Performance Evaluation and Enhancement of Surface Coverage Relay Protocol

Antoine Gallais and Jean Carle

IRCICA/LIFL, Univ. Lille 1, INRIA Futurs, France
{antoine.gallais,jean.carle}@lifl.fr
http://www.lifl.fr/POPS

Abstract. Area coverage protocols aim at turning off redundant sensor nodes while ensuring full coverage of the area by the remaining active nodes. Connectivity of the active nodes subset must also be provided so that monitoring reports can reach the sink stations. Existing solutions hardly address these two issues as a unified one and very few are robust to non ideal physical conditions. In this paper, we propose a deep analysis and some enhancements of a localized algorithm for area coverage, based on Surface Coverage Relays (SCR) and able to build connected active nodes sets that fully cover the area. We first enhanced the critical phase of our protocol (the relay selection) and show that the number of active nodes can be drastically reduced. We then raise the issue of the robustness of the protocol once a realistic physical layer is simulated. Our algorithm proved itself to be an interesting solution as it remained able to still ensure high coverage level under realistic physical layer conditions. We also added the possibility to finely tune the overall proportion of active nodes through a new parameter used during local relay selection phases.

Key words: Wireless sensor networks, area coverage, configurable localized algorithm, realistic physical layer.

1 Introduction

Wireless sensor networks are made up of hundreds of devices in which a battery, a sensing module and a wireless communication device are embedded. They are deployed over hostile or remote environments, in which they become one-use-only since their batteries cannot easily be replaced or refilled. Energy consumption is therefore balanced by taking advantage of the redundancy induced by the random deployment of nodes; some nodes are active while others are in sleep mode, thus using less energy. Such a dynamic topology should not impact the monitoring activity. The ensuing issue consists in the active nodes to fully cover the area. Furthermore, as monitoring reports should at least reach the sink stations, the set of active nodes should be connected. This is also known as the connected area coverage problem.

This work focuses on a localized algorithm, considering first that the availability of a deciding central entity is a strong assumption and also willing to show that such an organization can be obtained from networks without any infrastructure. We proposed in [1] a fully localized solution to the connected sensor area coverage issue. Both a neighbor detection phase and a relay selection phase are required to complete the activity decision process. In this article, we deeply study the relay selection phase, made from a new method described here. This latest solution allows us to strongly reduce the proportions of active nodes while not inducing any communication overhead nor any new information in control messages. Most of existing area coverage solutions rely on an ideal communication model (each node has a transmission threshold, beyond which no message may be directly sent). The radio channel randomness is hardly considered. We studied the robustness of our SCR protocol with such a new assumption. After showing that it was slightly impacted, we introduce a new parameter during the relay selection phase in order to vary the size of the relay set thus finely tuning the proportions of active nodes.

2 Preliminaries

We assume that randomly deployed sensors are time-synchronized, static and position-aware. Several solutions have been studied to achieve these goals [2]. The location of a sensor node is used as its unique identifier. We also consider that sensors are able to determine whether an area is covered by a set of sensors or not. Various *coverage evaluation schemes* have already been used for that purpose [3, 4].

2.1 Communication and sensing models

A wireless network is modeled by a graph $G = (V, E)$, V being the set of vertices and $E \subseteq V^2$ the set of edges that the available communications. An edge between two vertices u and v exists if u is physically able to send a message to v. The neighborhood set of u, noted as $N(u)$, is defined as:

$$N(u) = \{v \in V \mid v \neq u \wedge (u, v) \in E\}. \tag{1}$$

Each sensor has a communication range R_c and a sensing range R_s. We denote by $\mathrm{S}(u)$ the area covered by a node u and $\mathrm{S}(A)$ the area covered by a set of nodes $A = \{a_1, a_2, \ldots, a_n\}$ such that:

$$\mathrm{S}(A) = \bigcup_{i=1}^{i=|A|} \mathrm{S}(a_i). \tag{2}$$

2.2 Radio channel model

Given a graph $G = (V, E)$, a communication range R_c and the Euclidean distance between nodes u and v $\mathrm{dist}(u, v)$, the *unit disk graph* model defines the set of edges E as:

$$E = \{(u, v) \in V^2 \mid u \neq v \wedge \mathrm{dist}(u, v) \leq R_c\}, \tag{3}$$

Yet, in a practical context, transmission quality varies over the time and a node u may be able to communicate with node v at time t, but not at time $t+1$. Each communication link has a probability to exist which is influenced by a lot of factors such as the emitting power or the distance between the emitter and the receiver. Many models exist and some generators have already been proposed to allow model generation from parameters such as the kind of material used or the localization of the sensors themselves (see [5] for instance). To replace the unit disk graph, we opted for the *lognormal shadowing* model thus transforming G into a weighted graph, where the weight of each edge $(u, v) \in E$ is equal to the probability of correct reception $\mathrm{p}(\mathrm{dist}(u, v))$ for the two nodes u and v. An approximated function $\mathrm{P}(x)$ is described by Kuruvila et al. [6]:

$$\mathrm{P}(x) = \begin{cases} 1 - \frac{(\frac{x}{R_c})^{2\alpha}}{2} & \text{if } 0 < x \leq R_c, \\[2ex] \frac{(\frac{2R_c - x}{R_c})^{2\alpha}}{2} & \text{if } R_c < x \leq 2 \times R_c, \\[2ex] 0 & \text{otherwise,} \end{cases} \tag{4}$$

α being a factor that highly depends on the environment and x being the considered distance. This function assumes that the probability of correct reception for the range R_c is always equal to $\mathrm{P}(R_c) = 0.5$. Fig. 1 illustrates this function for $\alpha = 2$, which is frequently encountered in the literature. Being aware of the approximation that is made here, we are not pretending to use a highly realistic model. Yet, we believe that considering radio channel randomness during protocol analysis and simulation constitutes a first step towards facilitating further developments on real material.

3   Related Work

Many papers have already addressed the problem of connected area coverage in wireless sensor networks once sensors are allowed to self-schedule their activity. Some have assumed a given topology to start the study, such as a grid for instance [7]. In [8], authors
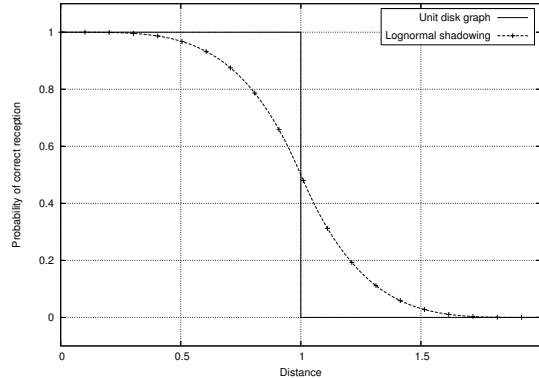
Fig. 1. The two considered physical models ($R_c = 1, \alpha = 2$).

study some deployment patterns in order to ensure both coverage and connectivity once wireless sensors are deployed over a sensing field. Meanwhile, the deployment might be random thus forcing sensor devices to make decisions based on non predictable information.

Tian and Georganas [9] have proposed a node scheduling scheme in which time synchronized and randomly deployed sensor devices regularly detect their neighbors and listen for retreat messages for a given timeout. A node can get passive if remaining neighbors fully cover its area. To ensure connectivity of the set of active nodes, a strong assumption is made regarding the ratio between sensing and communication radii ($R_s \leq R_c$). Such an assumption had already been used and proved as preserving network connectivity as long as area coverage was provided [3, 4].

Recently, a randomized algorithm was proposed in [10]. Based on a Markov model, this algorithm allows each node to probabilistically turn off while preserving both coverage and connectivity. Yet, the need for a new analysis when considering non homogeneous settings might be an obstacle in real deployments where neighbor information is required to ensure coherent node behavior once the neighborhood is modified for any reason (mobility, failure).

4 Base of the SCR-CADS protocol

While coverage might not be a strict requirement for wireless sensor networks, connectivity among the set of monitoring nodes is crucial as every collected piece of information should be able to reach one of the sink stations. This is why we propose a new look at our SCR-CADS protocol which has been proved to ensure both connectivity and coverage in an original manner [1]. We show how this algorithm could be configured and especially what its resistance to new communication models is. Our solution consists in three consecutive phases (a neighbor discovery mechanism, a relay selection phase and a decision making process), each being detailed in this section.

Regular neighbor discovery A classical neighbor discovery phase is first required. At the beginning of each round, a *hello* message is sent by each node. It contains information such as its geographical position and a priority, assumed to be unique among the set of nodes.

Relay selection When considering node coverage, Jacquet et al. [11] have defined the notion of multipoint relay (MPR). Each node selects a subset of its neighborhood; the set of relays. For any given node u, this set, denoted as $MPR(u)$, enables multi-hop communications with every 2-hop neighbors:

$$\text{MPR} : V \rightarrow V$$
$$u \longmapsto MPR(u)$$

$$\forall u, \exists\, \text{MPR} \,|\, N(\text{MPR}(u)) \backslash N(u) \;=\; N(N(u)) \backslash N(u), \tag{5}$$

V being the whole set of nodes. We can observe that $MPR(u) = N(u)$ satisfies the relation. Yet, trying to have a MPR set as little as possible leads to diminishing the size of the dominating set, denoted as MPR-DS. Each node must send its own relay set to its whole neighborhood so that a simple decision rule can be made. Authors have proved that some rules were providing MPR-DS that were connected.

As we aim at providing area coverage by connected sets, we have extended this solution to the problem of area coverage in wireless sensor networks. The relay selection phase has been modified. Each node computes its own relay set; this set must cover as large an area as the whole set of neighbors. This set is called Surface Coverage Relay (SCR) set. We defined a SCR function for this purpose:

$$\text{SCR} : V \rightarrow V$$
$$u \longmapsto SCR(u)$$

$SCR(u)$ standing for the subset of $N(u)$ for which the property holds:

$$\forall u, \exists\, \text{SCR} \,|\, S(\text{SCR}(u)) \backslash S(u) \;=\; S(N(u)) \backslash S(u). \tag{6}$$

Each node must be able to compute its SCR set. Finding the subset of minimal size is equivalent to the *minimum set cover problem*, which is NP-complete [12]. This is why heuristics have to be designed in order to compute SCR sets that are as optimal as possible considering the number of nodes involved.

Initially, a heuristic based on centroid computation was used [1]. In this article, we propose a new heuristic to compute SCR sets. Each node starts with an empty SCR set. Then, step by step, a neighbor is added to the SCR set if and only if it covers a portion of area yet uncovered by the SCR set. To get interesting subsets, at each step, the neighbor that brings the more new coverage should be added. Ideally, the neighbors should therefore be ordered according to their coverage potential regarding previously observed nodes.

The first heuristic was based on a simple idea; at each step, we try to find the node that should bring the more coverage. To avoid heavy computation, we opted for the one whose location is the furthest from the *center of mass* of already selected relays (every

relay having the same *mass*). It turned out that still large sets were computed. We now propose a new heuristic for selecting the relays, based on a simple intuition; the furthest neighbor is the one that covers the largest portion of area outside the sensing zone of the selecting sensor. Neighbors are therefore ordered according to their distance to the selecting node. The furthest is added to SCR set, then the second furthest is evaluated and so on.

Note that for any heuristic, considered neighbors have to be distinguished in case of equality when facing the choice criterion. Which node is considered first is determined with a simple random function.

Once SCR sets have been computed, each node has to decide its own activity status. We now explain the decision rule used for this purpose.

Making activity decision In the MPR-DS protocol [11], nodes apply a simple rule to decide their status (dominant or not). This rule is based on a key, assumed to be unique among the set of nodes.

Definition 1. *Every node u whose key is the lowest among the neighborhood or which belongs to the relay set of the neighbor with the lowest key must be active.*

As already shown in [11], if relay sets are computed so that the property 5 holds, then the induced set of active nodes is connected. We have already proved that using this rule with SCR sets was providing a set covering as large an area as the whole set of nodes. It is also obvious that SCR sets are MPR sets once $R_s = R_c$ since 2-hop neighbors are necessarily located within one or several sensing areas of the 1-hop neighbors. This means that the covering set is also connected once we every activity decision has been made. Now that we have presented SCR-CADS and a new heuristic to compute SCR sets, we are evaluating this protocol under several communication models. Indeed, it has already been largely spread that protocols evaluated under ideal communication assumptions (i.e. the unit disk model) were more prone to failures or to dysfunction once used in practice (see [13] for some instances among area coverage protocols). The remaining of this article is therefore focused on the analysis of SCR-CADS protocol and its performances with a realistic physical layer for communications.

5 Performance evaluation

We have simulated only homogeneous sensor networks. Experimental results were obtained from randomly generated connected networks with a discrete event simulator. Nodes are deployed over a 50×50 rectangle area, considering a Poisson point process of intensity $\lambda > 0$. We define the density d of the network to be the average number of nodes in a given communication area: we thus have $d = \lambda \times \pi R_c^2$. Both the communication range (R_c) and the sensing radius (R_s) are fixed at 10. Simulations were launched over densities varying from 20 to 90 nodes per communication zone, with a step of 10. For each density, the number of performed iterations is adjusted so that 95% of the results are in a sufficiently tight confidence interval. Each iteration consists of rounds. Each round starts with the neighbor discovery phase, followed by the SCR

sets computation and the decision making process. A sensing area is modeled as a disk while the communciation model varies (unit disk model or a realistic physical layer). When nodes have to compute the area covered by a given neighbor, they use a coverage evaluation scheme based on a disks intersections theorem, already applied in [3, 4]. A round ends with the sensing period which involves every node that has decided to be active. We essentially measure the percentage of active nodes and the percentage of preserved coverage of original area.

5.1   Under ideal communication assumptions

We give complementary results about the SCR protocol when simulated with ideal communication assumptions. By ideal, we mean that no message can ever be lost, neither due to message collisions nor to the communication environment itself.

Area coverage and connectivity As already mentioned, the algorithm has been proved as preserving both area coverage and connectivity of the set of active nodes. We therefore focus on the size of the active nodes sets in order to evaluate the impact of the heuristic used during the relay selection phase.

Size of SCR sets We have observed the average size of the SCR sets in order to compare the performances of the original heuristic with the ones the latest solution, based on distances. The size of the relay sets is a matter of prime importance first because the number of active nodes is strongly linked to these and second because they determine the communication overhead induced by the protocol. Concerning the first point,every node whose key is the lowest among its neighborhood is necessarily active. This number does not vary as long as the network density remains identical. Yet, the larger the relay sets, the higher the probability for a node to belong to the SCR set of the neighbor with the lowest key and so to become active. Concerning the second point, as relay sets must be sent to every 1-hop neighbor, the less large they are, the better for the communication overhead and thus the energy consumption.

   We could observed that the distance-based heuristic was actually performing better than the centroid-based one. As shown on Fig. 2(a), the average size of the SCR sets generated by the distance-based heuristic varies from roughly 8 neighbors to 11, respectively for densities 20 and 90, while the centroid-based method leads to 11 neighbors selected as relays at density 20 and nearly 24 at density 90, that is more than twice the size obtained by our distance-based heuristic.

   As already mentioned, the size of the relay sets straightly impacts the number of active nodes. As our new heuristic allows us to considerably reduce this size, we are now looking to what extent this improved relay selection impacts the average number of active nodes that are involved at every round.

Active nodes Figure 2(b) shows the average number of active nodes versus the network density.

   As expected, the percentage of active nodes decreases as the network density increases; the more nodes there are, the less the proportion of required sensors to achieve
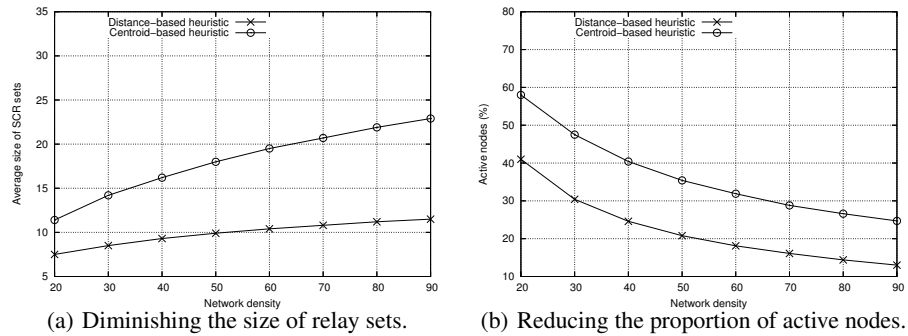
(a) Diminishing the size of relay sets.   (b) Reducing the proportion of active nodes.

Fig. 2. Main advantages of the distance-based heuristic for relay selection.

full coverage of the area. Interestingly, we can observe that the centroid-based heuristic leads to 24.7% of active nodes at density 90 while the distance-based method lowers this result to only 13% for the same density. We therefore obtained very interesting results with this new method.

5.2   Under realistic physical layer assumptions

We now try to evaluate the SCR-CADS protocol with a realistic communication model. The goal for such an investigation is twofold. First, we aimed at testing the robustness of this algorithm and then some potential enhancements were to be implemented and also evaluated.

Introducing a more realistic physical layer leads to more channel randomness, that is communication links that are probabilistic; two nodes can communicate with a given probability depending on the distance that separates them. Some area coverage protocols have already been studied under this new assumption and we aim at evaluating SCR protocol also. In our algorithm, once the relay sets have been computed, they are sent so that every node can make its activity decision from the rule previously described in 4. The so computed set of active nodes is connected and covers as large an area as the whole set of nodes once the unit disk model is used. When a realistic physical layer is used, we could observed some changes.

Area coverage Figure 3(a) shows the average percentage of covered area by the sets of active nodes computed by our SCR protocol.

We can observe that while area coverage was fully preserved with an ideal communication model, this property does no longer holds under our new assumptions. Area coverage goes from a bit more than 95% at density 30 to less than 90% at density 80. Yet, this loss of coverage is not as important as we could have expected, and especially not as dramatic as the ones we could already observed with other protocols [13].

This shows an inherent robustness for our SCR protocol. Let us detail the reasons why the impact is that low. During the neighbor discovery phase, *hello* messages from

some of the theoretical 1-hop neighbors may not be received by a given node u. Yet, this actually increases the probability for u to evaluate its own key as being the lowest among its neighbors. Then, the less neighbors are discovered, the higher the probability for u to get active, the worst case being no *hello* message ever received and u having an empty neighbor table, thus necessarily deciding to be active. In such a case, we would have all nodes being leaf and thus active, which is not the case here as some coverage losses have still been recorded.

We may now focus on the relay selection phase. Once relays have been selected, SCR sets are sent to the neighbors. If some of these messages are never received by a given node u, then it results in some nodes being aware of only a subset of its real SCR selectors. If one of these potentially missed selectors is the neighbor with the lowest key so far, such a wrong reception leads to u not knowing it has been selected by its neighbor with the lowest key, which would have meant it had to be active. In such a case, u decides to be passive and therefore jeopardizes area coverage. This is why some coverage losses are observed on Fig. 3(a). However, note that if a relay message is received from a yet unknown node whose key is the lowest among already collected ones, then the receiving node must consider this message as a *hello* message. It must also check if it has been selected as a relay by this node. This demonstrates the inherent robustness of SCR protocol since only coverage losses may happen only in case the relay message of the neighbor with the lowest key (and whose previous *hello* message was received) was never received. This means that only this message is crucial while others can only bring some new information that could help nodes make the right activity decision.
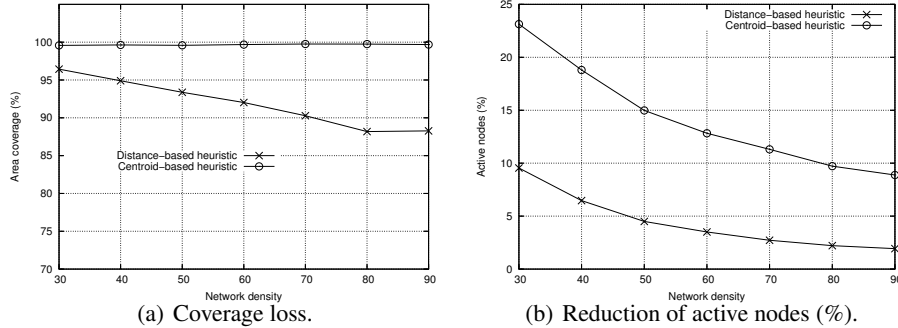

(a) Coverage loss.   (b) Reduction of active nodes (%).

Fig. 3. SCR-CADS algorithm and a realistic physical layer (LNS, see formula 4, with $\alpha = 2$).
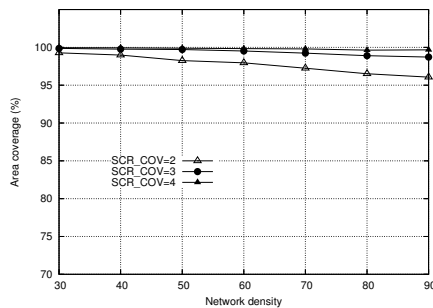
Active nodes We also observed to what extent the average percentage of active nodes was impacted by our new physical layer assumptions. Generally, when some coverage losses occur, it is either caused by a decrease of the number of active nodes or by a spatial incoherence (as many active nodes but concentrated on a given region at the

expense of another one which is left uncovered). The latter case can hardly take place thanks to the homogeneity of both our deployment and our random processes used during key computation. Indeed, fig. 3(b) shows that the percentage of active nodes is largely decreased (from less than 10% at density 30 to only 2% at density 90) when the distance-based heuristic is used.
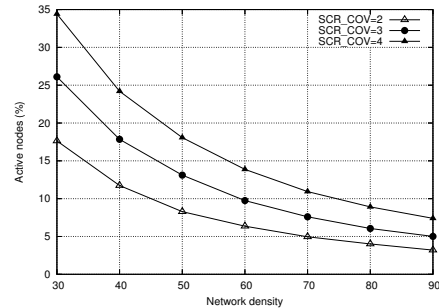
Although this tremendous drop of the percentage of active nodes could lead us to think that large coverage holes have occurred, we have already observed that the impact on area coverage was minor. The SCR-CADS protocol is inherently robust and does not much suffer from loss of active nodes since the remaining ones have self-elected with a coherent local process.

5.3   Tending to more robustness and customization of SCR protocol

Although the impact of a realistic physical layer is not important on area coverage, we could observe that the number of active nodes was largely decreased. As already mentioned, the percentage of active nodes is very much induced by the average size of the relay sets. We are now trying to tune the relay selection made by each node in order to be able to tune the global percentage of active nodes, and also the area coverage. We therefore introduced a coverage parameter, denoted as SCR_COV, which stands for the number of relays that should cover any portion of the sensing area of a neighbor currently evaluated for potential selection. Every neighbor is now added to the set of relays if and only if its sensing area is not k-covered by already selected relays, with k being SCR_COV. Several definitions of k-coverage exist in the literature. Our approach is to consider a k-covered piece of area as soon as k distinct sensor devices are able to sense it. This should help us to raise the proportions of active nodes as relay sets could be made larger. By this way, the resistance to some communication anomalies should be obtained thanks to the redundancy induced in the set of monitoring nodes. Note that this solution solely requires a modified *coverage evaluation scheme*.

(a) Increasing robustness to radio channel randomness.

(b) Configurable proportions of active nodes with enhanced relay selection.

Fig. 4. Configurability of SCR-CADS protocol (LNS, see formula 4, with $\alpha = 2$).

Area coverage The prime goal of our SCR_COV parameter is to be able to better resist to message loss. Figure 4(a) shows area coverage versus network density, with $SCR\_COV \in \{2, 3, 4\}$ when the distance-based heuristic is used. As long as $SCR\_COV > 1$, area coverage is at least 95%. This shows that, even in presence of a realistic radio channel, our algorithm can provide satisfying results in terms of area coverage. Our new parameter may also help us to customize the relay selection more finely. Until now, we were only using our two selection heuristics. The problem with such a method is that dynamically modifying the way a relay selection is conducted by a node would require changing the code it actually runs. Given the hardware constraints of the target devices, it would be more convenient to interact with the algorithm by the means of a simple parameter.

Active nodes We observed to what extent the SCR_COV value was able to impact the proportions of active nodes. The average percentage of active nodes is shown on Fig. 4(b), for several SCR_COV values. When $SCR\_COV = 2$ (which is sufficient to ensure more than 95% of covered area), we have between 17% and 4% of active nodes. This shows that our SCR_COV not only deserves to be used to improve area coverage but also to finely tune the proportions of active nodes.

6 Conclusion

We have enhanced our SCR-CADS protocol in order to have a more efficient relay selection that fits in today's requirements for wireless sensor networks algorithms. We have shown that our algorithm, originally designed under ideal communication assumptions, could work properly with a realistic physical layer, maintaining a high coverage level. We then introduced a new parameter during the relay selection, thus enabling full coverage in case this would be strongly required by the application. This allows to finely tune the proportions of active nodes as we tend to a more configurable protocol.

Future work will focus on characterizing connectivity under more realistic communication assumptions. We also aim at observing the impact of both synchronization and localization errors which could be induced either due to the environment or to the used algorithm itself.

Acknowledgments

References

1. Carle, J., Gallais, A., Simplot-Ryl, D.: Preserving area coverage in wireless sensor networks by using surface coverage relay dominating sets. In: Proceedings of IEEE Symposium on Computers and Communications (ISCC), Cartagena, Spain (2005) 347–352
2. Römer, K.: Time Synchronization and Localization in Sensor Networks. PhD thesis, ETH Zurich, Switzerland (2005)
3. Zhang, H., Hou, J.C.: Maintaining sensing coverage and connectivity in large sensor networks. Ad Hoc and Sensor Wireless Networks journal (AHSWN) **1** (2005) 89–123
4. Xing, G., Wang, X., Zhang, Y., Lu, C., Pless, R., Gill, C.: Integrated coverage and connectivity configuration for energy conservation in sensor networks. ACM Transactions on Sensor Networks (TOSN) **1**(1) (2005) 36–72
5. Cerpa, A., Wong, J.L., Kuang, L., Potkonjak, M., Estrin, D.: Statistical model of lossy links in wireless sensor networks. In: Proceedings of International Symposium on Information Processing in Sensor Networks (IPSN). (2005)
6. Kuruvila, J., Nayak, A., Stojmenović, I.: Hop count optimal position based packet routing algorithms for ad hoc wireless networks with a realistic physical layer. In: Proceedings of IEEE Mobile Ad hoc and Sensor Systems (MASS), Fort Lauderdale, FL, USA (2004)
7. Shakkottai, S., Srikant, R., Shroff, N.: Unreliable sensor grids: Coverage, connectivity and diameter. In: Proceedings of IEEE International Conference on Computer Communications (INFOCOM), San Francisco, CA, USA (2003)
8. Liu, C., Xiao, Y.: Random coverage with guaranteed connectivity: Joint scheduling for wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems (TPDS) **17**(6) (2006)
9. Tian, D., Georganas, N.D.: Connectivity maintenance and coverage preservation in wireless sensor networks. AdHoc Networks Journal (Elsevier Science) (2005) 744–761
10. Yener, B., Magdon-Ismail, M., Sivrikaya, F.: Joint problem of power optimal connectivity and coverage in wireless sensor networks. Wireless Networks **13**(4) (2007) 537–550
11. Adjih, C., Jacquet, P., Viennot, L.: Computing connected dominated sets with multipoint relays. Ad Hoc and Sensor Wireless Networks journal (AHSWN) **1**(3) (2005) 27–39
12. Karp, R.M.: Reducibility among combinatorial problems. In Miller, R.E., Thatcher, J.W., eds.: Complexity of Computer Computations. Plenum Press, New York (1972) 85–103
13. Gallais, A., Ingelrest, F., Carle, J., Simplot-Ryl, D.: Preserving area coverage in sensor networks with a realistic physical layer. In: Proceedings of IEEE International Conference on Computer Communications (INFOCOM), Anchorage, AK, USA (2007) short paper.