

Random Walks on Sensor Networks

Luísa Lima

João Barros

Instituto de Telecomunicações
Department of Computer Science
Universidade do Porto, Portugal
{barros, luisalima}@dcc.fc.up.pt

Abstract— We consider the mobile data gathering problem in large-scale wireless sensor networks with static sensor nodes and a mobile patrol node. Based on the assumptions that (a) the sensor positions are unknown and (b) the network may not be entirely connected, we formulate the problem as one of random walks in random geometric graphs and derive analytical bounds for the node coverage, i.e. the number of queried sensor nodes within a given time frame. Based on this metric, we propose an algorithm that improves the data gathering performance by generating *constrained* random walks, in which the probability mass function at each step reflects the available side information (e.g. the memory of past visited sites).

I. INTRODUCTION

Among the many research challenges posed by wireless sensor networks — i.e. networks of tiny, low-power devices that pick up samples from a physical process in a given area, process their observations, and send the collected data back to a remote data fusion center — the exploitation of mobile nodes for data gathering purposes is one that has so far received only limited attention from the research community¹.

Most contributions that address the data gathering problem in sensor networks focus on distributed data processing techniques combined with broadcast or gossiping protocols for reliable and power-efficient transmission across the network (see. e.g. [1], [2], [3], [4], and [5]). Some contributions attempt to maximize the energy lifetime of the whole network, instead of the lifetime of each individual sensor [6], whereas other focus on exploiting the correlation in the data collected by neighboring sensors [7]. Problems related to sensor nodes operating with low duty cycles are investigated e.g. in [8]. Since all of the previous contributions assume a static sensor network that is dense enough to provide at least one path between any sensor node and the data collection point, it is not surprising that mobility is not taken into consideration. On the other hand, in the area of robotics we find examples such as [9], in which mobile sensors are placed on robots that navigate by dropping and re-visiting sensor nodes serving as exploration markers, and [10], where the same authors present an algorithm for robotic coverage and exploration, which also deals with the problem of sensor network deployment in a simple yet ingenious way.

Our take is to consider a large-scale sensor network with two types of nodes: static sensor nodes with limited connectivity, taking the required measurements locally, and a mobile node,

which patrols the area and collects the desired data. Our main contributions are as follows:

- *Problem Formulation*: We formulate the problem as a classical random walk on a unit square, where a large number of sensor nodes are scattered at random. The key difference between this formulation and well-known problems involving random walks on graphs is that in our case it is not necessary for the mobile node to visit the exact location of a node on the network graph for it to count as visited — it suffices for the mobile node to enter the transmission range of a sensor node for its data to be collected and the node count to be increased.
- *Bounds for Random Walks on Sensor Networks*: Introducing the concept of *node coverage* to describe the effectiveness of different mobile data gathering strategies in terms of the expected number of sensor nodes captured within a given time frame (or, equivalently, a given delay), we provide a detailed mathematical analysis of random walks on sensor networks represented by random geometric graphs and present both inner and outer bounds for its node coverage depending on the given geometric parameters.
- *Algorithms for Constrained Random Walks*: To improve the performance of the node coverage, we propose a methodology in which the probability mass function of the directions the mobile node can choose from is altered in a dynamic fashion based on the available side information at each step. A simple example that introduces memory into the navigation process of the mobile node shows the potential towards improving the node coverage, as shown by numerical results.

The remainder of the paper is organized as follows. A formal problem statement can be found in *Section II*, followed by the theoretical analysis of random walks on random geometric graphs in *Section III*. *Section IV* discusses the addition of constraints on the random walk, and provides numerical examples. The paper concludes with *Section V*, offering a brief discussion of the main results and some directions for future work.

II. MOBILE DATA GATHERING WITH A PATROL NODE

In this section, we start by defining the problem of mobile data gathering in wireless sensor networks with a patrol node and provide the rationale for considering random walks in this context.

¹Parts of this work were carried out while the authors were with LI-ACC/UP. This research was partly supported by the Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology) under grant POSC/EIA/62199/2004.

A. Problem Statement

We model the topology of the wireless sensor network by a random geometric graph, which can be constructed in the following manner [11]: place n nodes uniformly at random onto the surface of a unit square, and connect all nodes within Euclidean distance γ of each other. Let the node v_i represent the sensor indexed by i , and assume that each sensor has the same maximum transmission radius, denoted γ . Since an edge (v_i, v_j) exists if and only if the distance between v_i and v_j is less than γ , we may view edge (v_i, v_j) as a communication link between sensor i and sensor j . For simplicity, in this paper we assume that the sensor nodes are always awake, but the problem setup can be easily extended to account for low duty cycles. Beyond *static* sensor nodes, there exists a *mobile* patrol node denoted by v_0 that performs a random walk on the unit square, as illustrated in *Figure 1*. It is further assumed that the patrol node is capable of communicating with any sensor node located within radius γ . The figure of merit chosen to evaluate the performance of the mobile patrol node is defined as follows.

Definition 1 (Node coverage): We say a sensor is *collected* if the area defined by its transmission radius was visited at least once. The *node coverage* of the random walk $\eta(t)$ is then defined as the expected number of distinct sensor nodes collected until time t .

The main goal of the patrol node, pursuing a random walk with or without constraints and possibly taking samples to complement the ones collected by the sensor network, is to gather the data of as many sensor nodes as possible within a given time frame, or equivalently to *maximize its node coverage*.

B. Random Walks on Sensor Networks

As a first approach towards a trajectory for mobile data gathering with unknown sensor positions we consider random walks on a lattice. Our motivation comes from the following observations:

- 1) the patrol has no information about the topology of

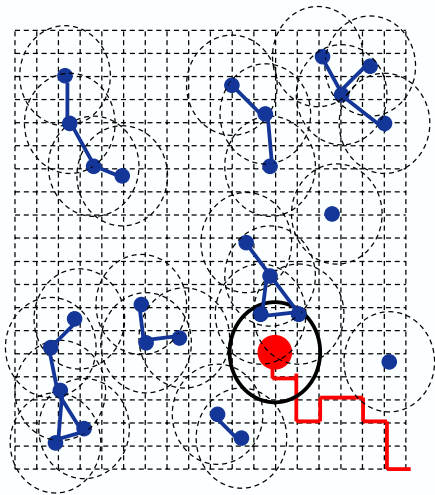


Fig. 1. Problem setup. The large dot represents a mobile node, the small dots represent the sensor nodes. Each dot is surrounded by a circle corresponding to its transmission range. The mobile node performs a random walk on a square lattice and queries every sensor node within its transmission radius.

- the network that would allow for the computation of an optimal trajectory (in fact, even when all the sensor positions are known, this task is likely to be intractable);
- 2) the patrol node is able to collect the data of all sensors within wireless range without actually having to cross each sensor position exactly (this requirement is much less strict than in classical random walk problems, and thus increases the number of nodes that can be visited by a random walk in a given time frame);
- 3) random walks can be equally applied, when the behavior of the sensor nodes is characterized by uncontrolled dynamics [12], e.g. random ON-OFF transitions to save power (this case will be considered elsewhere).

Note that these observations are true for other (more elaborate) mobility models, as well. For simplicity, we choose to run a random walk on a *square* lattice, but our analysis can be easily extended to other instances with more degrees of freedom, e.g. the hexagonal lattice. The ratio between the transmission radius and step size (or, equivalently, the lattice side), must be fixed carefully to avoid the undesirable effects of overlapping covering regions, as illustrated in *Figure 2*. Notice that in the case of a random walk with very short steps, i.e. if the patrol node queries the sensor nodes within range almost continuously as it moves, it is advisable to take some measure to prevent the patrol node from querying the same sensor node multiple times, as this results in unnecessary waste of energy. A simple way to achieve this is to fix the step size of the random walk in relation to the transmission radius and program the patrol node to query the sensor nodes only when it reaches the next position on the square lattice.

From a formal point of view, we may define the described random walk as follows. Let G be a square lattice, i.e. an unweighted graph with maximum degree of four, in which each node corresponds to a pair of coordinates on the field, and the edges are the equivalent to the lattices between the positions, or intuitively the direct paths from one node to its four neighbours. Thus, a random walk on G — clearly a connected graph — is the sequence of vertices visited by a particle that starts at a specified vertex and visits other vertices according to the following transition rule: if the particle is at vertex i at time t , then at time $t+1$ it moves to a neighbour of i picked uniformly at random. The probability of visiting any neighbor of a vertex is thus inversely proportional to the degree of the vertex.

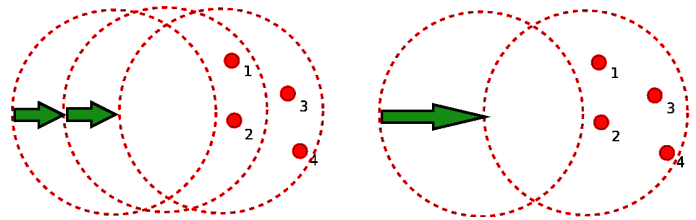


Fig. 2. Two examples to illustrate the relevance of the step size with respect to the mobile data gathering performance. In the case shown on the left, sensors 1 and 2 may be queried twice, thus wasting power. To avoid multiple queries we choose the step size such that each sensor will be visited only once, as shown on the right.

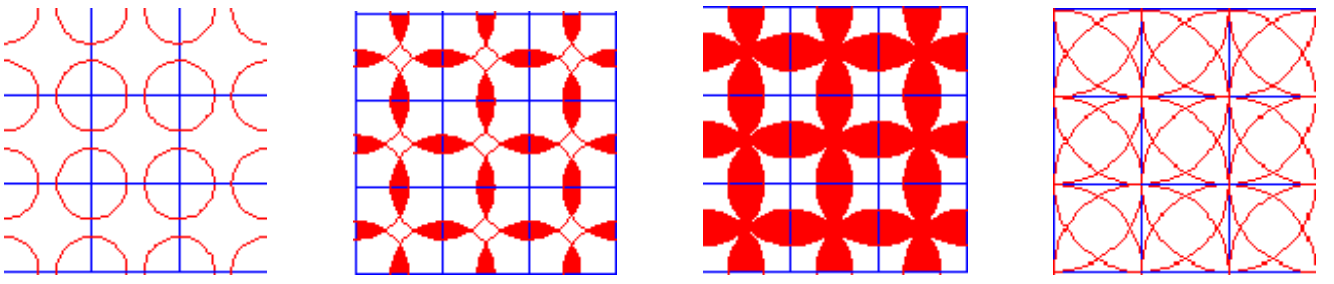


Fig. 3. From left to right, illustration of overlapping for the ratio (i) $\gamma/\mu < 1/2$, (ii) $1/2 < \gamma/\mu < 1/\sqrt{2}$, (iii) $\gamma/\mu = 1/\sqrt{2}$, and (iv) $\gamma/\mu > 1/\sqrt{2}$.

III. NODE COVERAGE FOR UNCONSTRAINED RANDOM WALKS

Our main result is a mathematical characterization of the node coverage for unconstrained random walks in terms of the following quantities (that are either given or easily computed):

- the number of sensors in the field, n ;
- the side of the lattice, μ ;
- the ratio Δ between the intersection of two circles of radius γ and the given area A ;
- the intersection area between a node's coverage area and the area of the field, $\phi = (\gamma^2\pi)/A$;
- a constant $c = 1.8456\dots$ (see [13]);
- the *support* of the random walk, $E[S(t)]$, i.e. the expected number of distinct lattice points (also called *sites*) visited until time t .

The last quantity is at the center of the following lemma, which will be used in our proof.

Lemma 1 (from [13]): Let $E[S(t)]$ be the support of the walk at time t . The average number of sites *not* visited until time t is given by

$$N - E[S(t)] = N(cN)^{-\sigma},$$

where N is the number of possible sites in a finite lattice, $c = 1.8456\dots$, and σ is a time scaling factor such that $t = \sigma\pi^{-1}N \ln^2 cN$.

We are now ready to state and prove the following theorem, which bounds the node coverage for an unconstrained random walk on a sensor network.

Theorem 1: The expected node coverage $E[\eta]$ of an unconstrained random walk until time t is bounded by

$$E[S(t)] n(\phi - 2\Delta) < E[\eta] \leq E[S(t)] n(\phi - \Delta),$$

with

$$E[S(t)] = \frac{1}{\mu^2} \left(1 - \left(c \frac{1}{\mu^2} \right)^{\left(-\frac{t\pi}{\mu^2 \ln^2 \left(c \frac{1}{\mu^2} \right)} \right)} \right),$$

and equality for $\gamma/\mu \leq \frac{1}{2}$.

Proof: Since the n nodes are placed uniformly and independently on the given area, the probability that the patrol node is within the transmission radius of one particular sensor node is given by $\phi = (\gamma^2\pi)/A$. Hence, taking into consideration first only the nodes that are at least γ away from the border of the given area, we have that the expected number of nodes within range of the patrol node results in $n_r = \phi \times n$.

Remark 1: For the remaining nodes that are close to the borders, we know from [14] that the expected degree μ_0 when

$\hat{r}_0 = \gamma/a$, with a denoting the side of the square (in our case, 1), is given by

$$\mu_0 = n\hat{r}_0^2\pi \left(1 - \frac{8}{3\pi}\hat{r}_0 + \frac{1}{2\pi}\hat{r}_0^2 \right).$$

Now, the ratio between the transmission radius and the lattice side will determine whether there is overlapping between the transmission disks at each lattice point, as illustrated in *Figure 3*. We must consider four different cases:

- $\gamma/\mu \leq \frac{1}{2}$ — no overlapping regions; the data of some sensor nodes might not be collected;
- $\frac{1}{2} < \gamma/\mu < \frac{1}{\sqrt{2}}$ — overlapping regions; the data of some sensor nodes might not be collected;
- $\gamma/\mu = \frac{1}{\sqrt{2}}$ — overlapping regions; all sensor nodes will be queried;
- $\gamma/\mu > \frac{1}{\sqrt{2}}$ — overlapping regions; multiple queries of the same sensor nodes (this case can be ignored).

From Lemma 1 it follows the number of sites visited until time t is given by

$$\begin{aligned} E[S(t)] &= N - N(cN)^{-\sigma} \\ &= \frac{1}{\mu^2} \left(1 - \left(c \frac{1}{\mu^2} \right)^{\left(-\frac{t\pi}{\mu^2 \ln^2 \left(c \frac{1}{\mu^2} \right)} \right)} \right) \end{aligned}$$

Based on the previous observations with respect to the ratio between the transmission radius and the lattice side, we obtain the following preliminary results for the node coverage. If $\gamma/\mu \leq \frac{1}{2}$, then there is no overlapping of the transmission disks of the patrol node at each lattice point, the node coverage is simply

$$E[\eta] = E[S(t)] \phi n$$

On the other hand, when $\frac{1}{2} < \gamma/\mu < \frac{1}{\sqrt{2}}$, we must take the overlapping regions into consideration. Bounding the node coverage for this case by the case of $\frac{1}{2} < \gamma/\mu < 1$ yields

$$E[\eta] = E[S(t)] \phi n - E[S(t)] \Delta n,$$

where Δ denotes the area of intersection of the two radiuses of transmission divided by the area A (which again in our case is 1), i.e.

$$\Delta = \gamma^2 (\Theta - \sin(\Theta)),$$

with $\Theta = 2 \arcsin \sqrt{1 - \left(\frac{\mu}{2\gamma} \right)^2}$. Notice that it is possible that the transmission disk of a patrol node at one lattice point also intersects with that of a previously visited lattice point. The number of intersections of this type is difficult to determine

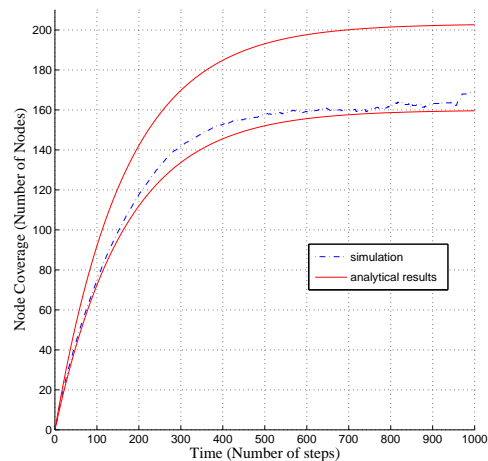
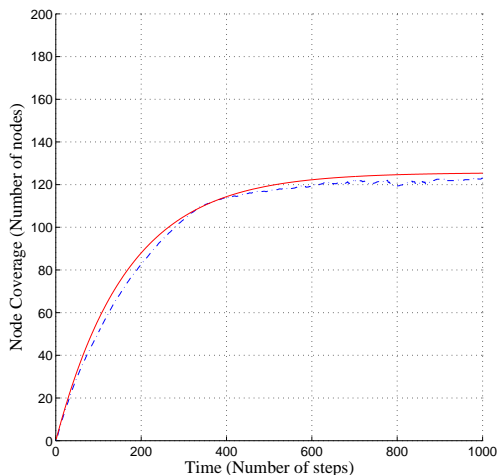


Fig. 4. Node coverage in function of time (number of steps). Left: analytical results versus 500 simulations with $\gamma = 0.05$, $\mu = 0.1$, 200 sensor nodes in the unit square. Only one curve is shown since there are no overlapping regions. Right: analytical results (lower and upper bounds) versus 500 simulations, both for $\gamma = 0.07$, $\mu = 0.1$, and 200 sensor nodes in the unit square (dashed).

as it depends on the topology of the support of the random walk; however, it tends to approximate 1 per distinct site as the number of distinct sites grows.

Thus, we obtain a lower bound for the node coverage, which can be written as

$$\begin{aligned} E[\eta] &= E[S(t)] \phi n - E[S(t)] \Delta n - E[S(t)] \Delta n \\ &= E[S(t)] \phi n - 2 E[S(t)] \Delta n, \end{aligned}$$

which concludes the proof. ■

A. Illustration and Validation of Theorem 1

To validate the previous result we carried out several sets of computer simulations, shown in *Figure 4*. We conclude that the derived bounds are considerably tight in both cases. As expected, the node coverage curves have a steep start and then stagnate. This is due to the fact that, as the number of visited sites increases, the mobile node is likely to spend more and more time in those instead of the not yet visited ones. As we increase the value of the ratio γ/μ , we can see that node coverage progresses at a faster pace, because a larger transmission radius allows more sensor nodes to be captured at each lattice site.

IV. CONSTRAINED RANDOM WALKS

A. Rationale

It is clear that, although classical unconstrained random walks allow for a pleasing analytical treatment, they are far from being an efficient and therefore satisfactory solution. On the one hand, it is possible that a classical random walk will not cover a significant percentage of the area in a reasonable amount of time by visiting the same locations repeatedly; on the other hand, classical random walks do not exploit the fact that sensors might form clusters or provide information to assist the navigation of the mobile patrol node. In order to improve the behavior of the classical random walk with respect to the node coverage, we propose adding the following constraints: at each position the probability of the next direction is altered dynamically based on the available side information (or even on the memory of the patrol node), i.e. the probability mass function of the next

direction is dynamic and not static as in the classical random walk (see *Algorithm 1*).

Algorithm 1 *c_walk* — Generic Constrained Random Walk

```

1: matrix  $\leftarrow$  0
2: while covered_area < area_to_cover do
3:   matrix(position)  $\leftarrow$  current_state
4:   for each quadrant do
5:     determine(probabilityquadrant) based on information of
       matrix)
6:   end for
7:   go to quadrant with probability probabilityquadrant
8: end while

```

As a simple example to illustrate the potential of this methodology, a possible first constraint would be to introduce memory in order to prevent the patrol node from going immediately backwards, i.e. if $a_i \sim a_{i+1}$ then the sequence

$$a_i \rightarrow a_{i+1} \rightarrow a_i$$

is *not* possible. A more sophisticated approach consists in dividing the area in four “quadrants”. The probability that the patrol node chooses one of the quadrants will be weighted according to the number of not visited sites in that quadrant. The probability of jumping to each section is then given by

$$p = \frac{1}{4} \frac{n_0}{(n_0 + n_1)},$$

where n_0 stands for the number of 0’s in the matrix (unvisited sites), and n_1 stands for the number of 1’s (visited sites).

A practical navigation algorithm based on this principle is described in *Algorithm 2*.

B. Numerical Results

The results of our experiments are shown in *Figure 5*. The simulator used was Matlab version R14. In order to ease the required computations instead of using diagonal lines to divide the quadrants, we rotate the lattice by 45°, and use the horizontal and vertical lines that pass through the current position to divide them. Notice that the lattice is still square — the only difference

Algorithm 2 *c_walk_area* — Constrained Random Walk — Area clusters

```

1: covered_area ← 0
2: matrix ← 0
3: while covered_area < area_to_cover do
4:   matrix(position) ← 1
5:   for each quadrant do
6:     sum_quadrant ← sum(matrix(quadrant))
7:     probability_quadrant ←  $\frac{(size_{quadrant} - sum_{quadrant})}{4size_{quadrant}}$ 
8:   end for
9:   go to quadrant with probability probability_quadrant
10: end while

```

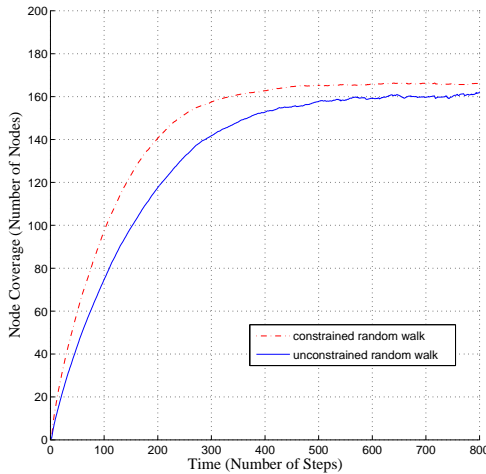


Fig. 5. Node coverage in function of time (number of steps); the simulation was run 500 times with $\gamma = 0.07$, $\mu = 0.1$, and 200 nodes.

are the four possible angles for the random walk, that will now be $\pi/4$, $3\pi/4$, $-3\pi/4$ and $-\pi/4$. It is also important to take into consideration that the lattice side becomes $\sqrt{\frac{\mu^2}{2}}$; and that, since we only use the diagonals of the lattice, the number of sites in the simple random walk is not N but $N/2$.

Adding constraints to the random walk does not alter the overall behavior (steep ascent and strong saturation), but clearly improves the data gathering performance in terms of node coverage. We conjecture that the optimum radius-lattice combination is one with ratio $\gamma/\mu = 1/\sqrt{2}$ — although it presents some overlapping, it is the smallest value for each the entire area is covered without any blank spaces.

V. CONCLUDING REMARKS

We presented a rigorous formulation of the problem of mobile data gathering in a wireless sensor network with one patrol node. Introducing a new concept of coverage — the *node coverage* — we derived analytical bounds for the performance of classical unconstrained random walks on random geometric graphs. To improve this performance we proposed a practical algorithm that constrains the random walk based on the available side information. The concept was illustrated with a constrained random walk based on the memory of previously visited sites.

There are several interesting directions to carry on this work:

- (a) obtain analytical bounds for constrained random walks based on a Markov chain description of their trajectory;
- (b) revise the results in terms of the hexagonal lattice and evaluate the impact of the added degrees of freedom on

the performance of unconstrained and constrained random walks;

- (c) take into consideration other mobility strategies (e.g. [15]) and wireless propagation models (e.g. [16]);
- (d) extend the model to account for dynamic sensor nodes with low duty cycles and increased power awareness;
- (e) analyze the impact of two or more multiple patrol nodes and define strategies that minimize the node coverage and consequently the delay.

Ultimately, we envision a clustered approach in which sensor nodes exchange information on their locations and measurements and guide the mobile node towards the most efficient data gathering path.

REFERENCES

- [1] I. F. Akyildiz, Su Weilian, Y. Sankarasubramaniam, and E. E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.
- [2] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin, “Directed diffusion: a scalable and robust communication paradigm for sensor networks,” in *MobiCom '00: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2000, pp. 56–67, ACM Press.
- [3] Joanna Kulik, Wendi Heinzelman, and Hari Balakrishnan, “Negotiation-based protocols for disseminating information in wireless sensor networks,” *Wirel. Netw.*, vol. 8, no. 2/3, pp. 169–185, 2002.
- [4] S. Lindsey and C. S. Raghavendra, “Pegasis: Power-efficient gathering in sensor information system,” in *Proc. of the IEEE Aerospace Conference*, March 2002.
- [5] H. O. Tan and Ibrahim Korpeoglu, “Power efficient data gathering and aggregation in wireless sensor networks,” *SIGMOD Rec.*, vol. 32, no. 4, pp. 66–71, 2003.
- [6] Konstantinos Kalpakis, Koustuv Dasgupta, and Parag Namjoshi, “Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks,” *Comput. Networks*, vol. 42, no. 6, pp. 697–716, 2003.
- [7] A. Scaglione and S. D. Servetto, “On the interdependence of routing and data compression in multi-hop sensor networks,” in *Proc. ACM MobiCom*, Atlanta, GA, 2002.
- [8] G. Lu, B. Krishnamachari, and C. Raghavendra, “An adaptive energy-efficient and low-latency mac for data gathering in wireless sensor networks,” in *Proc. of the 4th International Workshop on Mobile, Ad-hoc and Sensor Networks*, 2004.
- [9] Gaurav S. Sukhatme and Maxim A. Batalin, “Efficient exploration without localization,” in *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, Taipei, Taiwan, September 2003, pp. 14–19.
- [10] Maxim A. Batalin and Gaurav S. Sukhatme, “The analysis of an efficient algorithm for robot coverage and exploration based on sensor network deployment,” in *Proc. of the 2005 IEEE International Conf. on Robotics and Automation*, Barcelona, 2005.
- [11] Stephen Boyd, Arouta Ghosh, Balaji Prabhakar, and Devavrat Shah, “Mixing times for random walks on geometric random graphs,” in *Proc. of the Second Workshop on Analytic Algorithmics and Combinatorics*, Vancouver, January 2005.
- [12] Sergio D. Servetto and Guillermo Barrenechea, “Constrained random walks on random graphs: routing algorithms for large scale wireless sensor networks,” in *Proceedings of the First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA-02)*, New York, Sept. 28 2002, pp. 12–21, ACM Press.
- [13] S. Caser and H.J. Hilhorst, “Topology of the support of the two-dimensional lattice random walk,” *Phys Rev Lett.*, vol. 77, no. 6, pp. 992–995, August 1996.
- [14] Christian Bettstetter, “Mobility modeling, connectivity, and adaptive clustering in ad hoc networks,” 2003, PhD Thesis, Technische Universitaet Muenchen.
- [15] T. Camp, J. Boleng, and V. Davies, “A survey of mobility models for ad hoc network research,” *Wirel. Commun. Mob. Comput.*, vol. 2, pp. 483–502, 2002.
- [16] Christian Bettstetter and Christian Hartmann, “Connectivity of wireless multihop networks in a shadow fading environment,” in *MSWIM '03: Proceedings of the 6th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems*, New York, NY, USA, 2003, pp. 28–32, ACM Press.