

Path Diversity in Energy-Efficient Wireless Sensor Networks

Pascal Mérimodol

IP Networking Lab (INL)

Université catholique de Louvain (UCL), Belgium
pascal.merindol@uclouvain.be

Antoine Gallais

LSIIT (UMR CNRS 7005)

Strasbourg University (UdS), France
gallais@unistra.fr

Abstract—Energy efficiency is one of the most important issue to be tackled in wireless sensor networks. Activity scheduling protocols aim at prolonging the network lifetime by reducing the proportion of nodes that participate in the application. Among the vast range of criteria existing to schedule nodes activities, area coverage by connected sets is one of the most studied. Active nodes must ensure area coverage while remaining connected in order to guarantee proper data collection to the sink stations. As wireless communications stand for the main source of energy consumption, we investigated the communication redundancy of the active nodes set. We define a path diversity based metric that allows to characterize the communication redundancy of a given set of nodes. We show that one of the most used connectivity criterion is far from building minimal connected sets in terms of communicating nodes involved. Our results open new directions to design localized connected sets solutions.

I. INTRODUCTION

Wireless sensor networks [1] keep raising attention in the network protocols community. These tiny objects equipped with minimalistic features (e.g. low computation and storage abilities, low communication range and rate, limited sensing range) raise several issues, especially regarding the solutions that should be imagined to ensure long-lasting deployments. By long-lasting, we mean networks able to observe the deployment environment for as long as possible with minimal intervention. Turning some deployed sensors into a low energy consumption mode is one of the envisioned solutions for energy-efficient wireless sensor networks. Activity scheduling protocols indeed aim at constructing sets of active nodes while passive ones save energy. Such build sets must meet the goals initially imposed by the application for which the whole network has been deployed. For instance, considering a monitoring application, some nodes can be allowed to be passive if and only if active ones can monitor as large an area as all nodes do. With sensors limited in terms of computation and storage abilities, collected data are treated by more powerful machines, referred as the sink stations. One of the major concerns common to all wireless sensor networks deployments is then the reachability of at least one sink station from every active node. This is why most of activity scheduling solutions aim at preserving the connectivity of the active nodes set.

The prime objective of this article is to study these protocols from the communication point of view. Indeed, we wondered

whether requiring the connectivity of the active nodes set was a viable constraint regarding the real n -to- m communication paradigm of wireless sensor networks (n being the number of nodes and m the number of sink stations). We propose to use a metric quantifying the path diversity in order to evaluate the communication redundancy of an active nodes set. Our study relies on area coverage protocols. They aim at scheduling nodes so that the set of active sensors fully covers the target area and remains connected. We limited our study to localized solutions that describe nodes making decisions solely based on local information. Once all local decisions have been made, both area coverage and connectivity should be preserved. We detail one of the most widespread local connectivity criterion in order to show that solutions based on it actually leave the network with a very high connectivity redundancy. For all pairs of nodes, we also investigate the stretch of alternate path costs compared to optimal ones. Our evaluation shows that alternate paths costs are close to optimal ones, thus reinforcing the fact that the generated path diversity within the set of active nodes can be exploited for objectives such as fast re-routing, load balancing or energy efficiency. This paper therefore concludes that solutions relying on local mechanisms ensuring global connectivity of the graph remain sub-optimal, thus meaning that much more nodes could have their forwarding ability removed. As wireless communications are the main source of energy-consumption, we believe that our results pave the way towards new kinds of energy-efficient sensor networks.

The remainder of the paper is organized as follows. We give some definitions and notations in Sec. II-A. The connected area coverage problem is presented in Sec. II-B while our path diversity metric is detailed in Sec. II-C. In Sec. III, we describe the methodology that we use to evaluate the path diversity in energy-efficient wireless sensor networks. Evaluation results are discussed in Sec. IV while some concluding remarks and hints for future works are given in Sec. V.

II. CONTEXT AND MODELS

A. Notations

The communication area of a node u is modeled as a disk of range CR , centered at u . Two nodes u and v are said to be neighbors as long as $dist(u, v) < CR$, $dist$ being the euclidean distance between u and v . The neighborhood of a node u is noted as $N(u)$ ($|N(u)|$ being the degree of u).

The network is therefore represented as a graph $G(V, E, w)$, V being the set of nodes, E being the set of edges (the communication links of the network) and w being a positive valuation of edges. The valuation function w is used to define a metric related to the paths, generally additive, especially if $w = \text{dist}(u, v)$ as it is the case in our study. Two nodes are connected as long as a set of edges linking u and v (a path) exists. A network is connected if all pairs of nodes $\in V$ are connected. Section II-C gives the notations related to our evaluation of the path diversity. Concerning coverage modeling, the sensing ability of a sensor node u has for long be modeled as a disk of range SR , centered at u . Recently, this model has been enhanced to be made more realistic [2]. Here, we use the lognormal shadowing model, generally used to model radio communication links. In this context, an approximated function was described by Kuruvila et al. in [3]. We transpose this model for the needs of sensing modeling as detailed in [4].

B. Area coverage with connected sets

We decided to consider activity scheduling in wireless sensor networks through the area coverage criterion. Area coverage should be ensured by connected active sets in order to ensure reliable data collection by the sink stations. Existing protocols aim at reducing the proportions of active nodes while preserving full area coverage by connected active nodes sets. Several solutions have already been proposed while relying on purely local decisions. Nodes therefore make decisions solely based on available local information. The ensuing goal is to obtain a global coherent behavior (i.e. full area coverage by connected sets) while keeping small active nodes sets or low control traffic induced by the activity scheduling protocol itself. Depending on the target optimization (proportion of active nodes, induced control traffic), proposed contributions either aim at drastically diminishing the number of involved active sensors [5] or propose to allow relatively larger proportions with lower communication overhead [6]. All solutions are studied for networks of variable density, thus showing the scalability of these protocols. For a given density, some works have also proposed to vary the coverage degree [7] or to relax the coverage constraint once a local detection threshold is considered [8]. In every case, the connectivity of the active nodes set is provided along with full area coverage.

To ensure connectivity, several solutions have been proposed. One of them is to rely on the SR/CR ratio. In [9], authors show that ensuring full area coverage is enough to guarantee connectivity as long as $CR > 2 \times SR$. Then, most of existing contributions make this assumptions and solely focus on area coverage preservation. Other solutions are based on connected dominating sets [10] that are built with a local decision making process. A node u can indeed become passive provided that its set of active neighbors is connected, thus meaning that u is not essential to the local connectivity. Many reasons stand for not having the network connectivity relying on the $CR > 2 \times SR$ assumption. This ratio may either not stand in reality or could be heterogeneous over the set of

nodes. In cases where coverage is not to be fully preserved (e.g. probabilistic coverage), there is no guarantee that the involved graph is connected anymore. Still, solutions based on connected dominating sets impose further local knowledge (e.g. 2-hop information to locally compute the connectivity of the neighbor set), thus increasing the communication overhead induced by the activity scheduling protocol.

Solutions proposed in the field of wireless sensor networks require area coverage by connected active nodes sets without defining the connectivity in a sensor network context. Yet, a given sensor does not need to be able to reach any other sensor as it solely aims at reaching one of the sink stations ($n - to - m$ communication paradigm). In the context of localized protocols, local criteria help ensuring the graph connectivity. This means that there exists at least one path connecting two distinct nodes of the network. While using local criteria implies low communication overhead, it may overestimate the number of forwarding nodes. To the best of our knowledge, no evaluation has ever highlighted the number of existing paths for any pair of nodes among the active nodes set. Though, such an evaluation would allow a more precise characterization of the graph connectivity. Indeed, for a given pair of nodes, an excessive number of communication paths would mean that the set of forwarding nodes could be reduced. This would considerably ease the design of new energy-efficient solutions. Our contribution is to evaluate the graph connectivity through a localized path diversity metric.

C. Path diversity

In conventional networks, the diversity of forwarding paths can be defined according to several criteria (e.g. the total number of routes [11], the resulting maximum flow, the number of local deflections [12]) and strongly depends of the chosen forwarding mode (e.g., hop-by-hop, source routing, label-switching). When the routing plan is ensured by a hop-by-hop forwarding scheme, the deployed routes have to be loop-free to ensure a correct distributed composition [11]. Therefore, the path diversity depends on the constraints related to the underlying forwarding plan, and especially on the utilization of alternates paths (load balancing, fast re-routing). In our study, we focus on a local and static analysis of the number of viable next hops discovered by each sensor. Such criteria provide a general evaluation of the path diversity not relying on previously stated constraints. Let s be a source node and d a destination node. We focus on the number of neighbors of s ($|N(s)|$) able to forward the data collected by s towards d . Such a node $v \in N(s)$ provides an elementary path (a simple path without cycle) not containing the edge (v, s) nor s itself (no outgoing link towards s). The subset of nodes $v \in N(s)$ verifying this property is denoted as $F(s, d)$. We use $|F(s, d)|$ to quantify the diversity in terms of the number of local next hops. By this way, we can evaluate whether the space of forwarding nodes could be reduced or not. To refine this analysis, we also investigate the path costs distribution. Let $C_v(s, d)$ be the cost of the optimal elementary path between s and d , using v as the

first hop. The optimal cost over all neighbors is formulated as $\min(C_v(s, d))$, $\forall v \in N(s)$, $\forall s, d \in V$. First, we denote $m \in N(s)$ the neighbor minimizing the distance $C_v(s, d)$. Then, we analyse the distribution of the costs of alternate paths in contrast to $C_m(s, d)$. We now detail the methodology used to obtain this set of measures.

III. METHODOLOGY

In this paper, we use an activity scheduling protocol proposed in [6]. This area coverage solution relies on a local coverage evaluation mechanism. This phase is essential to ensure that a node would not decide to be passive without being fully covered locally. In order to tune the quality of the coverage ensured by the set of active sensors, our evaluation uses both a probabilistic detection model (see Sec. II-A) and an adapted coverage evaluation scheme that is detailed in [4]. A node evaluates its coverage depending on a local sensing threshold. If every physical point of its theoretical sensing area is covered with a probability greater than the given threshold, then the area is said to be covered. In the context of our evaluation, the variation of this threshold allows us to obtain several active nodes proportions for networks of a given density. Unlike most of existing solutions, this protocol preserves the network connectivity without relying on any assumption regarding the SR/CR ratio. Indeed, it relies on a local connectivity criterion that prevents any node to be passive if its set of active neighbors is not connected. This protocol is therefore symptomatic of the previously described solutions as it allows to tune the coverage constraint but does not propose any similar option for the overall network connectivity. In order to evaluate the connectivity redundancy induced by this local connectivity criterion, we evaluate the path diversity of the set of active nodes.

We decided to measure the diversity of the collecting data paths in a distributed and localized manner. For that purpose, from each node $s \in V$, $|N(s)|$ instances of the Dijkstra algorithm allow us to determine the set of the costs $C_v(s, d)$, $\forall v \in N(s), \forall d \in V$. For each neighbor v (for each instance of the Dijkstra computation), we remove all outgoing links of s except the link (s, v) . Therefore, when $C_v(s, d) < \infty$, we know that there exists an elementary path through v to reach destination d . If the number of neighbors v verifying $C_v(s, d) < \infty$ is not equal to $|N(s)|$, then there exists at least one destination that s cannot reach through each of its neighbors. In this case, that means that there exists at least one bridge link connected to s (a sufficient condition is the presence of a node of degree 1 connected to s).

In contrast, when each node s can use all its neighbors v to reach all destinations d (through a path containing neither s nor any cycle in general), that means that the considered graph is at least 2-edge connected (as it does not contain any bridge link). Indeed, for a given destination d , the ability of s to use any of its neighbors to reach d , implies that each neighbor can use a path towards d not containing a link connected to s . If we generalize this property to all pairs (s, d) , we know that there is no link whose suppression would imply the graph

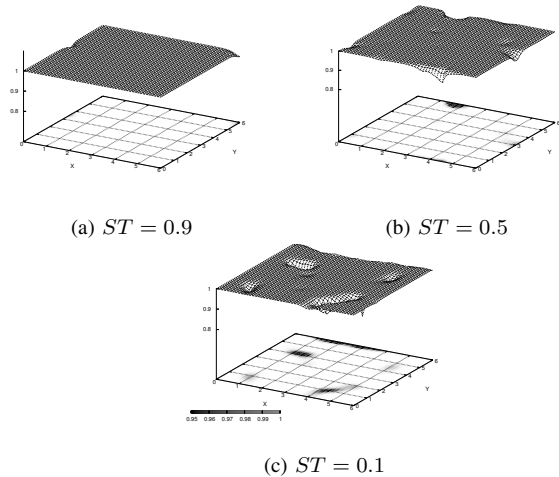


Fig. 1. Probability of point coverage all over the target area.

to be partitioned in several connected components. Thus, if k denotes the minimal degree of the graph, then the graph is k -edge-connected.

Our analysis is done at two levels. First, we measure the proportion of nodes connected to bridge links. We indeed quantified the number of nodes s verifying $N(s) = F(s, d)$, $\forall d \in V$. The set of nodes verifying this property is defined as the set H . Second, we evaluate the alternate path costs distribution as it highlights the dispersion of the alternate paths costs compared to optimal ones.

IV. EVALUATION

Our simulations use the Positive-Only (PO) protocol proposed in [6]. Nodes are deployed over a 50×50 area and have equal theoretical sensing and communicating ranges ($SR = CR = 10$). We also used a realistic sensing model (see Sec. II-A). During its activity decision, a node u uses the coverage evaluation scheme detailed in [4]. Then, for a given topology, the modification of a local sensing threshold (denoted as ST) allows us to vary the global proportion of active nodes. Finally, for the cost distribution analysis, a single destination randomly chosen among the set of active nodes was considered.

A. Controlling sensing activity with local criterion

Depending on the local sensing threshold, for a given density, the overall active nodes proportion may vary. Figure 1 shows the area covered by the same original network topology, reduced in an active nodes topology. The active nodes density directly depends on the local sensing threshold. We can observe that this local sensing threshold logically impacts the coverage quality. Indeed, once the threshold is kept low (0.1 for instance), more physical points have decreased coverage probabilities, thus leading to larger coverage holes on the figure. Depending on the application, this local sensing threshold can be adapted to relax the coverage constraint. In this paper, it allows us to consider various active nodes proportions for the same given density. Yet, as solely the coverage evaluation is locally modified, we wanted to observe to what extent this variation of the active nodes proportion impacted the overall connectivity of the sensor network. We

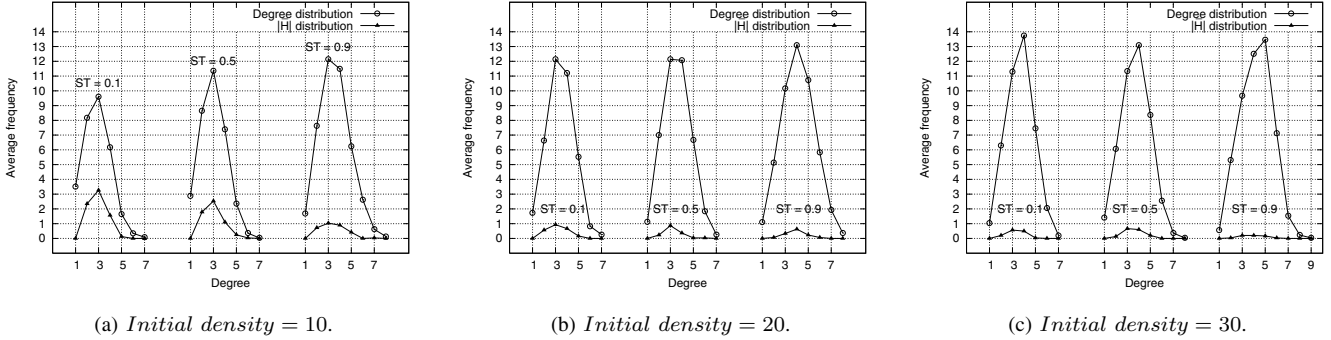


Fig. 2. Degree and $|\bar{H}|$ distribution with three distinct sensing thresholds.

therefore observed the path diversity as a connectivity indicator in the induced communication graph.

B. Path diversity induced by activity scheduling

In order to characterize the connectivity of our communication graphs, we decided to quantify the path diversity as defined in Sec. III. Fig. 2 represents this diversity for three distinct network initial densities (10, 20 and 30) once the activity scheduling protocol has determined the set of active nodes. Results were obtained with the metric using $dist(u, v)$ as a valuation function of edges. Note that it could be any other parameter such as the probability of correct reception for instance, in case of a more realistic communication model.

Fig. 2 shows both the degree distribution of the active nodes and the distribution of $|\bar{H}|$ according to the degree of s . Note that \bar{H} stands for the complementary of H . An element s of $|\bar{H}|$ verifies $\exists d \in V \mid F(s, d) \subset N(s)$. For a given degree, a given density and a given coverage requirement (ST), we can observe the average frequency of nodes having this degree and the average number of nodes whose the number of valid next hops for all potential destinations differs from its degree. When density increases or when the coverage requirement is made harder, we can observe a similar behavior. The path diversity increases as well, $|\bar{H}|$ being strongly reduced. The lower the $|\bar{H}|$ average frequency is, the more nodes in the neighborhood can act as forwarders to reach a potential sink. This is the case when the density is high or when the coverage constraint is strong (0.9 for instance), the active nodes proportion and so the communication density being increased.

Fig. 3 shows the three densities once full coverage is required. There are very few nodes of degree 1 (a leaf node implies the existence of a bridge link), a low distribution of $|\bar{H}|$ means that there is no bridge link, except those originated by leaf nodes. Our results shows that once full coverage is required, the path diversity is very high as nearly every neighbor is a potential forwarder to reach any other node of the network. This observation allows us to conclude that it is possible for all sensors to easily select a subset of forwarders neighbors among its neighborhood. Moreover, when a node is able to use all of its neighbors to reach any sink, that means that this node is not a bridge in the topology, and can deactivate its forwarding capability.

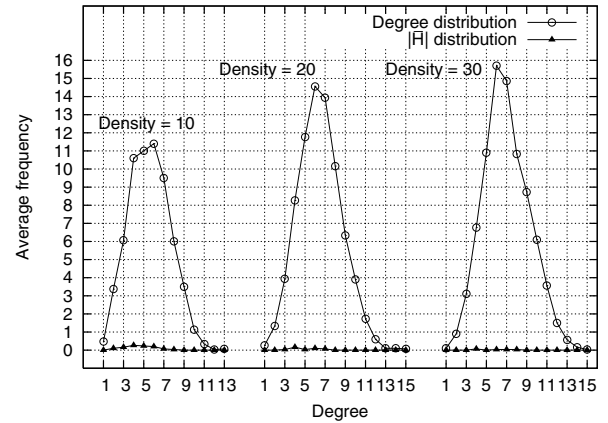


Fig. 3. $|\bar{H}|$ distribution once full area coverage is required.

Fig. 4 shows the cost distribution of every path (optimal and alternate) from every sensor to a given sink station. Note that this figure illustrates the paths of a given topology, with a given sink station randomly chosen among the set of active nodes. Results have indeed shown that neither the location of the sink station nor its degree had significant impact on this distribution. For any node s , we give the length of the optimal path towards the sink station d and the costs of all alternate paths for the same pair (s, d) . The upper plot is an abstract representation of a path whose cost is twice the one of the optimal path to give an insight of the quality of the alternate paths. Nodes are sorted according to their distance to the sink. The nodes closest to the sink have more scattered alternate paths. This means that messages have to go through much longer paths, which is quite logical regarding the fact that these nodes are located to one or two hops of the sink stations, thus imposing long roundabout paths once the optimal path is not used. This is mainly due to the area coverage protocol. Indeed, the communication area including both nodes s and d is also a sensing area ($SR = CR$). Nodes s and d being active (the source and the sink station), there is a high probability that the area coverage protocol has turned intermediate nodes into passive mode, those being useless for local coverage. Therefore, few nodes could stand for forming alternate paths, thus imposing long roundabout ways. Inversely, once nodes are further from the sink station, alternate paths remain of reasonable length. We could observe

that in most of topologies, each sensor can select any of its neighbor as a forwarder to any destination without significant loss in routing quality (the vast majority of alternate path costs being within $[C_m(s, d), 2 \times C_m(s, d)]$, m being the next hop of the optimal path).

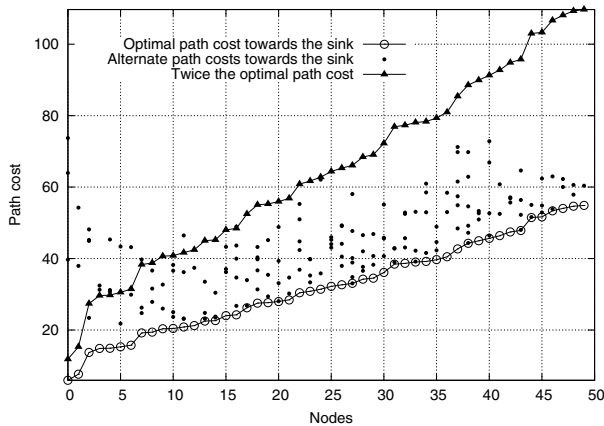


Fig. 4. Distribution of the path costs ($density = 20$, $ST = 0.9$).

Our results show that the local connectivity criterion used so far in most of activity scheduling solutions leaves networks with a very large path diversity. To end this study, we propose to give some hints in order to take advantage of this diversity. On the one hand, it may help the network to be more robust as the number of alternate paths from each sensor towards any sink stations is large. Indeed, alternate paths can be used for fast re-routing purposes to circumvent node failures for instance. Such a feature is of high interest, especially in wireless sensor networks where nodes could be prone to failures. It could also be used to balance the load during the routing task. Once some neighbors start running out of energy, the data traffic could be oriented toward nodes with higher capabilities, without having to wonder whether this node will be able to reach the sink or not. On the other hand, we can consider that this large diversity allows for more energy savings. We can indeed reduce the set of active nodes in order to minimize the path diversity induced by nodes required for connected area coverage. Yet, such a reduction would also impact the coverage quality in the same way as the modification of the local sensing threshold did.

V. CONCLUSION AND FUTURE WORKS

We observed activity scheduling through solutions tackling the connected area coverage problem. After detailing how the coverage constraint could be relaxed, we gave results showing that a widespread local connectivity criterion induces a large path diversity among the set of active nodes thus leaving several possibilities for future works. Indeed, this paper shows that most of the time, all neighbors of a node can act as a forwarder to any destination in the network thanks to a elementary path. We also observed that besides the optimal path, almost all alternate paths remain of good quality (almost each of them having a cost inferior to twice the cost of the optimal path).

As previously discussed, load-balancing and fast re-routing policies could take advantage of this diversity while the network would also be highly robust in terms of sink reachability from any given sensor node. An important future work is also to consider this large diversity as an opportunity to investigate activity scheduling policies that even more reduce the active nodes proportions by considering enhanced activity states. For instance, a *sensing-only* state could be considered. It would consist in solely sending sensed data while not participating in multi-hop communications, thus meaning that *sensing-only* nodes would never forward communications from other sensors but could still help ensuring full area coverage. With leaves being the *sensing-only* nodes, this refers to the Maximum Leaf Spanning Tree Problem, known to be NP-hard [13]. Hence, for a given sink d , a reverse shortest-path tree rooted at d and maximizing the number of leaf nodes can be approximated using ϵ -approximation algorithms in a centralized manner. Our evaluation emphasized the fact that the proportion of *sensing-only* nodes may be significantly high. In order to be coherent with our methodology, we would try to define local decisions able to build sets with minimized communication diversity. Those should ensure a minimal connectivity guarantee (a tree). This would result in tending to the near optimal bound achieved by a centralized solution, thus maximizing energy savings related to communications.

REFERENCES

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks (Elsevier)*, vol. 38, pp. 393–422, 2002.
- [2] M. Hefeeda and H. Ahmadi, "A probabilistic coverage protocol for wireless sensor networks," in *Proc. of IEEE ICNP*, 2007, pp. 41–50.
- [3] J. Kuruvila, A. Nayak, and I. Stojmenović, "Hop count optimal position based packet routing algorithms for ad hoc wireless networks with a realistic physical layer," in *Proc. of IEEE MASS, USA*, 2004, pp. 1267–1275.
- [4] A. Gallais, J. Carle, D. Simplot-Ryl, and I. Stojmenović, "Ensuring k-coverage in wireless sensor networks under realistic physical layer assumptions," in *Proc. of IEEE Sensors*, Korea, 2006, pp. 880–883.
- [5] D. Tian and N. D. Georganas, "Connectivity maintenance and coverage preservation in wireless sensor networks," *Ad Hoc Networks (Elsevier)*, vol. 3, pp. 744–761, 2005.
- [6] A. Gallais, J. Carle, D. Simplot-Ryl, and I. Stojmenović, "Localized sensor area coverage with low communication overhead," *IEEE Transactions on Mobile Computing (TMC)*, vol. 7, no. 5, pp. 661–672, 2008.
- [7] Z. Abrams, A. Goel, and S. Plotkin, "Set k-cover algorithms for energy efficient monitoring in wireless sensor networks," in *Proc. of IEEE/ACM IPSN, USA*, 2004, pp. 424–432.
- [8] Y. Cai, W. Lou, M. Li, and X.-Y. Li, "Target-oriented scheduling in directional sensor networks," in *Proc. of IEEE INFOCOM, USA*, 2007, pp. 1550–1558.
- [9] H. Zhang and J. C. Hou, "Maintaining sensing coverage and connectivity in large sensor networks," *Ad Hoc and Sensor Wireless Networks journal (AHSWN)*, vol. 1, pp. 89–123, 2005.
- [10] I. Stojmenović, M. Seddigh, and J. Zunic, "Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks," *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, vol. 13, no. 1, pp. 14–25, 2001.
- [11] P. Méridol, J.-J. Pansiot, and S. Catelain, "Improving Load Balancing with Multipath Routing," in *Proc. of IEEE ICCCN*, Virgin Islands, 2008, pp. 1–8.
- [12] X. Yang and D. Wetherall, "Source selectable path diversity via routing deflections," in *Proc. of ACM SIGCOMM*, Italy, 2006, pp. 159–170.
- [13] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, New York: W. H. Freeman & Co., 1979.