



<b>Title</b>	<b>Accuracy improvement of connectivity-based sensor network localization</b>
<b>Author(s)</b>	<b>Qiao, D; Pang, GKH</b>
<b>Citation</b>	<b>The 25th IEEE Canadian Conference on Electrical &amp; Computer Engineering (CCECE 2012), Montreal, QC., 29 April-2 May 2012. In IEEE Canadian Conference on Electrical and Computer Engineering Proceedings, 2012, p. 1-5</b>
<b>Issued Date</b>	<b>2012</b>
<b>URL</b>	<b><a href="http://hdl.handle.net/10722/165266">http://hdl.handle.net/10722/165266</a></b>
<b>Rights</b>	<b>IEEE Canadian Conference on Electrical and Computer Engineering Proceedings. Copyright © IEEE.</b>

# ACCURACY IMPROVEMENT OF CONNECTIVITY-BASED SENSOR NETWORK LOCALIZATION

Dapeng Qiao, Grantham K.H. Pang  
Department of Electrical and Electronic Engineering  
The University of Hong Kong  
Pokfulam Road, Hong Kong

## ABSTRACT

The early results from connectivity-based sensor network localization suffer from disappointing accuracy. The reason is partly due to the limited information of the problem, and also the deficiencies of the algorithms. This paper proposes a two-level range/indication of connectivity between each pair of nodes, which would indicate three levels of connectivity: strong, weak or nil. Theoretically, the two-level connectivity localization problem can be modeled as a non-convex optimization problem in mathematics, which contains the convex constraints and non-convex constraints. Besides using two-level range to enrich the given information, a two-objective evolutionary algorithm is also used for searching a solution. The simulation is carried out using five different topology networks all containing 100 nodes. Simulation results have shown that better solution can be obtained by using two-level range connectivity when compared with the usual one-level range connectivity-based localization.

*Index Terms* - wireless sensor network; localization; allocation; two-range; connectivity.

## I. INTRODUCTION

Wireless sensor network usually involves hundreds of low-cost sensor nodes, which can communicate with each other to make a network. Connectivity-based sensor network localization utilizes some anchors and the information on whether any two nodes are within connection or not to localize the nodes.

In a usual formulation of the connectivity-based localization problem, it is assumed that connectivity is established if the distance between two nodes is within a pre-defined distance. The current solutions of the connectivity-based localization problem can fall into two categories. The first class of methods tries to find the number of direct connections between two nodes. In other words, the number of hops from one node to another node needs to be found. Hence, the hop count value would roughly represent the distance between two nodes. The centroid method [1], the approximate point in triangulation (APIT) [2], the multidimensional scaling-MAP (MDS-MAP) [3], DV-Hop [4] all belong to this class. The other class of methods models the connectivity-based localization problem as a constrained optimization problem. The connectivity information becomes the constraints that the optimization result must satisfy. For example, convex position estimation (CPE) [5] selects the convex constraints to formulate the problem as a convex optimization and uses semi-definite programming (SDP) to solve the problem.

Nevertheless, the accuracy of the above algorithms is very coarse. Firstly, the given information of the connectivity-based network is limited. Different from the range-based localization which obtains an estimate of the distance measure between some pairs of nodes, the connectivity-based localization only knows whether any pair of nodes are within connection or not. Another reason of the coarse accuracy is that current algorithms have ignored information on the disconnections between any two nodes when calculating for a solution. Actually, if we just consider the connections between nodes, the problem would be a convex optimization problem. However, the disconnection situations, which indicate two nodes are not within a certain range, would lead to non-convexity. Therefore, for a complete solution, the problem should be modeled as a non-convex constraint satisfaction problem if all the connectivity constraints are considered.

This paper presents a two-level range method for modeling the localization problem which is formulated as a non-convex optimization problem. A two-objective evolutionary algorithm is used to solve the optimization problem.

## II. TWO-LEVEL RANGE

### A. Setting of Two-level Range

The main idea follows from our previous work [6] in which two-level range was proposed to improve the accuracy of the current algorithms. The connectivity information is determined by the communication range, which is the maximum distance within which the two nodes can communicate. We can modify the connectivity information by changing the value of range. Figure 1 shows a node that can operate at two different ranges. Range A is assumed to be a circle with radius of  $R_A$ , and range B is assumed to be a circle with radius of  $R_B$ .

An example of part of a network with such two-level range nodes is given in Figure 2. Range B for every node is expressed by dotted line. In Fig. 2, the distance between node 1 and node 2 is less than  $R_B$ . As another example, the distance between node 1 and node 3 is less than  $R_A$ , but more than  $R_B$ . When compared with the one-range connectivity-based sensor network, the two-level range network can give us more useful information on node connectivity and would potentially provide more accurate localization results.

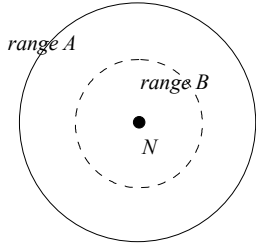


Figure 1. A node with two different ranges in a node  $N$

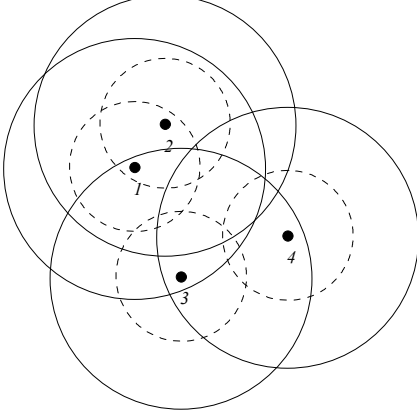


Figure 2. An example of part of a sensor network

As shown in an example in [7], the transmit power of the sensors can be controlled and modified. Hence, it is possible to arrange for nodes to operate at two different transmit power states. On the receiver side, the node device is always listening for incoming messages, regardless of transmission activity. Therefore, the transmission range between nodes would rely on the transmit power. The nodes can be operated to work on two ranges by changing the transmit power. Hence, the transmit power can have two states: one state for a longer range, and the other for a shorter range. The transmit power can also operate in these two states at different time instants in order to get the two different transmission ranges. The longer range is called range A, with radius  $R_A$ , and the shorter range is called range B with radius  $R_B$ . This two-level range configuration gives more information to sensor node localization.

The use of a two-level communication range is available in practice. The topic about controlling the range of transmission (including the transmission of omni-directional antenna) has been studied for many years [8] and there are techniques to change the range of transmission for two-level range nodes. An example of applying controllable transmission power on sensor network can be found in [7], which is a Mica2 sensor node developed by UC Berkeley. The radio part of this sensor node is a ChipCon CC1000 radio, which supports programmable transmission power levels ranging from -20dBm to +10dBm.

### B. Problem Definition

A formal definition of the two-range connectivity-based localization problem is given next. Let  $G = (V, E)$  be a given network, where  $V$  denotes the nodes of the network and  $E$  denotes the edge of the network. Let  $V$  be partitioned into two

sets:  $V_a = \{1, \dots, m\}$  of anchors, and  $V_b = \{m+1, \dots, m+n\}$  of sensors.  $E$  is also partitioned into two sets:  $E_{ab} = \{(i, j) \in E : i \in V_a, j \in V_b\}$  which are the edges between a sensor and an anchor.  $E_{bb} = \{(i, j) \in E : i, j \in V_b\}$  which are the edges between two sensors. For each anchor  $i \in V_a$ , the position  $a_i \in \mathbb{R}^2$  is assumed to be known. For each sensor  $i \in V_b$ , the position  $b_i \in \mathbb{R}^2$  is assumed to be unknown.

Let  $C_{ab} = \{(i, j, k) : i \in V_a, j \in V_b, k \in \{0, 0.5, 1\}\}$  be the connectivity information between a sensor and an anchor. Also let  $C_{bb} = \{(i, j, k) : i, j \in V_b, k \in \{0, 0.5, 1\}\}$  be the connectivity information between two sensors. The value  $k$  in  $C_{ab}$  or  $C_{bb}$  has the value 0, 0.5 or 1:

$k=0$  if there is no connection between node  $i$  and  $j$ .

$k=1$  for a strong connection between node  $i$  and  $j$ .

$k=0.5$  for a weak connection between node  $i$  and  $j$ .

Let  $a$  be a vector containing the positions of the anchors  $a = (a_i)_{i \in V_a} \in \mathbb{R}^{2m}$ . The goal of the two-range connectivity-based network localization problem is to determine the coordinates of all the sensors (unknown nodes)  $b = (b_i)_{i \in V_b} \in \mathbb{R}^{2n}$  such that  $b$  satisfies the following constraints:

If  $k=1$

$$\begin{cases} \|a_i - b_j\|_2^2 \leq R_B^2 \text{ for } (i, j) \in E_{ab} \\ \|b_i - b_j\|_2^2 \leq R_B^2 \text{ for } (i, j) \in E_{bb} \end{cases} \quad (1)$$

elseif  $k=0.5$

$$\begin{cases} R_A^2 \geq \|a_i - b_j\|_2^2 > R_B^2 \text{ for } (i, j) \in E_{ab} \\ R_A^2 \geq \|b_i - b_j\|_2^2 > R_B^2 \text{ for } (i, j) \in E_{bb} \end{cases} \quad (2)$$

else  $k=0$

$$\begin{cases} \|a_i - b_j\|_2^2 > R_A^2 \text{ for } (i, j) \in E_{ab} \\ \|b_i - b_j\|_2^2 > R_A^2 \text{ for } (i, j) \in E_{bb} \end{cases} \quad (3)$$

where  $R_B$  is the maximum distance (called the rangeB) within which strong connectivity can be established.  $R_A$  is the maximum distance (called the rangeA) within which weak connectivity can be established.

From the view of optimization, (1) involves convex constraints, (2) involves both convex and non-convex constraints, and (3) involves non-convex constraints. Therefore, the finding of the coordinates of the unknown nodes in this localization problem should be a non-convex optimization.

### III. AN EXAMPLE TO ILLUSTRATE TWO-LEVEL RANGE

The following simple example with 10 nodes is used to illustrate the two-level range connectivity-based localization problem. An sensor network with 4 anchors and 6 unknown nodes is placed in a square area of  $[0,10]$  by  $[0,10]$  (Figure 3), in which the squares represent the anchors, and the circles represent unknown nodes. The green lines indicate the connection within the range value of 5.

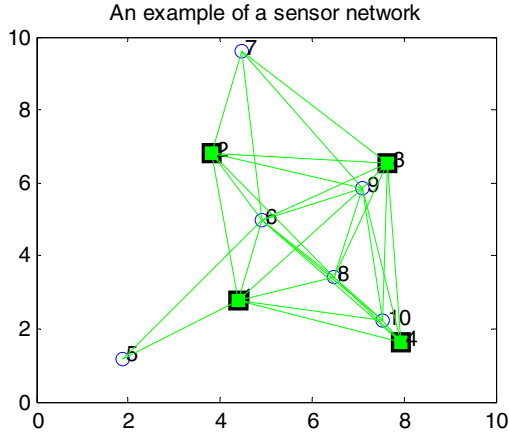


Figure 3. An example of a sensor network with 10 nodes

In the example, range A is 5 and range B is set as  $5\alpha$ , where  $\alpha$  is a real number less than 1 to indicate the ratio between the two ranges. When  $\alpha$  is set as 0.618, the distances between the nodes are shown in Figure 4, which are used for setting up the constraints in the localization problem. The strong connection is represented by the green lines; the weak connection is represented by the blue lines; the red line means there is no connection between two nodes. The edges corresponding to connections or disconnections can be divided into two categories: the first one includes edges between two unknown nodes and edges between an anchor and an unknown node. The second category is the edges between two anchors. The constraints in our algorithm are derived from the first category only, for the anchor locations are assumed known. The number of the connections and disconnections and their convexity are given in TABLE I.

TABLE I. NUMBER & CONVEXITY OF THE CONNECTIONS & DISCONNECTIONS

	1 <sup>st</sup> category	2 <sup>nd</sup> category	Nature of Constraint
# Strong connection	12	0	convex
# Weak connection	19	5	non-convex
# Disconnection	14	1	non-convex

An example of strong connection is the convex constraint for the distance between node 6 and 9 to be within  $R_B$  or  $d_{6,9} \leq R_B$ . For distance between node 7 and 9,  $R_B < d_{7,9} \leq R_A$ . Note that the distance between node 7 and 8 must be greater than  $R_A$ , and therefore  $R_A < d_{7,8}$ , which is another non-convex constraint. The localization problem

involves searching for a solution (which are the coordinates of the unknown nodes) which would satisfy all the convex and non-convex constraints in the first category.

In this paper, genetic algorithm has been used to search for the solutions. For the given node topology, a solution is obtained for the usual one-level range connectivity. Then, solution for the two-level range with ratio of 0.5, 0.618, 0.666, 0.75, 0.8 are obtained. For each case, the genetic algorithm is repeated for 30 times. The errors for the 30 solutions are drawn as box plot in Figure 5. It is clear that the solutions from the two-level range cases have better accuracy than the one level range. The errors from the solutions of the other algorithms such as SDP, DV-Hop and MDS are also shown. The errors of two-level range are about one third of that of the current algorithms.

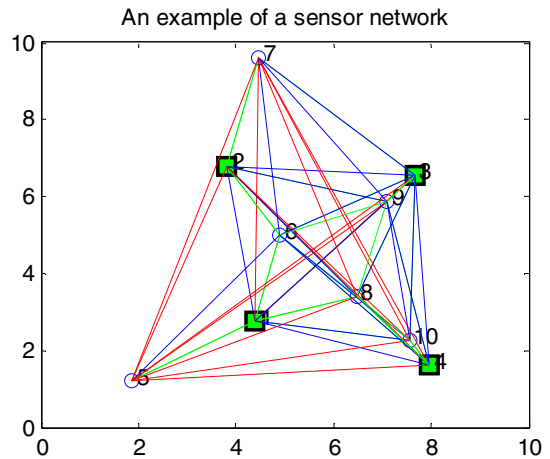


Figure 4. The constraints under two-level range

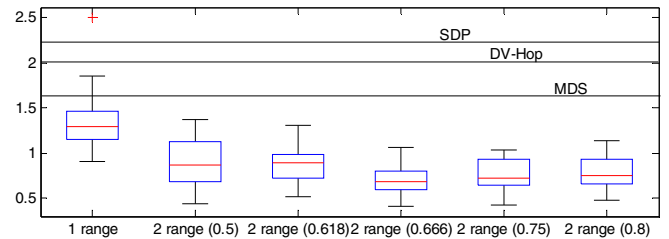


Figure 5. The error (shown by boxes) of two-level range localization of different ratio ( $\alpha=0.5, 0.618, 0.666, 0.75, 0.8$ )

### IV. EVOLUTIONARY ALGORITHM TO SOLVE THE OPTIMIZATION PROBLEM

An evolutionary algorithm is used to find a solution which would satisfy all the constraints. Evolutionary algorithms can potentially deal with general non-convex problems and to avoid, nevertheless in a stochastic, non-controllable way, local minima and related effects [9, 10].

In this two-range localization problem, the variable vector encodes the locations of all unknown nodes. There are two objectives that are needed to minimize. We set the first objective to be the number of cases that the estimated node positions fail the constraints, no matter it is a convex or non-convex constraint. If a set of variables containing the estimated

node position would make the target function value to be zero, the set of variables would be our answer. The second objective is the summation of the difference between the estimated distances and range A, if the estimated distances violate the connectivity. It is aimed at helping the first objective to converge. Obviously, if a variable vector would make the first objective zero, the second objective of this variable must also be zero.

The problem is a two-objective optimization problem, which can be solved by evolutionary algorithm. Examples of some popular algorithms include the Non-dominated Sorting Genetic Algorithm (NSGA) and its evolution (NSGA-II), Strength Pareto Evolutionary Approach (SPEA) and its evolution (SPEAS2). The Pareto Archived Evolution Strategy (PAES) mainly focuses on applying Pareto-based ranking schemes [11]. In this paper, we use PAES to tackle this problem.

The implementation of PAES used in our optimization can be found from jMetal [12]. PAES may represent the simplest possible nontrivial algorithm capable of generating diverse solutions in the Pareto optimal set. The simplest form: (1+1) evolution strategy, which is applied in this paper, employs local search but using a reference archive of previously found solutions in order to identify the approximate dominance ranking of the current and candidate solution vectors [13]. PAES comprises three parts: the candidate solution generator, the candidate solution acceptance function, and the non-dominated-solution (NDS) archive. The candidate solution generator is similar to simple random mutation hill-climbing, but subject to the preferring the less crowded solutions to keep the diversity preservation. It maintains a single current solution and, at each iteration, produces a single new candidate via random mutation. The (1+1)-PAES algorithm explanation can be found in [13].

## V. SIMULATION RESULT FROM 100-NODE NETWORK

Another result from the proposed two-level range connectivity localization with evolutionary algorithm is using a scenario of 100 nodes. The 100 nodes are randomly placed in a square area of [0,10] by [0,10], with 20% of them being anchors. The aim of the localization problem is to find the location of the other 80 unknown nodes. In the first simulation, range A is set to 1.5, and the range B is set to 1.5 multiplied with a ratio. Different ratios have been attempted, such as 0.5, 0.618, 0.666, 0.75, 0.8. For each case, in order to get a more convincing result, five different network topologies with different node locations are built, and the evolutionary algorithm is used for 30 trials for each case.

The accuracy of the estimation is evaluated by the difference between the estimated positions of the unknown nodes and the positions of the corresponding nodes when the problem is setup. The average error per unknown node is calculated by the formula below.

$$\text{error per node} = \frac{\sum_{i=m+1}^{m+n} \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}}{n}$$

where  $(\hat{x}_i, \hat{y}_i)$  is the estimated position of node  $i$ ,  $(x_i, y_i)$  is the real position of the node  $i$ .

As a nature of evolutionary algorithm, when the maximum number of evaluations is fixed, it cannot be guaranteed that all the constraints are satisfied in every trial. Figure 6 shows the number of convergent cases in which all constraints are satisfied, when the maximum number of evaluation is set to 1,000,000. Note that the average percentage for satisfying all the constraints is over 50%.

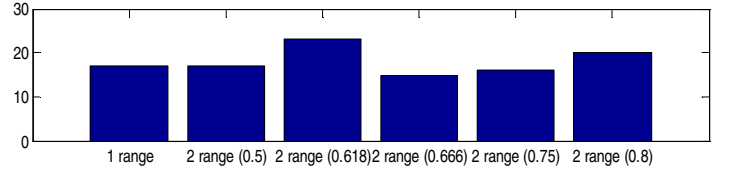


Figure 6. The number of the convergent cases in 30 trials

Boxplots of the average error for unknown nodes for the five topologies are shown in Figure 7. For each topology, there are 6 boxes. The 1<sup>st</sup> box is the error of the convergent trials for one-level range localization, and the 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, 5<sup>th</sup>, 6<sup>th</sup> box represents the two-level range localization with ratio  $\alpha$  being 0.5, 0.618, 0.666, 0.75, 0.8, respectively. On each box, the central mark is the median of the values of error per node for the convergent trials, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points the algorithm considers to be not outliers, and the outliers are plotted individually as red crosses '+'.

In Topology 1, the average error of the one range localization is 0.34, while the average errors of two-level range are smaller, i.e. 0.27, 0.23, 0.22, 0.2, 0.22. For each topology, the errors of five different ratios of two-level range are almost 60% of the error of the one-level range. The best ratio value is dependent on the value of range and topology of the network. The accuracy improvement of two-level range is more obvious when the range A value is set to be 2.5 as shown in Figure 8.

## VI. CONCLUSION

To improve the accuracy of the connectivity-based localization problem in sensor network, a two-level range method is proposed to enrich connectivity information. Range A is set to be the initial range of normal connectivity-based localization, while Range B, which is added in this method, is smaller than Range A, and can give additional connectivity information. The localization problem based on a two-level range connectivity can be solved as a non-convex optimization problem. In this paper, simulation result shows the two-level range connectivity can greatly improve the accuracy of connectivity-based localization.

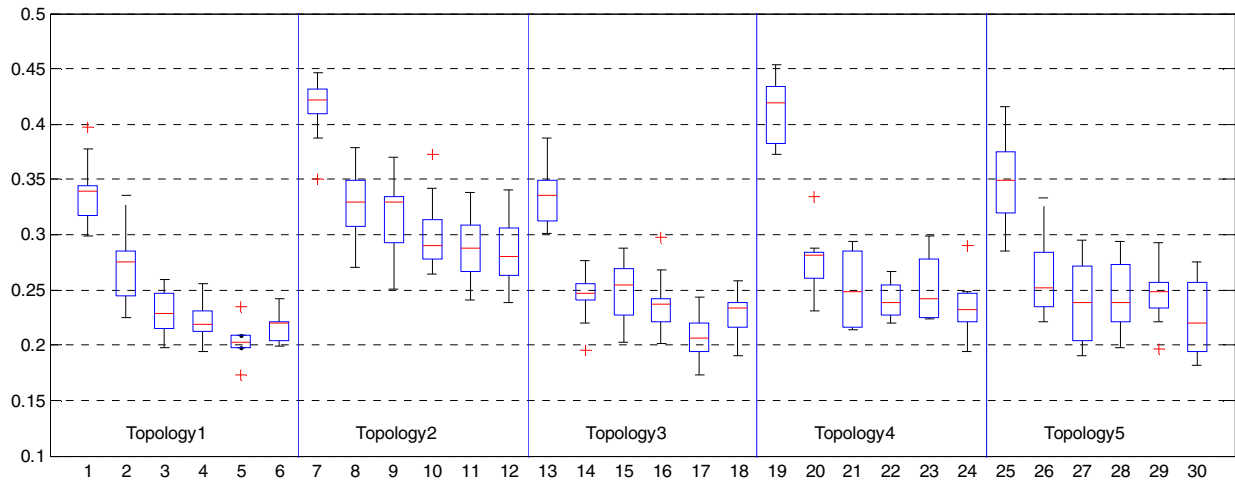


Figure 7. The error compare in the 5 topologies with range  $A=1.5$ , range  $B=1.5\alpha$

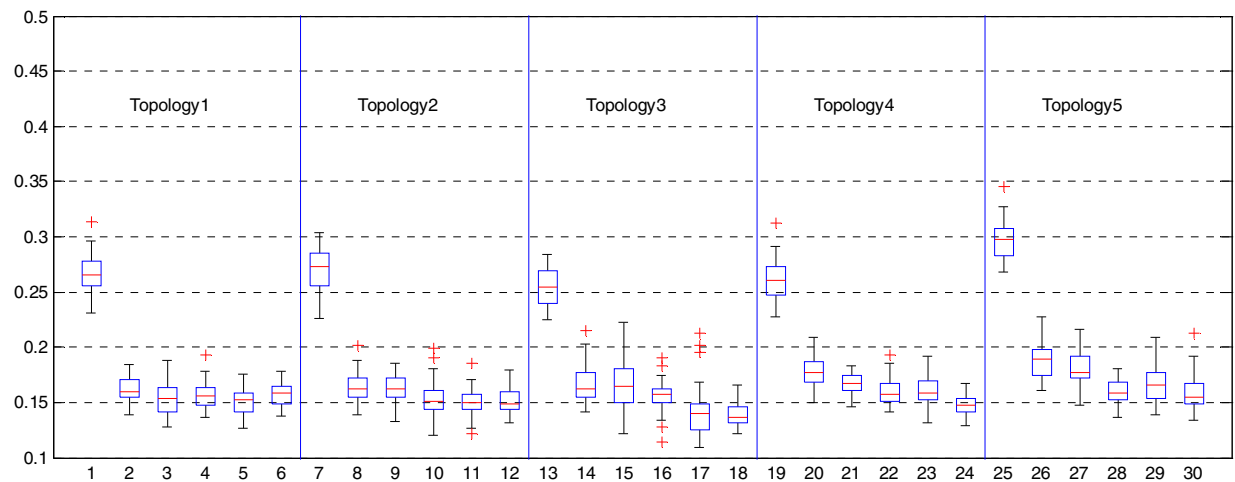


Figure 8. The error compare in the 5 topologies with range  $A=2.5$ , range  $B=2.5\alpha$

## REFERENCES

- [1] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low-cost outdoor localization for very small devices," *Personal Communications, IEEE*, vol. 7, pp. 28-34, 2000.
- [2] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th annual international conference on Mobile computing and networking San Diego, CA, USA: ACM*, 2003.
- [3] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, "Localization from mere connectivity," in *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking Annapolis, Maryland, USA: ACM*, 2003.
- [4] D. Niculescu and B. Nath, "DV Based Positioning in Ad Hoc Networks," *Kluwer journal of Telecommunication Systems*, pp. 267-280, 2003.
- [5] L. Doherty, K. S. J. Pister, and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, 2001, pp. 1655-1663 vol.3.
- [6] D. Qiao and G. K. H. Pang, "Two-range Connectivity-based Sensor Network Localization," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing Victoria, Canada*, 2011.
- [7] V. Shnayder, M. Hempstead, B.-r. Chen, G. W. Allen, and M. Welsh, "Simulating the power consumption of large-scale sensor network applications," in *Proceedings of the 2nd international conference on Embedded networked sensor systems Baltimore, MD, USA: ACM*, 2004.
- [8] H. Ting-Chao and L. Victor, "Transmission Range Control in Multihop Packet Radio Networks," *Communications, IEEE Transactions on*, vol. 34, pp. 38-44, 1986.
- [9] E. S. Mistakidis and G. E. Stavroulakis, *Nonconvex Optimization in Mechanics: Kluwer Academic Publishers*, 1998.
- [10] R. H. Byrd and R. A. Waltz, "An active-set algorithm for nonlinear programming using linear programming and equality constrained subproblems," *Optimization Technology Center, Northwestern University* 2002.
- [11] "[http://en.wikipedia.org/wiki/Multiobjective\\_optimization](http://en.wikipedia.org/wiki/Multiobjective_optimization)," 2011.
- [12] "<http://jmetal.sourceforge.net/>," 2011.
- [13] J. D. Knowles and D. W. Corne, "Approximating the Nondominated Front Using the Pareto Archived Evolution Strategy," *Evol. Comput.*, vol. 8, pp. 149-172, 2000.
- [14] E. Zitzler, M. Laumanns, and S. Bleuler, "A Tutorial on Evolutionary Multiobjective Optimization," in *Metaheuristics for Multiobjective Optimisation: Springer-Verlag*, 2003, pp. 3--38.