Investigating Coverage and Connectivity Trade-offs in Wireless Sensor Networks: the Benefits of MOEAs

Matthias Woehrle, Dimo Brockhoff, Tim Hohm, and Stefan Bleuler

Abstract How many wireless sensor nodes should be used and where should they be placed in order to form an optimal wireless sensor network (WSN) deployment? This is a difficult question to answer for a decision maker due to the conflicting objectives of deployment costs and wireless transmission reliability. Here, we address this problem using a multiobjective evolutionary algorithm (MOEA) which allows to identify the trade-offs between low-cost and highly reliable deployments—providing the decision maker with a set of good solutions to choose from. For the MOEA, we use an off-the-shelf selector and propose a problem-specific representation, an initialization scheme, and variation operators. The resulting algorithm is applied to a test deployment scenario to show the usefulness of the approach in terms of decision making.

Key words: Evolutionary multiobjective optimization, variable-length representation, wireless sensor networks.

1 Introduction

WSNs are a new form of pervasive and distributed computing infrastructure, deeply embedded into the environment. Providing remote access to the sensing devices, WSN technology is a radical innovation for many diverse application areas such as environmental monitoring (Mainwaring *et al.*, 2002), structural monitoring (Xu *et al.*, 2004), or event detection (Meier *et al.*, 2007). Monitoring phenomena in a given environment requires coverage of the area with the sensing devices. For remote access to the sensed data, sensor nodes provide an unreliable wireless communication infrastructure. Data

Matthias Woehrle · Dimo Brockhoff · Tim Hohm · Stefan Bleuler
Computer Engineering and Networks Lab, ETH Zurich, 8092 Zurich, Switzerland
e-mail: `firstname.lastname@tik.ee.ethz.ch`

is transmitted via multiple hops along a defined path via intermediate nodes. These paths are constructed based on neighborhood information of individual nodes, relying on the quality of node-to-node links. Reliable data transport providing the user with sensed data is of utmost importance. Connectivity requires operable links between nodes and redundant communication paths to compensate for node failures. Coverage needs to be established to provide quality of data. Cost considerations limit the number of deployed nodes.

The deployment of a WSN, i.e., placing nodes in a given environment, is a complex task. The decision for a node placement needs to consider the aforementioned conflicting constraints and objectives. In order to explore these non-trivial trade-offs, we propose to employ MOEAs. In this paper, we make the following contributions: based on a WSN deployment model by Woehrle *et al.* (2007), we propose objectives and constraints to be used for exploring the trade-offs in WSN deployments. In detail, we propose a variable length representation MOEA, including new variation operators and apply the MOEA to a test deployment.

2 Related Work

Although several approaches for the deployment of WSNs have been proposed in the literature, there is no work employing a realistic deployment model for nodes and the environment and at the same time exploring the intricate trade-offs between coverage, connectivity and cost.

For example, Dhillon *et al.* (2002) and So and Ye (2005) present algorithms to improve the deployment coverage. Both papers do not consider deployment connectivity and the according trade-offs. Wang *et al.* (2003) present the integration of communication and sensing coverage whereas Jourdan (2006) looks at coverage and lifetime. Both works use communication models which are limited to a simplistic homogeneous Euclidean distance model. Bai *et al.* (2006) prove the asymptotic optimality of a stripe-based deployment pattern for different ratios of sensing range to communication range. The latter approaches of Wang *et al.* (2003), Jourdan (2006), and Bai *et al.* (2006) are based on simplifying assumptions, as discussed by Kotz *et al.* (2004). Rajagopalan *et al.* (2005) use a more realistic communication model and, in addition, investigate the trade-offs with respect to energy consumption. However, only points on a spatial grid are considered as possible node positions.

None of this work considers the complex trade-off between reliability of communication and deployment costs. To the best of our knowledge, only Krause *et al.* (2006) consider both coverage and communication in a realistic scenario. The authors present a polynomial-time, data-driven algorithm using non-parametric probabilistic models called Gaussian Processes. Since their work requires sensor and link quality data collected at an initial deployment, the work of Krause *et al.* (2006) complements the present study, as we can determine an optimized deployment without any preceding data collection.

3 Problem Formulation

In this study, we consider the problem of how to distribute wireless sensor nodes in order to cover a certain area with as few nodes as possible but still provide reliable communication paths from each node to a data sink. Before we define the considered optimization criteria, we briefly describe the underlying model.

3.1 Model Description

The considered model is divided into two parts, an environment model and a model for the sensor nodes. The environment is represented by a data sink to which all the sensor readings need to be communicated and an area of interest which is to be monitored. This area of interest is outlined by a polygon and represented by a set of points of interest. We regard the area of interest as covered by sensors if every point of interest lies within the sensing range of at least one node. Note, that the proposed formulation explicitly allows sensor nodes outside the region of interest, although they only contribute to the enhancement of communication paths. The sensor nodes in turn are characterized by their position, a sensor range (here assumed to be circular), and their communication properties, i.e., their transmission probability is given by a radio function depending on the distance between transmitting and receiving node. A detailed description of both, the environment model and the sensor node model, is given by Woehrle *et al.* (2007)[1].

3.2 Optimization Criteria

To determine the quality of a given placement of sensor nodes, we propose the following objectives that have to be minimized:

[1] In contrast to Woehrle *et al.* (2007), we use the parameters $d_0 = 10m$, $P_t = 0dBm$, $\sigma = 4.0$, $\eta = 4.0$, and $P_n = -115dBm$ here.

3.2.1 Sensor Cost

Each sensor node that has to be placed causes costs, i.e., for production, deployment, and maintenance. Since one is interested in a cost-effective solution, the first optimization criterion is to minimize these costs and thereby the number of nodes. In a first approach, a cost of '1' is associated with each node. Therefore, we take the number of used nodes n as the first optimization criterion:

$$f_1 = n \tag{1}$$

3.2.2 Transmission Failure Probability

The sensor readings need to be continuously communicated from the nodes to the data sink. Thus, each of the nodes needs a reliable communication path to the data sink; if the sink lies outside of the radio range of a specific node, its communication path contains intermediate nodes which forward the message to the sink. Since wireless communication is susceptible to communication failures between nodes, e.g., due to interferences or node failures, not only the reliabilities of the best communication paths are necessary to be optimized but redundant transmission paths of high reliability as well. Instead of maximizing the connection reliability, here we consider the dual criterion of minimizing the transmission failure probability:

$$f_2 = \frac{1}{W} \cdot \sum_{j=1}^{N_{red}} w_j \cdot (1 - p_{worst,j}) \tag{2}$$

$$\textit{with } W = \sum_{j=1}^{N_{red}} w_j$$

Equation 2 scores the difference between the worst transmission path $p_{worst,j}$ on redundancy level j to an optimal path with transmission probability 1. Therefore, minimizing this criterion ensures that there is a preference for node placements resulting in high transmission reliabilities; we explicitly allow to assign different weights w_j to connections on different redundancy levels j. In turn, f_2 is normalized with the sum of these weights W.

The path reliabilities of the N_{red} most reliable paths between all nodes i and the sink are computed as follows: For each node i, we compute the most reliable path to the sink and store its corresponding reliability $p_{i,1}$, using Dijkstra's algorithm. Afterwards, we delete all nodes of this path except source and sink, and iteratively repeat this procedure until N_{red} paths are found or no longer a path exists (if less than N_{red} paths are found, all missing paths are assigned a probability of zero).

Algorithm 1 Variable Length Representation MOEA

```
initialize population P
set generation counter g = 0
set maximum number of generations G
while g ≤ G do
  M ← matingSelection(P)
  for each pair m1, m2 ∈ M do
    with probability p_c, create two offspring o1, o2 with 2D crossover
    else let o1 = m1 and o2 = m2
    draw a binary random number r distributed according to the ratio r_mut
    if r = 0 then
      V ← V ∪ voronoiMutation(o1) ∪ voronoiMutation(o2)
    else
      V ← V ∪ gaussianMutation(o1) ∪ gaussianMutation(o2)
    end if
  end for
  P ← environmentalSelection(P, V)
  g = g + 1
end while
```

4 An Evolutionary Multiobjective Algorithm with Variable Length Representations

The focus of this paper is to show the benefits of MOEAs for decision making with respect to WSN deployment. We use an off-the-shelf MOEA, namely IBEA by Zitzler and Künzli (2004), as it is provided in the PISA framework of Bleuler *et al.* (2003) and adapt the initialization and variation operators to the new search space. The resulting algorithm is sketched in Algorithm 1.

4.1 Representation

An individual represents an entire wireless sensor network as a set of sensor nodes and their positions. More precisely, an individual stores the node's plain x-y-positions as a real-valued vector with x- and y-positions alternating. Since the number of nodes is one of the optimization criteria, we explicitly allow vectors of variable length, i.e., sensor networks with a varying number of nodes. Since we assume that all sensor nodes are homogeneous, we do not explicitly have to include properties of the node model into the representation.

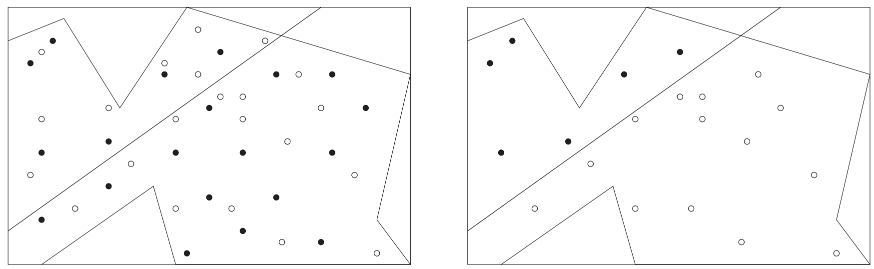

Fig. 1 Illustration of the crossover: (left) sensor node positions of the two parents (filled/empty); (right) first child, containing all positions from the first parent lying above the line and all positions from the second parent lying below the line.

4.2 Initialization

For each of the μ initial individuals, the number of sensor nodes n is randomly drawn from a Poisson distribution with mean

$$\lambda = 3 \cdot \frac{\text{area of polygon}}{\pi(\text{sensing range})^2}.$$

This ensures that initial points have enough sensor nodes to cover the region of interest. The n nodes are successively placed within the region of interest in the following way: The polygon defining the area of interest is first Delaunay triangulated. Then for each node to be placed within the polygon, a triangle is chosen randomly with a probability proportional to its area. Within this triangle, the node's position $\mathbf{x} \in \mathbb{R}^2$ is chosen uniformly according to the formula $\mathbf{x} = (1-\alpha)\mathbf{a} + \alpha(1-\beta)\mathbf{b} + \alpha\beta\mathbf{c}$ as proposed by Grimme (2005) (p. 79ff.) where $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^2$ are the triangle's vertices and α and β are chosen uniformly in the interval $[0, 1]$.

4.3 Crossover

The design of the crossover operator used in this work follows ideas earlier proposed by Schoenauer (1996) for Voronoi representations and by Zdarsky *et al.* (2005) for discrete search spaces. To create two offsprings from two parent individuals, we introduce a line intersecting the area of interest and take all sensor node positions from the first parent which reside on one side of the line and those positions from the second parent that lie on the other side of the line and vice versa to create two children (see Fig. 1). In effect, the number of nodes in the offsprings is not necessarily the same as in the parents. The intersecting line is placed randomly.

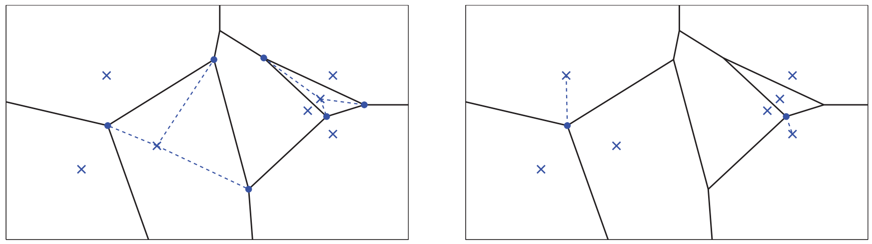

Fig. 2 Illustration of the Voronoi mutation: crosses denote sensor nodes, circles denote Voronoi vertices. The relative probabilities to delete (left) and add (right) a sensor node are depicted exemplary: on the left, the point to the right is more likely to be removed; on the right, the point to the left is more likely to be added.

4.4 Voronoi Mutation

For variable length individuals, we suggest a new mutation operator that only adds or removes points from an individual: For a network containing n sensor nodes, a number $n_{Voronoi}$ is drawn from a normal distribution $\mathcal{N}(0, 0.05n)$. This number is rounded to the next integer and according to its sign, nodes are removed (negative sign) or added (positive sign). The $|n_{Voronoi}|$ nodes are removed or added successively with the help of a Voronoi diagram on the current node set: The probability to remove a point is anti-proportional to the average Euclidean distance to all finite vertices of its corresponding Voronoi facet. The position where to add a new point always corresponds to a vertex of the Voronoi diagram: the probability to choose a vertex is proportional to its minimal Euclidean distance to all sensor nodes. Figure 2 illustrates this operator.

4.5 Gaussian Position Mutation

In case of a Gaussian position mutation, every node has a probability of 0.1 to be moved. If it is moved, a bivariate Gaussian distributed random vector with mean (0,0) and standard deviation $\sigma_{\text{mut}} = (\sigma_{\text{mut},x}, \sigma_{\text{mut},y})$[2] is added to the node's old position.

[2] To get a general operator, the σ_{mut}-values are adapted to the size of the polygon. To this end, we choose $\sigma_{\text{mut},x} = c_{\text{mut}} \cdot X/2$ where X is the width of the enclosing rectangle of the area of interest and $c_{\text{mut}} = 0.05$ is constant. The value of $\sigma_{\text{mut},y}$ is chosen similarly with respect to Y, the height of the enclosing rectangle.

5 Results and Discussion

To show the applicability of the proposed MOEA, we ran the algorithm with different parameter sets on a specific area of interest, shown in Fig. 3. Each optimization run took on average about $16h$ on a two chip dual core AMD Opteron $2.6GHz$ 64-bit machine with $8GB$ RAM. In total, we tested 25 different parameter sets, each with 22 independent runs. Using a fixed population size of $\mu = 50$ and evolving $G = 300$ generations, we tested five different crossover probabilities $\kappa \in \{0.0, 0.2, 0.5, 0.8, 1.0\}$ and five different ratios between Voronoi mutation and the Gaussian position mutation of $\rho \in \{0{:}1, 1{:}3, 1{:}1, 3{:}1, 1{:}0\}$. When comparing the hypervolume indicator values[3] for all runs, it turned out that the ratio between Voronoi mutation and Gaussian position mutation has a great influence on the results: using no Voronoi mutations resulted in a bad hypervolume while leaving out Gaussian position mutation had no clear influence on the results. The negative influence of skipping Voronoi mutation was intensified when no crossover was used, indicating that either crossover or Voronoi mutation are necessary to effectively variate the number of nodes used for the deployments. While crossover only has an influence on the number of nodes by either a beneficial cutting line or a differing number of nodes in the considered individuals, Voronoi mutation has a direct effect. Therefore, the loss in hypervolume is more drastic when leaving out Voronoi mutation than when omitting crossover. Overall, the runs using a decent amount of Voronoi mutation and crossover identified good solutions (see Fig. 4). As already indicated by the hypervolume indicator values, the identified Pareto front approximations showed a good coverage of the objective space (see Fig. 5).

[3] For the computation of the hypervolume indicator, we normalized the number of nodes with the maximal number of nodes occurring during the simulations. As reference point, $(1.01, 1.01)$ was chosen; resulting in a maximal indicator value of ≈ 1.02.

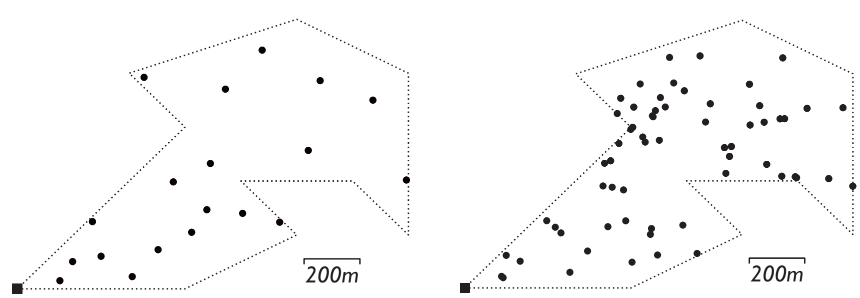

Fig. 3 Two examples for evolved node placements for the same area of interest outlined by a polygon. Nodes are marked by circles and the sink is indicated by a square. Both solutions are part of the respective Pareto front approximations found with the parameters $\kappa = 1.0$ and $\rho = 1{:}3$. The left solution corresponds to an objective vector of $(f_1, f_2) = (18, 0.5827)$ while the network on the right has objective values $(f_1, f_2) = (61, 0.0016)$.

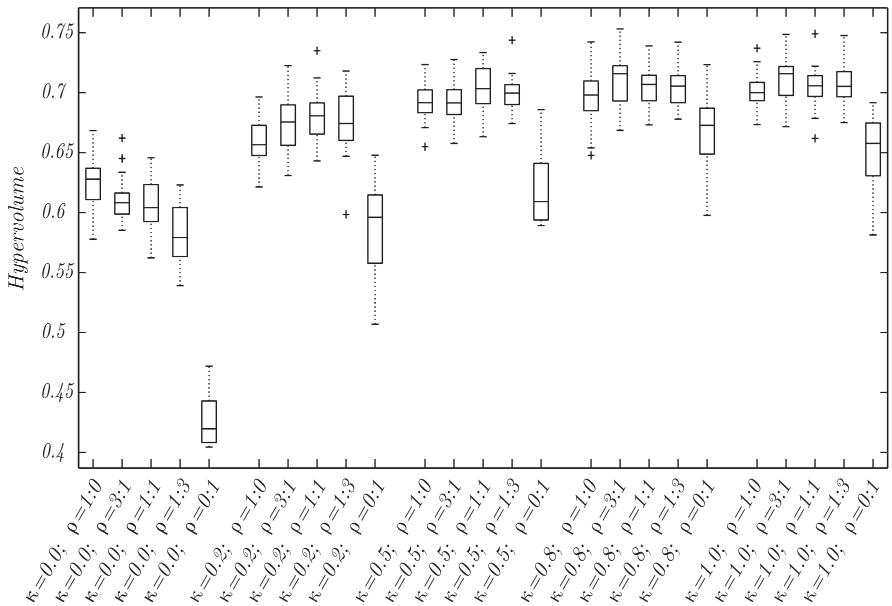

Fig. 4 Boxplot showing the results of parameter testing: each of the boxes summarizes the results of 22 independent optimization runs for the tested 25 parameter combinations. Here, κ refers to the crossover rate and ρ refers to the ratio between Voronoi mutation and Gaussian position mutation.

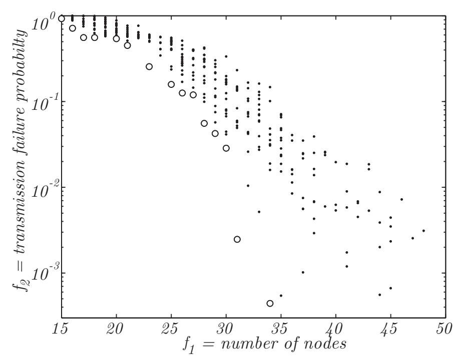

Fig. 5 Non-dominated solutions in population after 300 generations for all 22 runs with crossover probability $\kappa = 1.0$ and mutation ratio $\rho = 3 : 1$. To improve readability, the y-axis is plotted in log-scale. Solutions non-dominated over all runs are depicted as circles.

All in all, the proposed MOEA showed its capabilities in identifying a broad range of trade-offs between number of nodes and transmission reliability (see Fig. 3 and Fig. 5). Although the algorithm needs to be tested on further deployment scenarios in the future, the optimization results on the test scenario indicate that the MOEA is able to provide valuable support for a human decision maker. A specific real-world application where this support would be valuable is the design of a WSNs for fire detection: for this type of WSNs it is required to guarantee reliable data transmission on at least two redundant paths to ensure that no fire alarm gets lost. On the other hand, the number of nodes affects recurring costs: maintenance is expensive, especially when considering that some of the nodes may be placed in locations that are difficult to access. Without the knowledge about these trade-offs, it will be difficult for a decision maker to formulate an appropriate utility function as in classical optimization; this is where the proposed MOEA can help by providing good trade-off solutions that help the decision maker to decide on the most desirable solution.

6 Conclusion and Outlook

We applied a MOEA to the deployment of WSNs, considering the two objectives network cost and transmission reliability. Building on an off-the-shelf selector, we proposed new variation operators suitable for the considered problem. We gave a proof of concept that the proposed MOEA identifies a broad range of trade-offs between the considered criteria. Therefore, the MOEA can be used, e.g., to investigate the influence of positioning and the number of available data sinks on the resulting network topology—an upcoming research topic in WSNs.

In the future, we plan to adapt our approach to a range of real world setups. In the presented study, we considered link qualities depending on the sending node's distance to a receiving node, an assumption only valid without obstruction. To be able to model indoor scenarios accurately, we plan to extend the radio model by explicitly incorporating the effect of barriers. With respect to the environment model, we plan to introduce heterogeneous sensor coverage: for a monitored area, specific regions are often more important than others. This can be integrated into the model by requiring that the points representing the area of interest need to be covered by a set of sensors.

7 Acknowledgments

Matthias Woehrle and Dimo Brockhoff have been supported by the SNF under grant numbers 5005-67322 and 112079. Tim Hohm has been supported by the European Commission under the Marie Curie RTN SY-STEM, Project 5336.

References

Bai, X., Kuma, S., Xua, D., Yun, Z., and La, T. H. (2006). Deploying Wireless Sensors to Achieve Both Coverage and Connectivity. In *Symposium on Mobile ad hoc networking and computing (MobiHoc 2006)*, pages 131–142. ACM Press.

Bleuler, S., Laumanns, M., Thiele, L., and Zitzler, E. (2003). PISA—A Platform and Programming Language Independent Interface for Search Algorithms. In *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 494–508.

Dhillon, S., Chakrabarty, K., and Iyengar, S. (2002). Sensor Placement for Grid Coverage under Imprecise Detections. In *Conference on Information Fusion*, pages 1581– 1587.

Grimme, C. (2005). Räuber-Beute-Systeme für die Mehrkriterielle Optimierung. Internal Report of the Systems Analysis Research Group SYS–5/05, Dortmund University, Computer Science Section.

Jourdan, D. B. (2006). *Wireless Sensor Network Planning with Application to UWB Localization in GPS-Denied Environments*. Ph.D. thesis, Massachusetts Institute of Technology.

Kotz, D., Newport, C., Gray, R., Liu, J., Yuan, Y., and Elliott, C. (2004). Experimental evaluation of wireless simulation assumptions. In *Int'l Workshop Modeling Analysis and Simulation of Wireless and Mobile Systems (MSWiM 04)*, pages 78–82. ACM Press.

Krause, A., Guestrin, C., Gupta, A., and Kleinberg, J. (2006). Near-optimal Sensor Placements: Maximizing Information While Minimizing Communication Cost. In *Conference on Information Processing Sensor Networks (IPSN 2006)*, pages 2–10. ACM Press.

Mainwaring, A., Polastre, J., Szewczyk, R., Culler, D., and Anderson, J. (2002). Wireless Sensor Networks for Habitat Monitoring. In *Workshop on Wireless Sensor Networks and Application (WSNA 2002)*, pages 88–97.

Meier, A., Beutel, J., Lim, R., and Thiele, L. (2007). Design of a High-Reliability Low-Power Status Monitoring Protocol. In *Conference on Networked Sensing Systems (INSS 2007)*, pages 2–9.

Rajagopalan, R., Varshney, P. K., Mohan, C. K., and Mehrotra, K. G. (2005). Sensor Placement for Energy Efficient Target Detection in Wireless Sensor Networks: A Multi-objective Optimization Approach. In *Conference on Information Sciences and Systems*.

Schoenauer, M. (1996). Shape Representations and Evolution Schemes. In *Conference on Evolutionary Programming*, pages 121–129. MIT Press.

So, A. M.-C. and Ye, Y. (2005). On Solving Coverage Problems in a Wireless Sensor Network Using Voronoi Diagrams. In *Workshop on Internet and Network Economics (WINE 2005)*, pages 584–593.

Wang, X., Xing, G., Zhang, Y., Lu, C., Pless, R., and Gill, C. (2003). Integrated Coverage and Connectivity Configuration in Wireless Sensor Networks. In *Conference on Embedded Networked Sensor Systems (SenSys 2003)*, pages 28–39. ACM Press.

Woehrle, M., Brockhoff, D., and Hohm, T. (2007). A New Model for Deployment Coverage and Connectivity of Wireless Sensor Networks. Technical Report 278, Computer Engineering and Networks Lab, ETH Zurich, 8092 Zurich, Switzerland.

Xu, N., Rangwala, S., Chintalapudi, K., Ganesan, D., Broad, A., Govindan, R., and Estrin, D. (2004). A Wireless Sensor Network For Structural Monitoring. In *Conference on Embedded Networked Sensor Systems (SenSys 2004)*, pages 13–24.

Zdarsky, F. A., Martinovic, I., and Schmitt, J. B. (2005). On Lower Bounds for MAC Layer Contention in CSMA/CA-Based Wireless Networks. In *Workshop on Discrete Algorithms and Methods for MOBILE Computing and Communications*, pages 8–16. ACM Press.

Zitzler, E. and Künzli, S. (2004). Indicator-Based Selection in Multiobjective Search. In *Conference on Parallel Problem Solving from Nature (PPSN VIII)*, volume 3242 of *LNCS*, pages 832–842. Springer.