakakis
University of Athens, Department of Informatics and Telecommunications,
Panepistimiopolis 157 84, Athens, Greece,
{greg, ioannis}@di.uoa.gr

Abstract. The paper proposes an algorithm for collecting data from a wireless sensor network modeled as a random geometric graph in the unit square. The sensors are supposed to work in an asychronous communication mode (they store the measures they perform and transmit a message containing the data, upon receival of a trigger signal). The model assume a mobile sink passing near the sensors, and asking the sensors to transmit their data. The algorithm defines a route for the sink such that the number of messages that a sensor needs to transmit be as low as possible.
We also present a kind of "controled" random walk in a (connected) random geometric graph that is based upon the main idea of the sink routing algorithm, reducing the graph cover time to $\Theta(n \log \log n)$ instead of $\Theta(n \log n)$ needed when a simple random walk is utilized
The model can be generalized for the case of more than one mobile sinks, and the algorithm can be modified to deal with locally uniform density of sensors deployed in the field.

1 Introduction

The most common usage of wireless sensor networks is that of having them operate in hostile and unattended environments where there are no other means of surveillance or monitoring. One or more sensors (nodes) of a wireless sensor network are chosen to behave as sinks, that is to collect information transmitted by other sensors. However, due to limited transmission capabilities, sensors have to seek cooperation of neighboring nodes for relaying their gathered data to the sink. A common approach for the aforementioned situation is that of multi-hop routing.

An obvious drawback of this approach is that it leads to dissimilar rates of battery power depletion of nodes in the wireless sensor network. To be more specific, sensors around the sink are subject to rapid depletion of their battery power since they have to relay the data from hundreds of thousands of nodes all

across the network. On the other hand, nodes along the periphery of the network enjoy a low load of work since they are rarely asked to forward data to the sink. The power-saving property of peripheral nodes is due to their locality away from the chosen sinks.

To address this drawback of multi-hop routing, sink mobility has often been employed as an effective mechanism. Since, with mobility, the sink moves to new locations, either being time driven or event driven, different nodes come into the sinks vicinity, forcing them to forward messages.

The paper is organized as follows: section 2 contains the description of the model and its relation to prior work. In Section 3, the algorithm is described; the analysis of its performance in terms of time and number of exchanged messages, is presented in Section 4. A related result concerning the definition of a "controled" random walk on random geometric graphs that reduces the needed cover time is contained in subsection 4.1. Finally, concluding remarks and some issues for further research are presented in Section 5.

2 Description of our model and related work

Following numerous previous works ([1], [2], [4]), we model the wireless sensor network we will study, as a random geometric graph [5]. In such a model, denoted $G(n, r)$, the n sensor locations are chosen uniformly and independently at random in the unit square. Each pair of sensors is connected if their Euclidean distance is smaller than some fixed transmission radius r. As noted earlier ([2], [5]) to assure both good connectivity and minimal interference in such a graph model, the transmission radius $r(n)$ must scale like $\Theta(\sqrt{\frac{\log n}{n}})$. For the rest of our study we will assume that all communications within the aforementioned radius always succeed, thus we will not address issues caused by failures due to communication.

Having the above setting of sensors in the unit square, we assume a single node, the *sink* which moves inside the unit square and collects information transmitted by the sensors. Exploiting the mobility of some nodes in order to facilitate data delivery has been widely discussed for general ad hoc networks in different contexts (see for example [6]).

In this paper, we consider the following model of transmission and receival of information from the sensors to the sink: whenever the sink enters a state of collecting data it transmits a specific signal which upon receival from the sensors in transmission range triggers them to transmit a message including the information. At any other time, each sensor senses the surrounding environment and quantifies its observation or stays idle if such an observation is not to be performed. This implies an asychronous communication mode, as in the approach developped by Shah et al. in their works on *data mules* [3]; the idea there, is to save energy by having single-hop communication (from a sensor to the mule that is passing by) instead of the more expensive multi-hop routing (from the sensor to the sink), while the mule is supposed to take the sensed data to the sink.

In fact, what we consider here is a data mule identified to a mobile sink which is is less energy-constrained than sensor nodes.

Our main purpose is to define a specific *route* for the sink, in order to collect *all* the data from the sensors and to achieve that by limiting the total number of transmitted messages to the absolute necessary, thus minimizing total energy consumption. As stated above, the sink possesses enough energy to collect and store all information as well as transmit all relevant communication messages. Note that this is a real world situation in the sense that wireless sensors are of limited power and capacity. Under such a model, the crucial need is the transmission and receival of all measurements under minimum power consumption for each sensor. In our model, minimum power consumption is roughly equivalent to minimum (absolutely necessary) message exchange.

We assume that apart from transmission and receival capabilities the sink is also aware of the number n of sensors in the specified geographical area of concern. For both the sink and the sensor nodes we assume and utilize the fact that they typically know their locations ([4]). Moreover, we only require that apart from standard measurement operations, the sensor nodes can only perform simple mathematical calculations.

3 The proposed algorithm

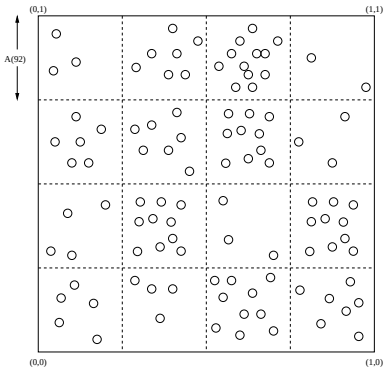


Fig. 1. Grid partition of the unit square

The proposed algorithm for the route of the sink, relies on the following crucial observation:

Lemma 1. *([4]) Let a graph be drawn randomly from the geometric ensemble* $G(n, r)$ *of sensor locations drawn uniformly and independently in the unit area*

(unit square), and a partition be made of the unit area into squares of length $A(n) = \sqrt{2\frac{\log n}{n}}$. *Then the following statements all hold with high probability*[1]*:*

1. *Each square contains at least one node.*
2. *If* $r(n) = \sqrt{10\frac{\log n}{n}}$ *then each node will be able to communicate (an edge exists) to a node in the four adjacent squares.*
3. *All the nodes in each square are connected with each other.*

The above result essentially states that our sensor network can be partitioned in a grid-like manner such as the one depicted in figure 1. Using this partitioning we are guaranteed to have (with high probability) at least one sensor in each square. Furthermore, in the presence of more than one sensors in each square, it is guaranteed that the subgraph induced by these sensors is very dense (with high probability: a clique), meaning that every sensor inside the square can directly communicate with every other sensor also inside the square. Furthermore, note that as the above lemma implies, starting from any sensor inside a square of the grid there is, with high probability, a path leading to a sensor in any neighboring square. On the other hand, the graph that is obtained if we take one vertex for every square and an edge for every couple of squares containing sensors within transmission range, contains a grid as an induced subgraph (in fact, with high probability it is a grid itself). But grids contain hamiltonian paths; hence, it is possible with high probability that there is a route traversing all squares once.

To fix the ideas, let's consider such a hamiltonian path starting from the upper left corner of the grid (that corresponds to the upper left corner square in the grid of squares) and ending to the lower right corner (ie, the lower right corner square in the grid of squares).

The algorithm below, devises such a route with these extremities, corresponding to a hamiltonian snake-like path on the grid.

Suppose that after a specified time needed for the network to stabilize (that is for all the sensors to quantify their observed measurements) we deploy the sink in the upper left square of the grid at the point $(0, 1)$ (the relative coordinates of the square are depicted in figure 1) and the following steps take place:

1. The sink s calculates the current square i that it is located at and consequently chooses as the *target* transition point t the center (x_{c_i}, y_{c_i}) of the square. A greedy geographic routing mechanism routes s to the nearest sensor in the target point t located at (x_t, y_t).
2. s forms the quadruple (x_l, y_l, x_u, y_u), sends a trigger message containing this quadruple and awaits sensors to send their measurements.
3. Each sensor in range, upon receival of the trigger messages transmits its computed measurement if its coordinates (x, y) satisfy $x_l \leq x \leq x_u$ and $y_l \leq y \leq y_u$ and if the sensor has not transmitted its measurements yet.

[1] Event $\mathcal{E}_n$ occurs with high probability if probability $P(\mathcal{E}_n)$ is such that $\lim_{n\to\infty} P(\mathcal{E}_n) = 1$

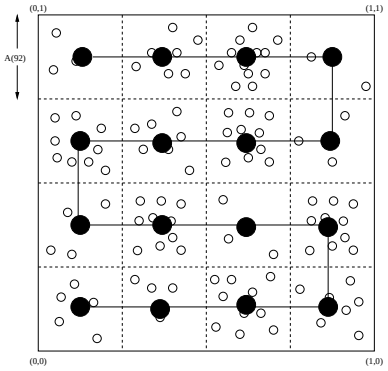


Fig. 2. Snake-like route of the sink.

4. Upon acceptance of the measurements, the sink calculates the center $(x_{c_{i+1}}, y_{c_{i+1}})$ of the neighboring square and after being transfered in that square (via geographic routing like Step 1) repeats the steps 2 and 3. The route of the sink is depicted in figure 2.

From now on we will refer to this procedure as a *single snake algorithm*. We begin by elaborating on the feasibility of this algorithm and in the next section we will provide analysis as far as the optimality of the number of exchanged messages is concerned. Since the sink (as well as any other sensor node) is always aware of the coordinates of its current location, the first step of the algorithm is rather easy and involves simple geometric calculations on behalf of s so as the number of the square and the coordinates of the square's center can be identified. In fact, the only extra information needed is the total number of nodes n in the unit square, information made available to the sink before its deployment in the area.

Proceeding to step 2 the sink is able to identify the lower and upper bounds (x_l, y_l, x_u, y_u) on the coordinates. More specifically, since the side of each square is $A(n) = \sqrt{2\frac{\log n}{n}}$ the bounds are

$$x_l = x_{c_i} - 1/2A(n), \quad x_u = x_{c_i} + 1/2A(n)$$

$$y_l = y_{c_i} - 1/2A(n), \quad y_u = y_{c_i} + 1/2A(n)$$

and from that point all sensors having this information in step 3, decide whether or not they should transmit. To resolve the case under which a sensor is located on the border line of a square of the grid decomposition we require each sensor to transmit not only if it falls inside the above bounds but also if it has not transmitted anything yet.

Finally the feasibility of step 4 is justified via the observation noted earlier that there always exists a path from one square to a neighboring one, an observation justified by lemma 1.

4 Analysis of messages and time

We proceed to the analysis of the single snake algorithm. We begin by elaborating on the total number of messages exchanged throughout a single traversal of the unit square by the sink s.

Lemma 2. *The total number of messages exchanged up to completion of the sink's route, is* $\frac{n}{2\log n} + n$ *the least possible. Furthermore, each sensor transmits only once.*

Proof. The fact that each sensor transmits only once is straightforward via step 3 of the single snake algorithm. Thus, this procedure involves n messages. The trigger messages sent by the sink are as many as the squares of the grid decomposition, that is $1/A^2(n) = n/(2\log n)$. □

As for the time needed for the sink s to complete the route, traversing each square of the grid decomposition the following result holds:

Lemma 3. *The single snake algorithm needs* $O(\frac{n^{3/2}}{\log n^{3/2}})$ *steps to complete.*

Proof. The sink s will traverse a total of $(n/2\log n)$ squares. To proceed from a square to a neighboring one a greedy geographic routing procedure needs to be applied which routes the sink to the target point in $O(\sqrt{n/\log n})$ steps([4]). This totals to $O(\frac{n^{3/2}}{\log n^{3/2}})$ steps. □

Note that as stated in the previous section, under the assumption of $r(n) = \sqrt{10\frac{\log n}{n}}$ there is an *edge* between two sensor nodes of two adjacent squares. One might think that since this edge exists, greedy geographic routing is not necessary for the sink to be relocated to an adjacent square. Although this is true, it is not hard to see that absence of routing would result in a significant increase of the number of messages exchanged for the sink to decide which such *edge* to follow. Furthermore, presence of such a routing mechanism pleads for as much simplicity as possible to the proposed algorithm.

4.1 Fast cover time of the underlying random geometric graph

The good performance of the single snake algorithm is related to the possibility of designing a suitably "forced" random walk that covers the underlying random geometric graph more fast than the simple random walk[2].

Recall the following theorem from [8]:

Theorem 1. *[8] For* $c > 1$*, if* $r^2 \geq \frac{c8\log n}{n}$*, then with high probability* $\mathcal{G}(n,r)$ *has cover time* $\Theta(n\log n)$*. If* $r^2 \leq \frac{\log n}{\pi n}$*, then* $\mathcal{G}(n,r)$ *has infinite cover time with positive probability (bounded away from zero).*

[2] The cover time of a graph G is the expected time taken by a random walk on G to visit all vertices in G.

Now consider the grid partition of the whole $\mathcal{G}(n,r)$ into squares of side length $\sqrt{2\frac{\log n}{n}}$, eachone containing an average of $2\log n$ nodes. It is possible, using state space decomposition techniques (see for example [9]) to devise a "controled" random walk, that starting from the upper left square, is kept into the square until the nodes of the square get covered, then passes (using greedy geographic routing) to the neighbor square, and so on, covering one square after another according to the route between squares used by the single snake algorithm described before.

Then, the following proposition holds:

Proposition 1. *There is a controled random walk for which $\mathcal{G}(n,r)$ has cover time $\Theta(n\log\log n)$.*

Proof (Sketch) The subgraph inside a square satisfies the conditions of theorem 1; hence it takes an average of $\Theta(\log n\log\log n)$ time for the controled random walk to cover it. The total cover time is then $\Theta(\sum_{i=1}^{k} c_i)$, where $k = n/2\log n$ and c_i is $\Theta(\log n\log\log n)$; the result is straightforward.

5 Conclusions and further research

This paper adresses the issue of data collection from a wireless sensor network using a mobile sink, in order to delay as long as possible depletion of battery power of the sensors. The network is modeled as a random geometric graph, with sensor locations uniformly drawn from the unit square, and density depending on the sensors transmission range in such a way that the correponding graph be connected with high probability.

Based on the property that, for a suitable grid partition of the unit area where the network is deployed, the subgraphs within the squares of the grid are with high probability fully connected, an algorithm is proposed for routing the sink from one square to another; each time the sink passes from a square of the grid, it transmit a specific signal to the sensors that are triggered to respond by sending their data to the sink. This model supposes an asynchronous communication mode and guarantees that only an absolutely necessary number of messages are transmitted; hence it guarantees a slow rate of battery power depletion for each sensor.

As a byproduct of the analysis of the algorithm, we also present the idea of a "controled" random walk in a (connected) random geometric graph which reduce the graph cover time to $\Theta(n\log\log n)$ instead of $\Theta(n\log n)$ needed when a simple random walk is utilized.

The single snake algorithm presented can be further generalized to treat the case of data collecting from wireless neworks of sensors deployed in a locally uniform manner.

It would be also of interest to extend this model for multiple sinks, in order to deal efficiently with energy constraints posed to the sinks.

References

1. S. Boyd, A. Ghosh, B. Prabhakar and D. Shah, "Gossip algorithms: Design, analysis and applications", *Proceedings of the 24th Conference of the IEEE Communications Society (INFOCOM 2005)*, pages 1653-1664.
2. P. Gupta and P. R. Kumar, *The capacity of wireless networks*, IEEE Transactions on Information Theory, vol. 46, number 2,2000, pages 388-404.
3. R.C. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling a three-tier architecture for sparse sensor networks", *Proceedings of the first IEEE international workshop for sparse sensor networks ans applications, SNPA 2003*, Anchorage, AK, 2003, pages 30-41.
4. A. G. Dimakis, A. D. Sarwate and M. J. Wainwright, "Geographic gossip: efficient aggregation for sensor networks", *Proceedings of the Fifth International Conference on Information Processing in Sensor Networks (IPSN 2006), 2006*, pages 69-76.
5. M. Penrose, *Random geometric graphs*, Oxford University Press, 2003.
6. I. Chatzigiannakis, S. Nikoletseas, and P. Spirakis, "An efficient communication strategy for ad-hoc mobile networks", *Proceedings of the Twentieth Annual ACM Symposium on Principles of Distributed Computing, PODC 2001*, Newport, RI, 2001, pages 320-332.
7. Z. Maria Wang, S. Basagni, E. Melachrinoudis and C. Petrioli, "Exploiting sink mobility for maximizing sensor networks lifetime", *Proceedings of the 38th Hawaii International Conference on System Sciences, HICSS-38 2005*, Big Island, HI, 2005.
8. C. Avin and G. Ercal, "On the cover time and mixing time of random geometric graphs", *Theoretical Computer Science Vol. 380 , Issue 1-2* (June 2007), pages 2-22.
9. N. Madras and D. Randall, "Markov Chain decomposition for convergence rate analysis", *Annals of Applied Probability* (to appear).