Node Deployment by Stochastic Point Processes in Wireless Sensor Networks

Heitor S. Ramos[1,2], Marcelo G. Almiron[1], Alejandro C. Frery[2], Eduardo F. Nakamura[3], Antonio A. F. Loureiro[1]

[1]Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil

[2]Institute of Computing, Federal University of Alagoas, Maceió, AL, Brazil

[3]FUCAPI, Manaus, AM, Brazil

{hramos, malmiron, loureiro}@dcc.ufmg.br, acfrery@pq.cnpq.br, eduardo.nakamura@fucapi.br

Abstract—**Node deployment plays an important role in the design of wireless sensor networks (WSNs). Many important properties such as coverage, connectivity, data fidelity and lifetime are directed influenced by the way nodes are placed in the sensor field. In this work we show that the use of a stochastic point process, namely the $\mathcal{C}$ model, is able to increase the network lifetime by addressing the "energy hole" problem. Among the $\mathcal{C}$ model properties, this process can be easily simulated and, for any number of nodes in arbitrary areas, it is able to describe situations that range from deterministic (regular) placement of nodes to clusters of sensors in a small region. The model is herein described and is assessed in terms of the coverage, connectivity, data fidelity and lifetime by using computer simulation. In particular, we compare the $\mathcal{C}$ process with the mostly used deployment strategy, namely the uniform randomly placement (URP), being the latter a particular case of the former.**

I. INTRODUCTION

Wireless sensor Networks (WSN) are ad hoc wireless networks consisting of spatially distributed autonomous devices that cooperatively monitor environmental conditions such as temperature, pressure and pollutants. WSNs have been used in various application areas (e.g., health, military, home) [1] where human presence is not possible nor desired [2], [3].

The sensors scattered in a sensor field have the capability to collect and aggregate data and route them to a base station [1]. The base station provides the users of the WSN with the result of these operations, which could be used to reconstruct the phenomena of interest and provide information for making decisions.

Many techniques have been proposed and used for traditional wireless ad hoc networks, but they are not well suited for many WSN applications [1]. For this reason, different protocols and algorithms specific for WSNs have been proposed.

Wu et al. [4] show that the lifetime of a uniformly deployed WSN is limited by the sensors at the first-hop from the base station that is collecting data, a feature known as "energy hole" analytically characterized by [5]. The authors show that increasing the number of nodes cannot desirably prolong the system lifetime under the binomial deployment. Actually, they conclude that the entire system lifetime can be improved by spreading more nodes close to the base station.

Thus, the resulting topology realized by the node deployment phase plays an important role in the WSN design process.

Despite the important role that deployment-induced topology plays in WSNs, studies in this venue are seldom found

in the literature. Actually, uniform random placement (URP) is the most used deployment strategy for the WSN simulations [6]. In order to fill this gap, this work:

- provides a framework to simulate WSNs with different topologies by using Spatial Point Process models as presented in [7], namely the $\mathcal{C}$ process. This model is able to generate three deployment scenarios: (i) attractive, (ii) repulsive (including regular deterministic placement) and (iii) binomial;
- assesses different topologies generated by the $\mathcal{C}$ process by means of: coverage, connectivity, the quality of information reported to the base station, and network lifetime.

The paper unfolds as follows. Section II analyzes the related work. Section III presents the main models we employ in this work, Section IV presents the simulation results, and Section V concludes the paper.

II. RELATED WORK

Optimal node placement is a challenging task since it has been proven to be a NP-Hard problem for most formulations in WSN [8]. Younis and Akkaya [8] present a comprehensive survey on strategies and techniques for node placement in wireless sensor networks. Among them, we will discuss randomly placed static nodes. Most studies consider that this strategy leads to a uniform node distribution but, as stated in [8], such assumption is seldom realistic.

Ishizuka and Aida [9] studied fault-tolerant properties of stochastic node deployment in flat networks. They assumed three models: simple diffusion (bivariate Gaussian distribution), uniform, and R-random. The R-random model is characterized by the density function $f(r, \theta) = 1/(2\pi r)$, where $0 \leq r \leq R$, $0 \leq \theta < 2\pi$, and the sensor position is given in polar coordinates within a distance R from the base station. They conclude that the initial placement of sensors has significant effect on network dependability in terms of tolerance of a node failure, and that the R-random deployment model yields better fault-tolerance properties. This fact leads to the need of concentration of nodes around the base station.

Lloyd and Xue [10] considered a two-tier network model where the nodes are grouped around relay nodes that directly communicate to the base station. They seek for the node deployments that maximizes network lifetime. They show that URP does not extend the network lifetime, since the

relay nodes deplete energy at different rates, depending on the distance to the base station. They then propose a weight random node distribution that concentrates more nodes away from the base station to split the load among the nodes that will consume more energy. This strategy may lead to connectivity problems because some relay nodes can be placed out of the transmission range of the base station. To mitigate such problem, they propose a hybrid deployment strategy that aims to balance the network lifetime and connectivity goals.

We propose a new approach to random deployment by means of a stochastic point process, and we study the same properties Younis and Akkaya [8] discuss that should be optimized. We assume that no adjustments in a sensor node position happens after deployment, thus, all properties studied here are determined only by the deployment strategy. We analyze coverage, connectivity, reconstruction error and network lifetime, and show that the "energy hole" problem can be diminished using the attractive process herein described. A module for the Sinalgo simulation package is provided.

III. The Models

1) The data: Sensors measure continuously varying real functions that describe, for instance, the illumination on the ground of a forest or the air pressure in a room. We model the phenomenon with a zero-mean isotropic Gaussian random field assuming isotropy and a covariance function of the form $\exp(-\boldsymbol{x}^s)$, where $\boldsymbol{x} \geq 0$ is the distance between sites. The scale $s > 0$ is the parameter that characterizes this model; it is related to the granularity of the process.

Figure 1 shows four situations, from fine ($s = 5$) to coarse ($s = 20$), in a red-yellow-white color table. granularities. Samples from this process can be readily obtained using the `RandomFields` package for R [11].

| (a) $s = 5$ | (b) $s = 10$ | (c) $s = 15$ | (d) $s = 20$ |
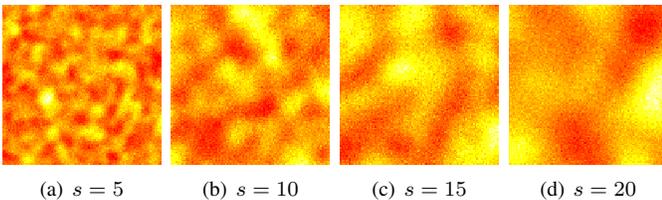
Fig. 1. Gaussian random fields

2) Signal sampling and reconstruction: Sensor i, located at $(x_i, y_i) \in W$, $W \subset \mathbb{R}^2$ being the area of interest, captures a portion of f: the mean value observed within its area of perception p_i, i.e., it stores the value $v_i = \int_{p_i} f$. We chose to work with isotropic homogeneous sensors, where

$$p_i = \{(x, y) \in W : x^2 + y^2 \leq r_s^2\},$$

being $r_s > 0$ the perception or sensing radius.

Kriging is the reconstruction technique we employed [12]. It is a geostatistical method, whose simplest version ("simple Kriging") is equivalent to minimum mean square error prediction under a linear Gaussian model with known parameter values. We used "ordinary Kriging", where the local mean value is estimated from the data and their location.

A. The Deployment Model

A point process is a stochastic model that describes the location of points in space. It is very useful in a broad variety of scientific applications such as ecology, medicine and engineering [13].

The isotropic stationary Poisson model, also known as "fully random", is the basic point process. The number of points in the region of interest follows a Poisson law with mean proportional to the area, and the location of each point does not have influence on the location of the other points. The other process we will use is a repulsive one, where points cannot lie at a position lower than a specified distance. Using these two processes we build a composed point process able to describe many practical situations.

If the number of points is known, and their locations are collectively independent, we have a binomial point process or "uniformly distributed" points, denoted $B(n)$. Samples from a Poisson point process on a rectangular window $W = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$, with $a_i < b_i$, $i = 1, 2$, can be drawn by first sampling from the random variable N_A. If $N_A(\omega) = n$ (assume $n > 0$), then draw (x_j, y_j), $1 \leq j \leq n$, independent samples from independent random variables $X \sim \mathcal{U}(a_i, b_1)$ and $Y \sim \mathcal{U}(a_2, b_2)$ uniformly distributed on the sides of the window. These samples are the coordinates of the n points.

The Matérn's Simple Sequential Inhibition process can be defined iteratively as the procedure that tries to place n points in W [13]. The first point is placed uniformly, and until all the n points are placed or until the maximum number of iterations $t_{\max}$ is reached, a new location is chosen uniformly on W. A new point is placed there if the new location is not closer than r to any previous point; otherwise the location is discarded, the iteration counter is increased by one and a new location is chosen uniformly. At the end, there are $m \leq n$ points in W that lie at least r units from each other. This process describes the distribution of non-overlapping discs of radii $r/2$ on W; denote it $M(n, r)$ assuming W is known.

We build the "step process" by merging two Poisson processes with different intensities on different supports W' and W, such that $W' \subsetneq W$. A step point process, with parameters a and $\lambda > 0$, is defined as two independent Point processes: one with parameter λ on $W \setminus W'$, and the other with parameter $a\lambda$ on W'. Without loss of generality we assume $\lambda = 1$ and denote this process as $S(n, a)$. The compound process on $W = [0, 100]^2$, $W' = [0, 25]^2$ and $\eta = 1$, is

$$\mathcal{C}(n, a) = \begin{cases} M(n, r_{\max}(1 - e^a)) & \text{if } a < 0 \\ B(n) & \text{if } 0 \leq a \leq 1 \\ S(n, a) & \text{if } a > 1, \end{cases}$$

where $r_{\max}$ is the maximum exclusion distance, set to $r_{\max} = n^{-1/2} = 1/10$. Negative values of a yield repulsive models; when $a \to -\infty$ the process becomes more repulsive, tending to the regular deployment. When $0 \leq a \leq 1$ there is no interaction between points, and we have the binomial model.

When $a > 1$ the model is attractive, with the concentration of proportionally more points in W' due to the step process.

Repulsive processes are able to describe the intentional but not completely controlled location of sensors as, e.g., when they are deployed by a plane at low altitude. Sensors located by a binomial process could have been deployed from high altitude, so their location is completely random and independent of each other; this is the URP model. Attractive situations may arise when sensors cannot either be deployed or could function everywhere as, for instance, when they are spread in a swamp: those that fall on a dry spot survive, but if they drown they fail to function. An attractive process is also able to describe the deployment when the designer intentionally increases the node density in a given region. This situation is desired to address the "energy hole" problem by increasing the node density around the base station.

Once the signal $f = F(\omega)$, outcome of the Gaussian random field with parameter $s \in \mathbb{R}$ is available, it will be sampled at positions $(x_1, y_1), \ldots, (x_{100}, y_{100})$ which, in turn, are the outcomes of the compound point process $\mathcal{C}(100, a)$, $a \in \mathbb{R}$.

B. The Energy Model

We consider the general energy model described in [6]. This model assumes that transmission and reception dominate sensors energy consumption and, thus, disregards any other energy spent in other tasks. In particular, we used the first-order model, which considers that a node dissipates energy to run the radio electronics and the power amplifier to transmit data, and dissipates only the radio electronics to receive data. The following equations describe the energy dissipation involved in the communication task:

$$E_{Tx}(\ell, d) = E_{elec}\,\ell + \epsilon_{amp}\,\ell\,d^2, \text{ and } E_{Rx}(\ell, d) = E_{elec}\,\ell,$$

where $E_{Tx}(\ell, d)$ is the total amount of needed energy to transmit ℓ bits to distance d, E_{elec} is the dissipated energy to run the radio electronics and ϵ_{amp} is the energy necessary to run the power amplifier. Since the transmission range of each sensor is r_c (a constant value), the dissipated energy to transmit and receive data is proportional only to the amount of data involved in the communication.

We only accounted the total amount of transmitted and received messages of each node involved in the communication. This metric is related to the total energy depletion, regardless the sensors and the wireless channels they are operating.

C. The Routing Model

We used a simple routing algorithm to estimate the lifetime of the WSN induced by the topologies generated by the $\mathcal{C}(n, a)$ point process and by the communication radius. This algorithm is a variation of the gossip routing algorithm [14].

Each sensor reports its collected data by using a minimum cost path to the base station, being the cost the number of hops towards the base station. However, to distribute the energy depletion as equally as possible, we constructed a routing tree where the nodes store all neighbors that provide the minimum cost to the base station. Each time the node transmits data

to the base station, it will choose, randomly, one of those neighbor nodes that presents the same distance to the base state to be its ancestral.

D. Coverage and Connectivity Models

We considered a WSN with homogeneous sensors with the same sensing (r_s) and transmission (r_t) ranges, both perfect circles. We calculate coverage as the union of the areas of the sensing circles, intersected with the sensor field.

Connectivity is the percentage of sensors able to report their information to the base station using a routing path. Two sensors are able to reach each other if and only if they are located within each other's transmission range.

IV. Simulation Results

A. Preliminaries

The main goal of our simulations is to evaluate (i) coverage, (ii) connectivity, (iii) reconstruction error and (iv) lifetime. While the first three metrics tend to favor repulsive deployments, the fourth has the opposite behavior. Due to the energy hole problem, attractive deployments lead to better energy distribution, extending the network lifetime. The default scenario is presented in Table I.

TABLE I
DEFAULT SIMULATION SCENARIOS PARAMETERS

Parameter	Value
sink node	1 (bottom left-most node)
network size	100 nodes
communication radius	15 m
event duration	50 min
data rate	1 packet/min
sensing radius	5.4 m
sensor field	$100 \times 100 \, \mathrm{m}^2$

We used R v. 2.8.1 [15] for topology generation, and SinalGo v. $0.75.3^1$ for discrete event simulation. Details can be obtained from the first author. Each situation was replicated independently 30 times.

For the rest of this paper, $a \in \{-30, -15, 0, 5, 15, 30\}$ will be referred to as strongly repulsive, fairly repulsive, independent, slightly attractive, fairly attractive and strongly attractive processes, respectively. We studied 80 scenarios varying $n \in \{100, 150, 200, 512\}$, $a \in \{0, 5, 15, 30\}$ and $rc \in \{5, 15, 20, 30, 50\}$.

Coverage and connectivity behaved as expected: improve with repulsive processes, which tend to fill the sensed area.

B. Reconstruction error

We discretized the signals on a $100 \times 100 \, \mathrm{m}^2$ regular grid, so the absolute value of the relative error is

$$\varepsilon^{(s,a)}(F, \widehat{F}) = \frac{1}{10^4} \sum_{1 \le i,j \le 100} \left| \frac{F(i,j) - \widehat{F}(i,j)}{F(i,j)} \right|,$$

[1] Sinalgo, "Simulator for network algorithms," 2008, Distributed Computing Group - ETH-Zurich, 2008. [Online]. Available: {http://dcg.ethz.ch/projects/sinalgo}

provided $F(i, j) \neq 0$, which is granted with probability 1 by the continuous nature of the Gaussian random field. We studied 24 different scenarios , namely for the factors $s \in \{5, 10, 15, 20\}$ and $a \in \{-30, 15, 0, 5, 15, 30\}$.

Figure 2 depicts the simulation results for the reconstruction error: cells represent the scale factor s, and the curves represent the deployment process. We observe that, regardless the deployment, the greater the scale of the Gaussian random field, the lesser the reconstruction error. Regarding the deployment, we observe that the more repulsive the process, the lesser the reconstruction error; this is predictable, since there are only 100 sensors and the coverage rate is less than 50% under the attractive deployment process. The reconstruction error decreases when the coverage increases; thus, for this metric the more repulsive the deployment, the better the quality.
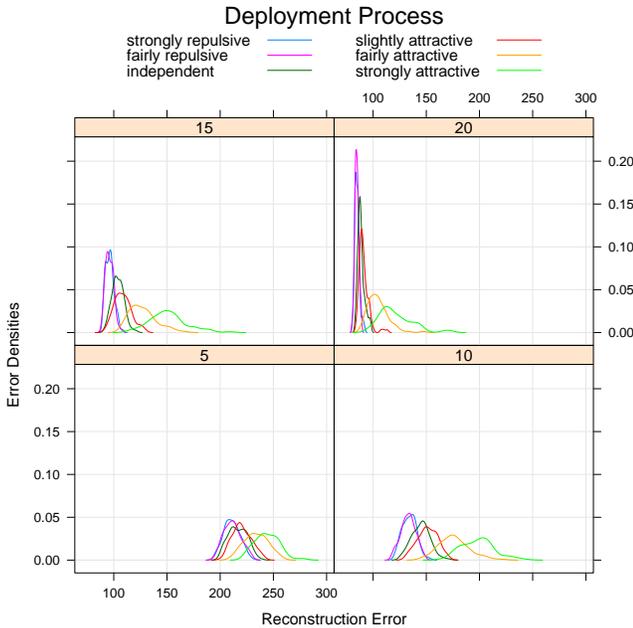
Fig. 2. Estimated Reconstruction Error Densities

C. Lifetime

The lifetime of a sensor is the time it can receive and transmit data until it runs out of energy. The whole network lifetime is a more sophisticated subject and, for the sake of simplicity, we only assess the total amount of messages transmitted and received by the nodes situated one hop away from the base station. For this, let us define the random variables that model the total amount of transmitted and received messages, respectively. We studied 72 different scenarios regarding those random variables, namely, the factors $n \in \{100, 150, 200\}$, $a \in \{-30, 15, 0, 5, 15, 30\}$ and $r_c \in \{15, 20, 30, 50\}$ for the nodes located one hop away from the base station.

Figures 3(a) and 3(b) depict the simulation results in semi-logarithm scale, respectively. The columns show the number

of nodes, the lines show the communication radius and the colored curves show the deployment processes.

Figure 3(a) shows that when the transmission range is high, e.g., 50 m, the curves present similar shape when the network size grows. This follows from the fact that when the communication radius grows, more sensors are able to send their data directly to the base station, and a few nodes will perform the relay role.

Strongly and fairly attractive deployment strategies behave alike, i.e., they exhibit the lowest amount of data sent for all situations. Slightly attractive processes also present low amount of sent data, but the density is more spread, indicating that the relay task is more concentrated in some nodes. When the communication radius decreases, the variability increases for all deployment processes, however, the variability increases slower when the deployment process is attractive.

For communication radii 15, 20 and 30 under repulsive (strongly and fairly) and independent processes one observes multimodal curves. They indicate that the energy dissipation is concentrated in few sensors and, therefore, the network lifetime will be degraded in such situations.

Even for low values of the communication radius, such as 15 m, attractive deployments split the relay task among the nodes near the base station. For instance, with 100 nodes the independent deployment tends to present few nodes distant one hop from the base station. Therefore, these few sensors will relay all data sent by all other sensors. This is the behavior that we intend to avoid because these few sensors tend to run out of battery early diminishing, thus, the overall network lifetime.

For communication radii 15, 20 and 30 m with repulsive deployment, curves are spread. It means that some nodes relay lots of data packets, while others relay a few. Only when the communication radius is high, repulsive deployments tend to present few relayed data by the sensors.

Under attractive deployments, the number of sensors has little influence on the transmitted data relayed by the nodes close to the base station: though the number of data packets increases, the number of sensors near the base station grows proportionally, and they will split the relay task more evenly. Although sensor density has little influence on the transmitted data, it has strong influence on the coverage and connectivity.

Figure 3(b) shows that deployment has little influence on the received messages in nodes one hop far from the base station. The number of received data packets increases with the number of nodes. This is important, since the dissipated energy receiving messages is of the same order of magnitude as that required to transmit. Therefore, the more attractive the deployment, the better the distribution of energy among the sensors situated close to the base station.

To improve lifetime without compromising data fidelity, an attractive deployment may be used increasing the number of sensors, but with care because when the number of sensors grows, the number of received data packets also increases.

(a) Transmitted Messages

(b) Received Messages
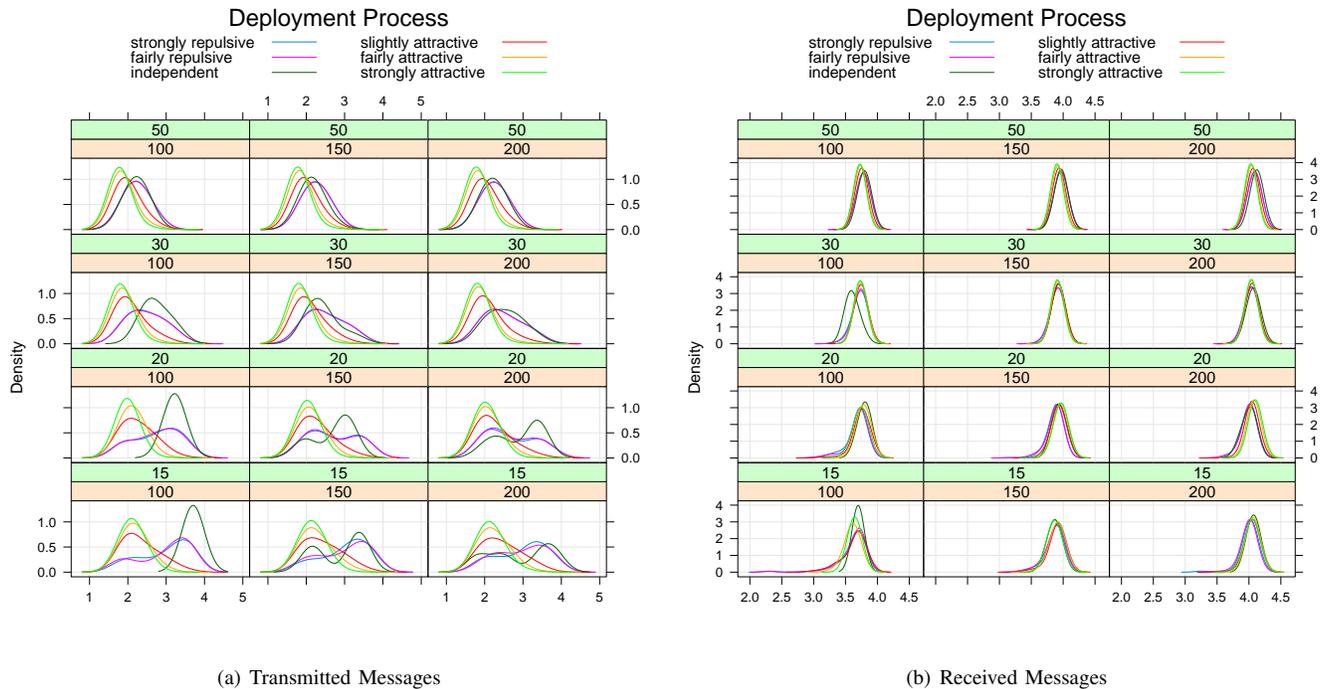
Fig. 3. Estimated Transmitted and Received Messages Densities

V. CONCLUSIONS AND OUTLOOK

We presented a comprehensive evaluation of the influence of the deployment strategy in four aspects of wireless sensor networks namely, coverage, connectivity, reconstruction error and lifetime. We conclude that: (i) coverage and connectivity diminish with the attractivity of the deployment, although this fact can be alleviated with the deployment of more sensor nodes; (ii) reconstruction error increases with the attractivity of the deployment, although this fact can be alleviated with the deployment of more sensor nodes, as well; (iii) network lifetime can be improved with the attractivity of the deployment.

A typical optimization problem is stated here, once coverage, connectivity and error conflict with network lifetime. To address this problem some actions can be taken depending, among other factors, on the project budget. Actually, the network designer will decide how to deploy the nodes and how many nodes will be necessary to maintain the QoS in acceptable levels for the application at hand.

This work can be used as a guide for the network designer to deal with those issues. The topology generator proposed in this work, an open-source project, is available at http://www.dcc.ufmg.br/~hramos/topology. It is ready to be used in the Sinalgo, but it can be extended to other simulation packages.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cyirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, March 2002.

[2] J.-H. Cui, J. Kong, M. Gerla, and S. Zhou, "The challenges of building mobile underwater wireless networks for aquatic applications," *IEEE Network*, vol. 20, no. 3, pp. 12–18, December 2006.

[3] O. Younis, M. Krunz, and S. Ramasubramanina, "Node clustering in wireless sensor networks: Recent developments and deployment challenges," *IEEE Network*, vol. 20, no. 3, pp. 20–25, December 2006.

[4] X. Wu, G. Chen, and S. K. Das, "Avoiding energy holes in wireless sensor networks with nonuniform node distribution," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 5, pp. 710–720, 2008.

[5] J. Li and P. Mohapatra, "Analytical modeling and mitigation techniques for the energy hole problem in sensor networks," *Pervasive and Mobile Computing*, vol. 3, no. 3, pp. 233–254, 2007.

[6] D. M. Wang, B. Xie, and D. P. Agrawal, "Coverage and lifetime optimization of wireless sensor networks with Gaussian distribution," *IEEE Transactions on Mobile Computing*, vol. 7, no. 12, pp. 1444–1458, Dec. 2008.

[7] A. C. Frery, H. S. Ramos, J. Alencar-Neto, E. Nakamura, and A. A. F. Loureiro, "Data driven performance evaluation of wireless sensor networks," *Sensors*, vol. 10, no. 3, pp. 2150–2168, 2010.

[8] M. Younis and K. Akkaya, "Strategies and techniques for node placement in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 6, no. 4, pp. 621–655, june 2008.

[9] M. Ishizuka and M. Aida, "Performance study of node placement in sensor networks," in *ICDCSW '04: Proceedings of the 24th International Conference on Distributed Computing Systems Workshops - W7: EC*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 598–603.

[10] E. L. Lloyd and G. Xue, "Relay node placement in wireless sensor networks," *IEEE Transactions on Computers*, vol. 56, no. 1, pp. 134–138, 2007.

[11] M. Schlather, "Simulation and analysis of random fields," *R News*, vol. 1/2, pp. 18–20, 2001.

[12] P. J. Diggle and P. J. Ribeiro Jr., *Model-based Geostatistics*. Springer, 2007.

[13] A. Baddeley, "Spatial point processes and their application," in *Stochastic Geometry*, ser. Lecture Notes in Mathematics, W. Weil, Ed. Berlin: Springer, 2006, vol. 1892, pp. 1–75.

[14] J. Al-Karaki and A. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

[15] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2009, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org