Layout Optimization for a Wireless Sensor Network Using a Multi-Objective Genetic Algorithm

Damien B. Jourdan, Olivier L. de Weck
Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, MA 02139, USA

***Abstract*-This paper examines the optimization of wireless sensor network layouts. To transmit their data to the base, all the sensors are required to be connected to a high-energy communication node, which serves as a relay from the ground to a satellite or to a high-altitude aircraft. The sensors are assumed to have a fixed communication and a fixed sensing range, which *can significantly vary depending on the type of sensing* performed. This simple framework serves to benchmark a Multi Objective Genetic Algorithm (MOGA) for the sensor placement, where the two competing objectives considered are the total sensor coverage and the lifetime of the network. The MOGA is then used to show that, for different relative sensing ranges, two fundamentally different types of layouts are obtained: one with the sensors closely packed together, the other with the sensors organized in a hub-and-spoke manner. The ratio of sensing to communication range is shown to be the discriminating factor.**

I. Introduction

RECENT military operations demonstrated the limitations of surveillance missions performed by high-altitude platforms (UAV, U2, satellite) even when equipped with state of the art sensors. Most of these limitations are inherent to this type of long-distance surveillance and cannot be resolved with any improvement in the onboard-sensor technology.

In order to gain a clear understanding of the situation on the ground, it is becoming vital to observe from close range, using remote sensing devices placed in the area of interest, e.g. Wireless Sensor Networks (WSN). Since these missions will be performed in hostile areas, the placement of such sensors needs to be done without human personnel involved, e.g. via aerial deployment from an aircraft. Once the sensors are deployed on the ground, their data is transmitted back to the home base to provide the necessary situational awareness.

The deployed units (the wireless sensors, called sensors in the following) fulfill two fundamental functions: sensing and communicating. The sensing can be of different types (seismic, acoustic, chemical, optical, etc.), and the communication is performed wirelessly. However, the small size and energy storage capacity of the sensors prevent them from relaying their gathered information directly to the base. It is therefore necessary that they transmit their data to a high-energy communication node (HECN) able to provide the transmission relay to an aircraft or a satellite. All sensors must be able to transmit their data to this node, either directly or via hops, using nearby sensors as communication relays (Fig. 1). The aircraft carrying the sensors has a limited payload, so it is impossible to randomly drop thousands of sensors over the area of interest, hoping the necessary communication connectivity would arise by chance; thus, the mission must be performed with a fixed maximum number of sensors. In addition, the terrain will often be complex (urban, wooded or mountainous environments) and will influence both sensing and communication abilities. Finally, the mode of deployment (airdrops) introduces uncertainty in the final sensor positions and this must be taken into account. These criteria motivate the creation of a planning system that automates the WSN deployment process and incorporates those requirements.

This paper addresses the automated placement of sensors, using an idealized model for the two characteristics of the sensors; they can communicate with one another if they are within a fixed distance R_{COMM}, and they can sense anything within their sensing radius R_{Sensor}. All sensors in the WSN are assumed to be identical. They also must be connected to the HECN in order to transmit their data to the base. An optimization technique (a Multi-Objective Genetic Algorithm (MOGA)) is used with this simplified model to provide the end-user with a set of Pareto-optimal (non-dominated) network designs with coverage and lifetime of the network as the two objectives to be maximized. We also use the MOGA to demonstrate the fundamental relationship between the ratio R_{Sensor}/R_{COMM} and the final Pareto-optimal layouts. This is important in practice since, depending on the type of sensor deployed, this ratio will vary greatly. For example, a seismic sensor may have R_{Sensor} much greater than R_{COMM}, while the opposite may be true for an acoustic sensor. We show that this factor will fundamentally affect the optimal layout of the network.

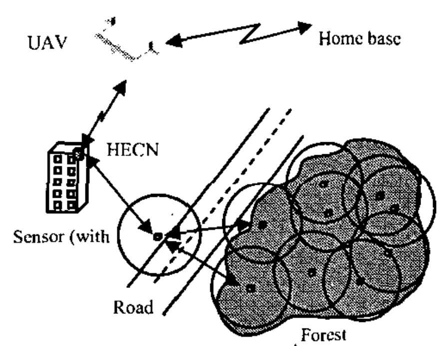

Figure 1. Example of the use of a WSN to monitor a region, with the High Energy Communication Node (HECN) placed on top of a building.

This research is funded by the Charles Stark Draper Laboratory through the MIT/Draper Technology Development Partnership Program.

II. LITERATURE REVIEW

Several bodies of research directly deal with the placement of nodes in network design.

From the early 1990's to a few years ago, a large body of research was devoted to the Base Station (BS) location problem for cellular phone networks. At that time the problem was to find the optimal location of BS (transmitters) in order to satisfactorily cover subscribers. Although this problem differs in many aspects from the sensor network planning problem (notably because in WSN the sensors ("BS") also need to communicate with each other (connectivity)), it is insightful to review the methods used. These range from Dynamic Programming [1], to Genetic Algorithms [2], [3] and Tabu Search [4]. Virtually every type of optimization technique was tested on this problem, many of which dealt with multiple objectives (though often blended into a single objective function, except in [3] which uses Pareto optimality) while using non-trivial communication models taking the terrain into account.

The BS location problem is part of the larger area of Facility Location in Operations Research [5]. Here a set of demand points must be covered by a set of facilities (which corresponds in WSN to covering an area with a set of sensors). The goal is to locate these facilities so as to optimize a certain objective (e.g. minimize the total distance of demand points to their closest facility). A classic example close to the WSN problem is the Maximal Covering Location Problem (MCLP) [6], [7], where as many demand points as possible must be covered with p sensors of fixed radius. It is also referred to as a location-allocation problem, since each demand point must be assigned to a certain sensor. Again in all these discussions, the main difference with WSN is that the nodes are not required to be connected. Another problem of interest is the Facility Location-Network Design problem, where facilities positions need to be determined (just as in MCLP) and the network connecting these facilities must also be optimized. Unfortunately, in WSN design it is impossible to decouple sensor placement and network design, since the location of the sensors determines the network topology.

The past three years have seen a rising interest in sensor network planning, focusing mostly on optimizing the location of the sensors in order to maximize their collective coverage (a problem almost identical to the BS location problem). Several techniques were used, but the research on BS location is never mentioned. Chakrabarty [8] used Integer Programming, while Bulusu [9], Dhillon [10] and Howard [11], [12] devised different greedy heuristic rules to incrementally deploy the sensors. Zou [13] adapted Virtual Force Methods (often used for the deployment of robots [12]) for the sensor deployment.

As was mentioned before, current work on WSN mainly focuses on the maximization of sensing coverage, with little or no attention given to the communication requirement between sensors. Meguerdichian [14] assumes that the communication radius of the sensors will be much larger than the sensing radius, so that connectivity will arise naturally. But this assumption is unrealistic for two reasons. First there exist sensors where the sensing range is of the same order or larger than the communication range (e.g. seismic sensors), so that maximizing the coverage without caring about the communication range will result in a disconnected network. Second, if the areas to be covered are disjoint, the network will be partitioned. In addition, in our WSN model the sensors must be connected not only to each other, but also to the HECN. Therefore the communication connectivity requirement cannot be trivialized, and both aspects of the sensors (sensing and communication) must be taken into account for the network planning. Also, only a single objective is considered (almost always coverage), whereas it seems other considerations are also of vital practical importance in the choice of the network layout (lifetime, robustness to node failure, etc.).

Current work on WSN does not deal with multiple objectives and pays little attention to the communication connectivity requirement, essential for the data relay. This work attempts to start addressing these gaps.

III. MODELING

A. WSN Modeling

The area considered is a flat square surface where sensor nodes can monitor anything within R_{Sensor}, and where they can communicate with any other node located within R_{COMM}. The HECN, with which every sensor must communicate (either directly or via hops through nearby sensors), is placed in the center of the area. This assumption is for convenience and does not prevent generalizability. Each sensor initially has the same energy available in its battery (assumed to be the sole source of energy), and we assume it decreases by one arbitrary unit for every data transmission.

The design variables are the horizontal and vertical coordinates of the sensors. In this paper we only consider WSN with a predetermined number n of sensors, so that the vector of design variables $\underline{DV}$ is of constant size 2n.

$$\underline{DV} = [x_1 \quad y_1 \quad \dots \quad x_n \quad y_n] \tag{1}$$

Similar layouts rotated about the origin will have similar objectives, but different design vectors. It may therefore be legitimate to talk about a global optimal *layout*, but not about an optimal *design vector* of coordinates.

B. Objectives Calculation

Two objectives are considered, coverage and lifetime of the network. Only the sensors that are connected to the HECN (i.e. those that can relay their data to the HECN) are taken into account in the calculation of these objectives. The coverage is equal to the area of the union of the disks of radii R_{Sensor} centered at each connected sensor, normalized by the total area.

0-7803-8255-2/04/$20.00 ©2004 IEEE.

$$Coverage = \left[\bigcup_{i=1}^{n} R^2_{Sensor}(x_i, y_i) \right] / Area \qquad (2)$$

The lifetime is defined as the ratio of the time to first sensor failure (no more energy) and the maximum lifetime of a sensor.

$$Lifetime = \min_{i=1...n}(T_{failure,i}) / T_{\max} \qquad (3)$$

It is assumed that all sensors gather data at the same time and then relay it to the HECN (we call this a sensing cycle). In order to do so, the data may need to be relayed by several sensors before it reaches the HECN (hops). Therefore, at every sensing cycle the sensor nodes need to transmit their own data and possibly the data from other sensors, in which case they act as communication relay. Thus at each sensing cycle, the data from every sensor needs to be routed to the HECN in a way that will maximize the remaining energy in the nodes. In order to find these routes, the outgoing edges of every node are weighted by the inverse of the node's remaining energy, and then the Dijkstra algorithm is used to find the route of minimum weight [15]. Repeating this calculation until the energy of at least one node is depleted gives the maximum number of sensing cycles a particular WSN layout can perform. This number is then normalized by T_{max}, the maximum number of sensing cycles possible (obtained when all sensors are directly connected to the HECN, so that none act as a communication relay).

These two objectives are competing. On the one hand the coverage objective will desire "spread out" network layouts, where sensors are as far apart from each other as possible in order to minimize the overlap between sensing disks. This implies a large number of relay transmissions for sensors communicating directly with the HECN, so that their failure will happen sooner due to the consumption of all their energy – the network lifetime will then be small. On the other hand, in order to get a lifetime of 1 all the sensors must communicate directly to the HECN, so that their energy is used only for their own data transmission. This implies a clustered configuration around the HECN with a lot of overlap between sensing disks, yielding a poor coverage value.

IV. MOGA RESULTS

The design space of the WSN layout optimization is highly non-linear, due notably to the binary nature of the communication connectivity between sensors; moving a sensor by a small amount can cause large changes in both objectives (e.g. if it becomes disconnected). Genetic Algorithm (GA) was chosen to perform the optimization since it has proven to work well with non-linear objectives. The MOGA is aimed at providing the end-user with a set of Pareto-optimal layouts from which to choose from. This is interesting because it shows for example how much lifetime must be given up in order to gain some coverage (trade-off). Because it explores the whole search space, the MOGA will find Pareto-optimum network topologies (structure of the network). Local search methods can then be used to refine these "raw" results by fine-tuning the position of each sensor.

The GA starts with a "parent" population of N network layouts randomly generated, where each individual is represented by its chromosome, the vector $\underline{DV}$ as in (1). Each parent individual is randomly mated with another to produce two children (crossover). The children are then mutated with probability m, where each coordinate of the design vector representing each child is modified randomly with probability m. The coverage and lifetime of the children are then calculated, and a fitness value is assigned to every parent and child. This fitness is based on the Pareto dominance developed by Fonseca and Fleming [16], and is proportional for each individual to the number of individuals that dominate it (in the Pareto sense). The N individuals with best fitness are then passed on to the next generation, and the process continues until the maximum number of generations is reached. In the end a well-populated Pareto Front is obtained, as shown in Figure 2. These results were obtained for a WSN of 10 sensors, with R_{COMM} and R_{Sensor} both equal to 2 (in arbitrary units). The area is a 10 by 10 square. The Pareto Front obtained is shown and two Pareto-optimal layouts are plotted

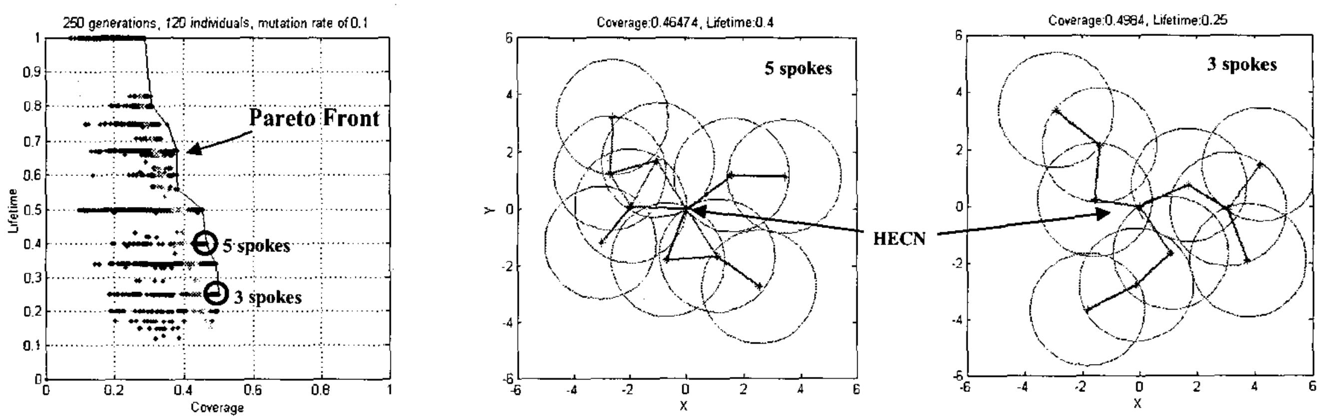

Figure 2. MOGA results for a WSN with 10 sensors and R_{Sensor}=R_{COMM}=2. The Pareto Front is shown on the left, and two Pareto-optimal layouts are shown, with respectively 5 and 3 spokes stemming out of the HECN

to illustrate the variety of designs available.

V. ANALYSIS OF LAYOUT WITH BEST COVERAGE

In this section we investigate the influence of the ratio R_{Sensor}/R_{COMM} on the Pareto-optimal layouts, focusing on the layouts at the two ends of the Pareto Front. The layout with best lifetime (upper end of the PF) is always the same, irrespective of the value of R_{Sensor}/R_{COMM}, with all the sensors clustered around the HECN. As mentioned before, in this configuration none of the sensors has to act as communication relay and they can devote all their energy to transmitting their own data, yielding a network lifetime of 1.

A. *Change of Optimal Layout depending on R_{Sensor}/R_{COMM}*

It is more interesting to look at the layout with best coverage (bottom end of the PF). We show in Figure 3 the layouts of best coverage for R_{Sensor}/R_{COMM} equal to 1/2, 1 and 2. In the case where the ratio is equal to 1/2 (Fig. 3a), the sensors are arranged in a beehive fashion around the HECN, whereas when it is equal to 1 and 2 (Fig. 3b and 3c) the sensors form a hub-and-spoke configuration, with respectively 3 and 2 spokes stemming from the HECN.

It is interesting to note that in the first case (Fig. 3a), hub-and-spoke layouts are not Pareto-optimal. For a given number of sensors, the maximum coverage will be obtained when, if possible, there is no overlap between the sensing disks. Viewing the sensors as marbles with a radius equal to the sensing radius, if we pack them as tightly as possible, each marble touches 6 neighbors (except for the peripheral ones). This beehive configuration ensures a maximum coverage, while providing the maximum number of neighbors for each sensor. Thus as long as $R_{COMM} \geq 2R_{Sensor}$, or $R_{Sensor} / R_{COMM} \leq 1/2$, this configuration is the Pareto-optimal layout with maximum coverage.

However, once $R_{Sensor} > R_{COMM} / 2$, if the sensors remain in this configuration they cannot communicate anymore (they are all outside of their neighbors' communication range). Therefore overlap between the sensing disks becomes necessary, which explains the change in optimal layout observed in Fig. 3b and Fig. 3c, from a beehive to a hub-and-spoke (which minimizes the overlap). The number of spokes stemming from the HECN also seems to vary as the ratio increases. The ratio of 1/2 is critical since the communication circles must encompass the center of at least one neighboring node, while for sensing, nodal circles are not required to overlap.

B. *Analysis of the Hub-and-spoke Layout*

Given a ratio R_{Sensor}/R_{COMM} greater than 1/2 (for smaller values the optimal layout is a beehive), we would like to determine the optimal number of spokes stemming from the HECN so that the coverage is maximized. For this purpose we assume the area to be the infinite plane. First we note that if n sensors are to be placed, the layout is completely determined by the number of sensors at the root (i.e. directly connected to the HECN); for example if we have 10 sensors to place and there are 3 of them at the root, then the maximum coverage will be obtained by attaching the remaining 7 sensors to these 3, forming a hub-and-spoke with 3 spokes. We determine the optimum number of sensors at the root so as to maximize the coverage of the WSN. When R_{Sensor}/R_{COMM} equals 1/2, 6 sensors can be placed at the root without overlap between the sensing disks (beehive). As this ratio increases, fewer sensors can be placed at the root if we are to avoid any overlap there. However, it might be beneficial to allow overlap at the root in order to have less overlap in the spokes, as illustrated in Fig. 4 for 3 sensors with R_{Sensor}/R_{COMM} equal to 1.

For a given number of sensors, we should only be willing to give up coverage at the root if it results in more overall coverage. Let n_r be the number of sensors at the root, Cn_r the coverage provided by the n_r sensors at the root, and C_{spoke} the added coverage provided when a sensor is connected to one of the sensors at the root (i.e. when it does not create a new spoke, but just adds onto an existing one). Assuming we only have n_r sensors at the root, we have two choices for the placement of an additional sensor: we either add it to the root, or we connect it to a sensor belonging to the root. The total

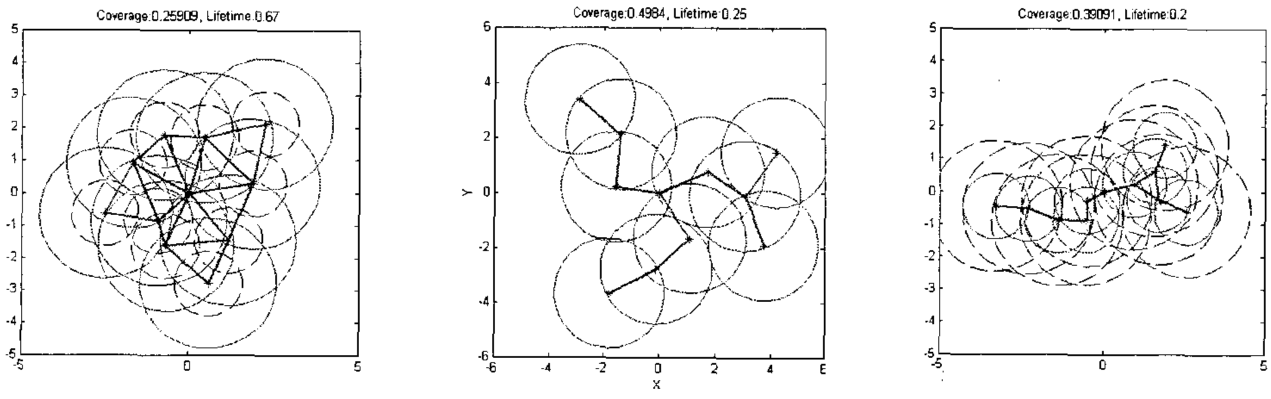

Figure 3. Layouts with best coverage for a WSN of 10 sensors and a ratio R_{Sensor}/R_{COMM} of 1/2 (a, left), 1 (b, middle) and 2 (c, right). The sensing circles are dashed, the communication circles solid, and the links between connected sensors are shown

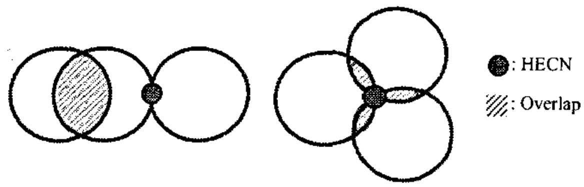

Figure 4. A layout with 2 spokes (a, left) generates more overlap than one with 3 spokes (b, right) in this example with 3 sensors.

coverage is Cn_{r+1} in the former case, and Cn_r+C_{spoke} in the latter (see Fig. 4 for an illustration where n_r is 2). We choose the option yielding the most coverage (the "root" option in Fig. 4a). Note that once Cn_{r+1} is smaller than Cn_r+C_{spoke}, sensors will always be added to the spokes (i.e. placing additional sensors at the root will provide less and less new coverage because they will overlap more and more with the sensors already present there). Fig. 5 shows the result of this numerical analysis repeated over different ratios. The number of spokes starts from 6 when R_{Sensor}/R_{COMM} is 1/2, and then decreases rapidly to 3. When the sensing radius is about twice the communication radius, the optimal layout only has 2 spokes stemming from the HECN; essentially the network degenerates into a linear array. These results, valid for any number of sensors, confirm the optimality of the layouts obtained by the MOGA; Fig. 3b has 3 spokes for a ratio of 1, and Fig. 3c 2 spokes for a ratio of 2. The GA thus has revealed rules of "optimal" WSN's in terms of coverage and lifetime for a given ratio R_{Sensor}/R_{COMM}.

VI. Conclusions and Future Work

In this paper we presented a Multi-Objective Genetic Algorithm to optimize the layout of WSN. The algorithm aims at maximizing the coverage and lifetime of the network, yielding a Pareto Front from which the user can choose. We also investigated the influence of the ratio between sensing range and communication range on the optimal layout with best coverage. When this ratio is below 1/2, this layout is formed of polygons and resembles a beehive, whereas for ratios above 1/2, hub-and-spoke layouts become optimal. An analysis showed how the number of spokes in the hub-and-spoke layouts varies with this ratio, from 6 to 2; the optimality of the results provided by the MOGA was thus verified in this simple idealized model. The true value of the MOGA will become apparent in more realistic conditions, e.g. with mixed sensors, uneven terrain and non-ideal sensing and communication boundary conditions.

In future work we will include in the MOGA the number of sensors as a design variable, as well as account for the uncertainty in the position of the sensors due to the mode of deployment (airdrops). Additionally, different sensing objectives, such as remote surveillance of a facility will be explored.

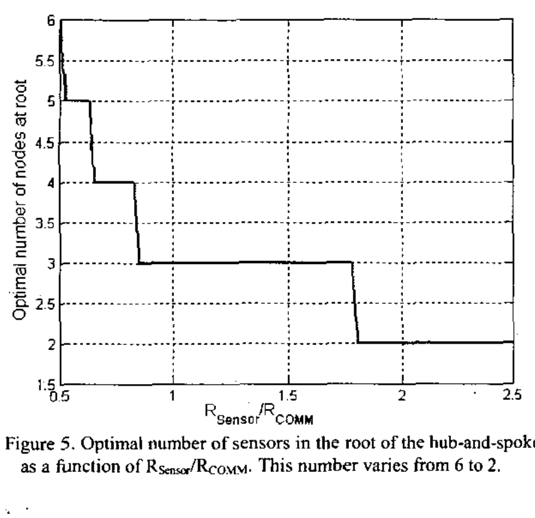

Figure 5. Optimal number of sensors in the root of the hub-and-spoke as a function of R_{Sensor}/R_{COMM}. This number varies from 6 to 2.

References

[1] R. Rose, "A smart technique for determining base-station locations in an urban environment," *Proc. IEEE Vehicular Technology Conference*, vol. 50, pp. 43-47, Jan. 2001.

[2] J. K. Han, B. S. Park, Y. S. Choi and H. K. Park, "Genetic approach with a new representation for base station placement in mobile communications," *Proc. IEEE Vehicular Technology Conference*, vol. 4, pp. 2703-2707, Oct. 2001.

[3] H. Meunier, E. Talbi, P. Reininger, "A multiobjective genetic algorithm for radio network optimization," *Proc. Congress on Evolutionary Computation*, vol. 1, pp. 317-324, July 2000.

[4] E. Amaldi, A. Capone, F. Malucelli, F. Signori, "UMTS radio planning: optimizing base station configuration," *Proc. IEEE Vehicular Technology Conference* , vol. 2, pp. 768-772, Sept. 2002.

[5] Z. Drezner, H. W. Hamacher (editors), *Facility location: applications and theory*, Springer-Verlag, Berlin, 2002.

[6] R. Church, C. ReVelle, "The maximal covering location problem," *Papers of The Regional Science Association*, 32: 101-118, 1974.

[7] A. Mehzer, A. Stulman, "The maximal covering location problem with facility placement on the entire plane," *Journal of Regional Science*, vol. 22, no. 3, 1982.

[8] K. Chakrabarty, S. S. Iyengar, H. Qi and E. Cho, "Grid coverage for surveillance and target location in distributed sensor networks," *IEEE Transactions on Computers*, vol. 51, pp. 1448-1453, Dec. 2002.

[9] N. Bulusu, J. Heidemann and D. Estrin, "Adaptative beacon placement," *Proc. Int. Conf. on Distributed Computing Systems*, pp. 489-498, Apr. 2001.

[10] S. S. Dhillon, K. Chakrabarty and S. S. Iyengar, "Sensor placement for grid coverage under imprecise detections," *Proc. Int. Conf. on Information Fusion*, vol. 2, pp. 1581-1587, Jul. 2002.

[11] A. Howard, M. J. Mataric and G. S. Sukhatme, "An incremental self-deployment algorithm for mobile sensor networks," *Autonomous Robots Special Issue on Intelligent Embedded Systems*, vol. 13, no. 2, pp. 113-126, 2002.

[12] A. Howard, M. J. Mataric and G. S. Sukhatme, "Mobile sensor network deployment using potential fields: a distributed, scalable solution to the area coverage problem," *Proc. Int. Conf. on Distributed Autonomous Robotic Systems*, pp. 299-308, 2002.

[13] Y. Zou and K. Chakrabarty, "Sensor deployment and target localization based on virtual forces," *Proc. IEEE Infocom Conference*, vol. 2, pp. 1293-1303, Apr. 2003.

[14] S. Meguerdichian, M. Potkonjak, "Low power 0/1 coverage and scheduling techniques in sensor networks," *UCLA Technical Reports 030001*, January 2003.

[15] T. H. Cormen et al., *Introduction to algorithms*, MIT Press, Cambridge, MA, 2001.

[16] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for Multiobjective optimization: formulation, discussion and generalization", *Genetic Algorithms: Proc. Fifth International Conference*, pp 416-423, Morgan Kaufmann, 1993.