



ELSEVIER

Contents lists available at SciVerse ScienceDirect

# Engineering Applications of Artificial Intelligence

journal homepage: [www.elsevier.com/locate/engappai](http://www.elsevier.com/locate/engappai)

## On the design of a novel two-objective harmony search approach for distance- and connectivity-based localization in wireless sensor networks

Diana Manjarres<sup>a</sup>, Javier Del Ser<sup>a</sup>, Sergio Gil-Lopez<sup>a</sup>, Massimo Vecchio<sup>b</sup>, Itziar Landa-Torres<sup>a</sup>, Sancho Salcedo-Sanz<sup>c,\*</sup>, Roberto Lopez-Valcarce<sup>b</sup>

<sup>a</sup> Tecnalia Research & Innovation, Parque Tecnológico de Bizkaia, Ed. 202, Zamudio, Bizkaia, Spain

<sup>b</sup> Department of Signal Theory and Communications, University of Vigo, Spain

<sup>c</sup> Department of Signal Processing and Communications, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain

### ARTICLE INFO

#### Article history:

Received 25 January 2012

Received in revised form

4 May 2012

Accepted 2 June 2012

Available online 23 June 2012

#### Keywords:

Wireless Sensor Networks

Localization

Multi-objective

Flip ambiguity

Harmony Search

### ABSTRACT

In several wireless sensor network applications the availability of accurate nodes' location information is essential to make collected data meaningful. In this context, estimating the positions of all unknown-located nodes of the network based on noisy distance-related measurements (usually referred to as *localization*) generally embodies a non-convex optimization problem, which is further exacerbated by the fact that the network may not be uniquely localizable, especially when its connectivity degree is not sufficiently high. In order to efficiently tackle this problem, we propose a novel two-objective localization approach based on the combination of the harmony search (HS) algorithm and a local search procedure. Moreover, some connectivity-based geometrical constraints are defined and exploited to limit the areas in which sensor nodes can be located. The proposed method is tested with different network configurations and compared, in terms of normalized localization error and three multi-objective quality indicators, with a state-of-the-art metaheuristic localization scheme based on the Pareto archived evolution strategy (PAES). The results show that the proposed approach achieves considerable accuracies and, in the majority of the scenarios, outperforms PAES.

© 2012 Elsevier Ltd. All rights reserved.

### 1. Introduction

The increasing attention attracted by the so-called wireless sensor networks (WSNs) during the last decade stems from their capacity to efficiently sense physical phenomena, without the need of any wired link and at a reduced per-node computational complexity (Akyildiz et al. (2002)). In addition to the classical monitoring applications for which WSNs were initially targeted, their proliferation and the value of their sensed information has ignited the interest in context-aware services and applications, in which the knowledge of nodes' location becomes essential so as to provide valuable spatio-temporal information on the physical world and its dynamics. Hence, in such applications it is essential to associate the captured data with the location of the node at hand in order to make such sensed data meaningful (Hu and Evans, 2004). Indeed, the importance of localization is exemplified by a large number of applications such as vehicle and animal tracking, habitat monitoring, precision agriculture, location-based routing (Ko and Vaidya, 2000), industrial sensing, infrastructure

security and control of machinery, all of which inherently rely on location information of the constituent nodes of the WSNs.

From a theoretical point of view, *effective* location awareness in WSNs might be based on installing a global positioning system (GPS) on each compounding node of the network. However, the size constraints, the limited energy budget and the reduced cost imposed on such nodes make this solution unfeasible in practice. Another drawback of GPS-based schemes lies on the fact that they are not suitable for indoor and underground deployments, because of the lack of satellite coverage in these areas. Consequently, *effective*, yet also *efficient* node localization in WSNs is still considered to be a challenging paradigm in the scientific community, which has unchained an upsurge of alternative approaches to this problem.

In the majority of such alternative schemes, it is assumed that only a fraction of the nodes of the network, referred to as *anchor* nodes, know their coordinates either by resorting to the aforementioned locally installed GPS devices, or by being placed at pre-defined geographical locations. The remaining nodes (denoted as *non-anchor* nodes) are able to infer their positions based on (1) the information on the location of the anchor nodes, (2) connectivity-based information, and (3) noisy pairwise distance estimations obtained through distance-ranging techniques, e.g. received signal strength (RSS) (Alippi and Vanini, 2006), time difference of arrival

\* Corresponding author. Tel.: +34 91 885 6731; fax: +34 91 885 6699.

E-mail address: [sancho.salcedo@uah.es](mailto:sancho.salcedo@uah.es) (S. Salcedo-Sanz).

(TDoA) (Savvides et al., 2001) or angle of arrival (AoA) (Niculescu and Nath, 2003). In this work we will focus on RSS based techniques, for which the maximum likelihood (ML) estimation is the most natural approach to tackle the localization problem. However, the ML formulation of the problem, i.e. locating the nodes from a set of pairwise measured distances, results in a multivariable non-convex optimization problem (Aspnes et al., 2004) for which different methods have been proposed in the literature, namely multidimensional scaling (MDS) (Ji and Zha, 2004; Costa et al., 2006), semi-definite programming (SDP) (Biswas et al., 2006) and stochastic optimization (Kannan et al., 2005, 2006). MDS, which is widely applied in data analysis, hinges on obtaining the relative coordinates of each node based on a starting distance matrix. On the other hand, SDP relaxes the original ML problem formulation in order to obtain an approximate solution at a reduced computational effort (Biswas et al., 2006; Tseng, 2007). Since the relaxation may increase the estimation error (Wang et al., 2008), additional refinements, such as gradient descent search procedures (Liang et al., 2004), are often employed to refine the initial solutions obtained by SDP (Biswas and Ye, 2004). Finally, the third class of techniques considers heuristic optimization algorithms to efficiently solve the localization problem, such as simulated annealing (SA) (Kannan et al., 2006) or genetic algorithms (Vecchio et al., 2012). In this paper we concentrate on this last set by deriving a connectivity- and distance-based heuristic localization technique.

The localization problem is generally further complicated by the non-uniqueness of the nodes' location in the network. Indeed, when tackled from a graph-theoretic perspective, the problem of finding a unique graph realization compatible with the inter-node distance measurements (i.e. the so-called *network localizability* problem) has gained momentum in the last years (Jackson and Jordan, 2005; Yang and Liu, 2011). Dense connectivity is obviously a principal requirement for a network to be localizable, even though a necessary and sufficient condition for the network localizability still represents an open research issue. The lack of localizability may lead to the so-called *flip ambiguity* phenomenon, which generally occurs due to the almost collinear placement of the neighbors of a particular sensor node. In this case, although the node location is flipped with respect to the virtual line connecting its neighboring nodes, the estimated topology is still compatible with the inter-node distance measurements. This effect can propagate catastrophically to subsequent iterations affecting, in turn, the location estimates of other nodes of the network. Therefore, it is of utmost importance to identify possibly flipped nodes, and to perform some specific processing during the localization task in order to avoid, or at least alleviate, its effect on subsequent location estimates. The flip ambiguity phenomenon is exemplified in Fig. 1 where, since sensor nodes A, B, G and H are nearly collinear, node F can be reflected to position F', while satisfying both distance and connectivity constraints. The effects of the flip in the estimated position of F can lead nodes C, D and E to be estimated at location C', D' and E', at their turn.

So far a number of heuristic localization approaches have specifically incorporated strategies for efficiently dealing with flip ambiguities in sparse networks. For instance, the authors in Kannan et al. (2005) proposed a two-phased simulated annealing (SA) approach tackling the localization problem in presence of the flip ambiguity phenomenon. During the first phase, SA is used to obtain an accurate estimate of the node locations by minimizing the squared error between the estimated and the measured inter-node distances while, in the second phase, an additional error term is added to the cost function when the estimated node location violates the connectivity constraints imposed by the network topology. Likewise, the same authors in Kannan et al. (2010) propose a robust criterion to detect flip ambiguities and enhance the reliability of the estimations by

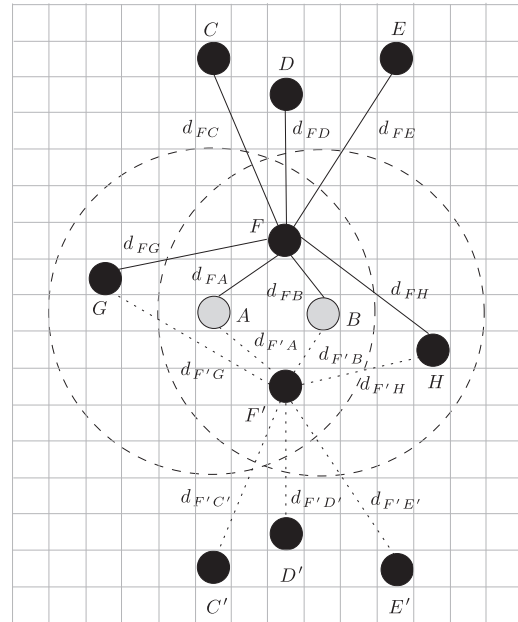


Fig. 1. An example of the flip ambiguity problem.

rejecting flipped nodes from being included in the localization process. More recently, a two-objective evolutionary algorithm based on the Pareto archived evolution strategy (PAES) is presented in Vecchio et al. (2012). There, the first objective function to be minimized, denoted as CF, is given by the squared error between the estimated and the measured inter-node distances. The second objective function, referred to as CV, is defined as the sum of neighborhood violations in the candidate topology. The obtained results show that the multi-objective treatment of the problem outperforms SA in terms of normalized location error (NLE) for a wide set of network topologies.

This paper advances over the state of the art by proposing a novel two-objective heuristic localization algorithm based on the combination of the Harmony Search algorithm with a novel local search procedure that alleviates the effects of the flip ambiguity phenomenon. Similar to the PAES-based counterpart, each candidate solution is evaluated in terms of both localization accuracy (CF) and connectivity constraint violation (CV). Our proposal, hereafter referred to as non-dominated sorting harmony search (NSHS), resorts to the same ranking and crowding estimation operators used by the non-dominated genetic algorithm-II (NSGA-II, Deb et al., 2002) in order to achieve a wide and uniformly covered Pareto-optimal front approximation. The obtained simulation results show that our approach outperforms the multi-objective PAES-based approach proposed in Vecchio et al. (in press) for a number of different network topologies and performance metrics.

The rest of the paper is structured as follows: in Section 2 the node localization problem is mathematically formulated, whereas Section 3 delves into the proposed NSHS multi-objective algorithm. Section 4 presents the simulation framework and discusses the obtained experimental results in terms of normalized localization error (NLE) and different unary multi-objective quality indicators, i.e. the hypervolume, the unary- $\epsilon$  and the  $R_2$  indicators. Finally, concluding remarks are discussed in Section 5.

## 2. Problem statement

The node localization problem can be mathematically defined by considering a WSN composed by a set of  $n$  nodes uniformly

deployed in  $T \subset \mathbb{R}^2$ , out of which the first  $m < n$  nodes represent the set of anchor nodes, whose coordinates  $\mathbf{p}_i = (x_i, y_i) \in T$  ( $i \in \{1, \dots, m\}$ ) are fixed and assumed to be known a priori for the localization algorithm. The problem consists of estimating the positions  $\hat{\mathbf{p}}_i = (\hat{x}_i, \hat{y}_i)$ ,  $\forall i \in \{m+1, \dots, n\}$  of the remaining  $n-m$  non-anchor nodes, assuming that any two sensor nodes in the connectivity range of each other, say  $i$  and  $j$ , can estimate their inter-node distance  $d_{ij}$  by resorting to any of the measurement techniques listed in Section 1. In the following,  $d_{ij}$  will be modeled as

$$d_{ij} = \begin{cases} r_{ij} & \text{if } (i, j) \in \{1, \dots, m\} \times \{1, \dots, m\}, \\ r_{ij} + e_{ij} & \text{otherwise,} \end{cases} \quad (1)$$

where  $r_{ij} \triangleq \|\mathbf{p}_i - \mathbf{p}_j\|$  represents the actual (real) distance between node  $i$  and  $j$  ( $\|\cdot\|$  denotes the Euclidean norm), and  $e_{ij}$  stands for the measurement error.

We assume that each node knows which nodes it can communicate with. Let us define the  $n \times n$  binary connectivity matrix  $\mathbf{C}$ , such that  $c_{ij} = 1$  if sensor nodes  $i$  and  $j$  are within the connectivity range of each other, i.e. if  $r_{ij} \leq R$ , where  $R$  represents the radius of the circular coverage area of every node of the network. Hence, we can define the set of neighbors of node  $i$  as

$$\mathcal{N}_i \triangleq \{j \in \{1, \dots, n\}, j \neq i : r_{ij} \leq R\} \quad (2)$$

and its complementary set  $\bar{\mathcal{N}}_i$ , composed by those nodes placed outside the coverage area of node  $i$ . Observe that the positions of anchor nodes and the connectivity range determine the regions in which each non-anchor node can (or cannot) be located. This information can be exploited in the initialization phase and during the localization process to further refine the position estimates of the non-anchor nodes.

The goal of the localization algorithm is to estimate the positions of all non-anchor nodes as accurately as possible given  $R$ ,  $\mathbf{C}$  and  $\{\mathbf{p}_i\}_{i=1}^m$ . To this purpose, two different, though related, objective functions to be minimized can be defined: CF (cost function) and CV (constraint violation). In detail, the first objective function CF is defined as

$$CF \triangleq \sum_{i=m+1}^n \left( \sum_{j \in \mathcal{N}_i} (d_{ij} - \hat{d}_{ij})^2 \right), \quad (3)$$

where  $d_{ij}$  represent the measured inter-node distances given in expression (1), while  $\hat{d}_{ij}$  represent the estimated distances between nodes  $i$  and  $j$  and computed as

$$\hat{d}_{ij} \triangleq \sqrt{(\hat{x}_i - \hat{x}_j)^2 + (\hat{y}_i - \hat{y}_j)^2}. \quad (4)$$

Therefore, CF represents the squared error between the estimated and the measured inter-node distances of those nodes that are in the connectivity range of each others. The second objective function, CV, takes into account the number of connectivity neighborhood constraints violated by the candidate topology. In words, a node  $j$  fulfilling  $j \in \mathcal{N}_i$  for a given  $i$  causes a violation of the connectivity constraints given by  $\mathbf{C}$  if its estimated position lied outside the transmission range of node  $i$  or, equivalently, if  $\hat{d}_{ij} > R$  provided that  $j \in \mathcal{N}_i$ . Thus, CV can be formally defined as

$$CV \triangleq \sum_{i=1}^n \left( \sum_{\substack{j=1 \\ \hat{d}_{ij} > R}}^n c_{ij} + \sum_{\substack{j=1 \\ \hat{d}_{ij} \leq R}}^n (1 - c_{ij}) \right). \quad (5)$$

Thus, both CF and CV leads to the multi-objective formulation of the localization problem

$$\min_{\{\hat{\mathbf{p}}_i\}_{i=m+1}^n} [CF, CV] \quad \text{s.t.} \quad \hat{\mathbf{p}}_i \triangleq (\hat{x}_i, \hat{y}_i) \in T. \quad (6)$$

Finally, in order to a posteriori evaluate the accuracy of an estimated topology we use the *normalized localization error* (NLE) defined as

$$NLE \triangleq \frac{100}{R} \sqrt{\frac{1}{(n-m)} \sum_{i=m+1}^n \|\mathbf{p}_i - \hat{\mathbf{p}}_i\|^2}, \quad [\%], \quad (7)$$

where  $NLE=0\%$  corresponds to the case of a perfect estimation.

### 3. The proposed NSHS algorithm

The bi-objective minimization problem posed in expression (6) will be tackled by means of a multi-objective formulation of the so-called harmony search (HS) algorithm (Geem et al., 2001). HS is a relatively new population-based metaheuristics that imitates the music improvisation process when seeking the most harmonious melody. Since its first publication in 2001, it has been widely applied for solving several optimization problems in different application fields such as vehicle routing (Geem et al., 2005), multicast routing (Forsati et al., 2008), multiuser detection (Zhang and Hanzo, 2009; Gil-Lopez et al., 2009), engineering design (Liao, 2010), radio resource allocation (Del Ser et al., 2012, 2011) and the access node location problem (Landa-Torres et al., 2011). A thorough review of the HS algorithm, its comparison with other metaheuristic approaches and a number of related applications can be found in Geem (2009). However, to the best of our knowledge, no previous work has been reported in the scientific community dealing with the application of multi-objective HS to the node localization problem.

Given its population-based nature, HS relies on a set of candidate solutions  $\{\mathbf{H}(k)\}_{k=1}^K$  (harmony memory), which are iteratively refined by means of intelligent combination and mutation operators applied note-wise. Adopting the classical notation of the related literature, in the following we will refer to a possible candidate set  $\mathbf{H}(k)$  as *harmony* or *melody*, whereas a *note* denotes any of its compounding entries  $h_l(k)$ , with  $k \in \{1, \dots, K\}$  and  $l \in \{1, \dots, L\}$ . In our optimization framework, each melody represents the positions of all the nodes of the network, thus the harmony memory can be redefined as  $\mathbf{H}(k) \triangleq \{\mathbf{p}_i(k)\}_{i=1}^n$ , with  $k=1, \dots, K$ . The first  $m$  notes (i.e.  $i=1, \dots, m \forall k$ ) are pairs of real numbers representing the  $(x, y)$  coordinates of the anchor nodes (which are assumed to be known a priori), whereas the remaining  $n-m$  pairs ( $i=m+1, \dots, n$ ) correspond to the estimated coordinates of the non-anchor nodes of the network. Thus, the improvisation process of the proposed HS algorithm will be only applied to the estimated positions of the non-anchor nodes (i.e. the second part of the melody or, equivalently,  $\{\mathbf{p}_i(k)\}_{i=m+1}^n$ ), which can be located in the areas defined by the topological constraints described in the previous section.

Briefly, the improvisation procedure is controlled by three different stochastic operators, each driven by its corresponding probabilistic parameter: the *harmony memory considering rate* (HMCR), the *pitch adjusting rate* (PAR) and the *random selection rate* (RSR)<sup>1</sup>. After the improvisation procedure, the value of the two objective functions (CF and CV) are separately computed for each improvised harmony and the *best K* harmonies (with respect to the fitness values and the spread, the latter measured in terms of *crowding distance* as in Deb et al., 2002) are kept in the harmony memory for the next iteration. This procedure is repeated until a fixed number of iterations  $\mathcal{I}_{HS}$  is completed.

<sup>1</sup> The RSR operator does not belong to the original HS implementation, but allows for an improved control of the trade-off between the explorative and exploitative behavior of the algorithm.

The flow diagram of the proposed NSHS algorithm is depicted in Fig. 2, and consists of five steps:

- A. The *initialization* process is executed only at the first iteration. At this step, the initial  $K$  harmonies are arranged by fixing  $\forall k=1, \dots, K$  the first  $m$  notes to the actual positions of the anchor-nodes (i.e.  $\{h_i(k)\}_{i=1}^m = \{\mathbf{p}_i\}_{i=1}^m \triangleq \{\hat{x}_i(k), \hat{y}_i(k)\}_{i=1}^m$ ) and by generating the remaining  $n-m$  notes at random, within the topological constraints imposed on the corresponding  $n-m$  non-anchor nodes.
- B. During the *improvisation* procedure, the stochastic operators are sequentially applied to each note of the harmonies so as to produce the new set of  $K$  improvised harmonies (i.e. candidate positions). In detail,
  1.  $\text{HMCR} \in [0, 1]$  sets the probability that the new value for a certain note  $\{\hat{x}_i(k), \hat{y}_i(k)\}$  is drawn uniformly from the values of this same note in all the other  $K-1$  melodies.

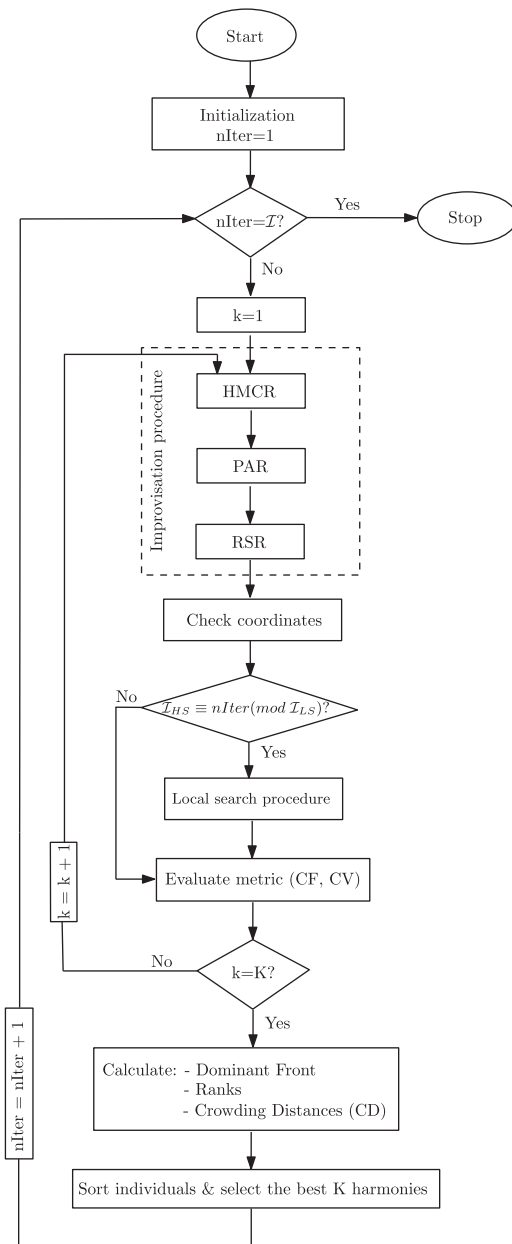


Fig. 2. Flow diagram of the proposed NSHS algorithm. Note that  $a \equiv b \pmod c$  denotes arithmetic congruence (i.e.,  $a$  and  $b$  are congruent modulo  $c$  if the difference  $(a-b)$  is an integer multiple of  $c$ ).

2.  $\text{PAR} \in [0, 1]$  refers to the probability that the new value for a given note  $(\hat{x}_i^k, \hat{y}_i^k)$  is taken randomly from its coverage area considering the geometrical constraints posed by anchor nodes for the non-anchor node at hand. Specifically, when operating on  $(\hat{x}_i^k, \hat{y}_i^k)$ ,  $\theta$  points (i.e.  $(x, y)$  coordinates) are drawn uniformly at random from the intersection of the circular coverage areas corresponding to every anchor node  $j$  such that  $c_{ij} = 1$  ( $j \in \{1, \dots, m\}$ ). The note  $(\hat{x}_i^k, \hat{y}_i^k)$  is then substituted by the geometrically closest point among the picked  $\theta$  coordinates.
3. The random selection rate,  $\text{RSR} \in [0, 1]$ , acts in a similar manner to the PAR procedure. However, in this case the note at hand is replaced with any point picked at random from the same intersection as above (rather than choosing the closest point in terms of Euclidean norm).
- C. Additionally, a *local search procedure* is performed every  $\mathcal{I}_{LS}$  iterations. The details of this refinement procedure will be given in Section 3.1.
- D. At each iteration the generated harmonies are evaluated in terms of CF and CV (expression (3) and expression (5), respectively). Then, as explained in Kalyanmoy et al. (2000), each current and improvised harmony is associated with a rank equal to its non-dominance level (1 for the best level, 2 for the next-best, and so on). Then, within each front, a specific crowding measure representing the sum of distances to the closest harmony along each objective is used to define an ordering among the harmonies: in order to cover the overall objective space, harmonies with large crowding distance are preferred to harmonies with small crowding distance. Thus, the harmony memory is filled by selecting the best  $K$  harmonies (considering first the ordering among the fronts and then among the harmonies).
- E. If the iteration counter  $nIter < \mathcal{I}_{HS}$ , the algorithm iterates by setting  $nIter = nIter + 1$  and returning to step B. Otherwise, the algorithm stops and the set of harmonies belonging to the first non-dominance level represents the Pareto front approximation.

### 3.1. Local search procedure

The proposed local search procedure is triggered every  $\mathcal{I}_{LS}$  iterations, at the end of the improvisation procedure. When invoked, it is executed on each non-anchor node  $i$  (1) lying outside the coverage area of any anchor node; and (2) whose any of its neighbors in the estimated topology violates the connectivity constraints imposed by the  $i$ -th row of  $\mathbf{C}$ . Fig. 3 depicts an example of the application of the local search procedure to the estimated node  $1'$  in a simplistic scenario. In this setup, anchor nodes A, B and C are represented by crosses ( $\times$ ), whereas real and estimated positions of non-anchor nodes are denoted with circles ( $\circ$ ) and squares ( $\square$ ), respectively. Real and estimated positions of node 1 (to which the local search procedure is applied) are highlighted by filling the marks in black ( $\bullet$  and  $\blacksquare$ , respectively). In this setup, the *real* connectivity matrix  $\mathbf{C}$  and the *estimated* connectivity matrix  $\mathbf{C}'$  are given in Fig. 4. Note that  $\mathbf{C}$  is provided as an input parameter to the algorithm, whereas  $\mathbf{C}'$  can be inferred for every candidate topology.

The fact that non-anchor node  $1'$  lies outside the coverage range of any anchor node and has some false neighbors implies a violation in the connectivity constraints imposed by  $\mathbf{C}$  (i.e.  $7'$  and  $8'$ ). Hence,  $1'$  can be selected by the local search procedure to be refined. The next step is to identify the group of nodes to be moved together with node  $1'$ , so as to avoid generating new connectivity violations. Specifically, this group will be composed by selecting those non-anchor nodes that are not inside the coverage range of any anchor node, but are real neighbors of



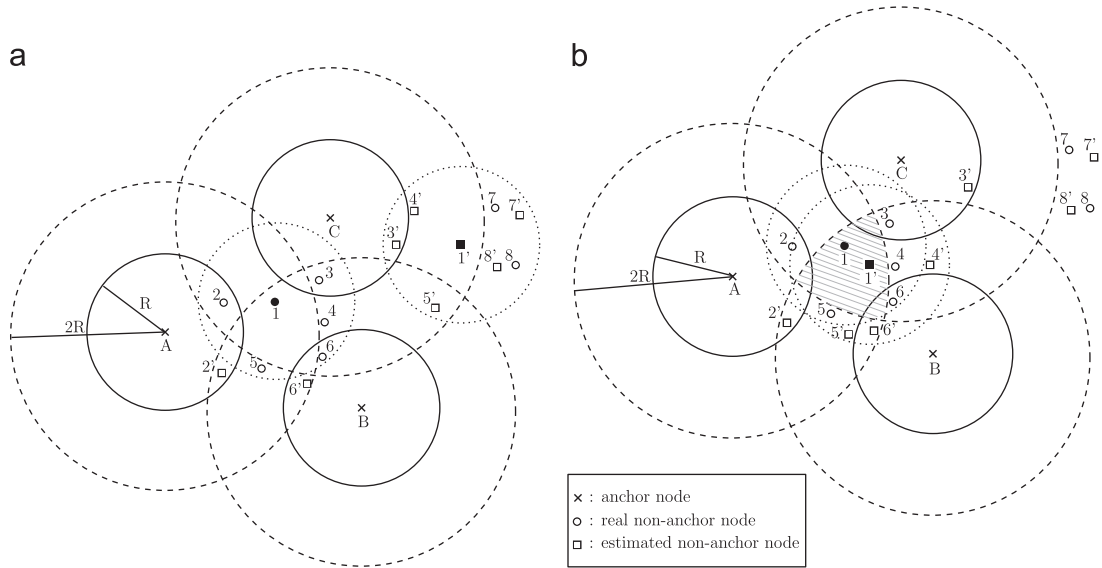


Fig. 3. An example of scenario obtained at the end of the improvisation process, (a) before and (b) after the local search procedure.

$$C = \begin{pmatrix} A & B & C & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$C' = \begin{pmatrix} A & B & C & 1' & 2' & 3' & 4' & 5' & 6' & 7' & 8' \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Fig. 4. The real connectivity matrix  $C$  and the estimated connectivity matrix  $C'$  for the local search example described in Section 3.1.

node 1. Thus, in this specific case, the group will be composed by nodes 4' and 5'. On the other hand, those anchor nodes connected to any actual neighbors of node 1 (nodes A, B and C) define the region in which node 1' can be located. This region comprises the intersection of the annuli with inner and outer radii  $R$  and  $2R$  respectively, centered in the selected anchor nodes, under the condition that the number of false neighbors decreases. Finally, the group of nodes 4' and 5' are randomly located inside the coverage region of node 1' (see Fig. 3).

#### 4. Numerical results

In order to assess the performance of the proposed multi-objective NSHS algorithm in tackling the node localization problem in WSNs, a comparison study between the NSHS and the PAES-based approaches has been performed based on extensive Monte Carlo simulations. In these simulations the localization technique proposed in Vecchio et al. (in press) will be used as a reference for the comparison, as it was shown to be more effective with respect to the Simulated Annealing-based counterpart in Kannan et al. (2006).

##### 4.1. The PAES-based approach to the node localization problem

The PAES algorithm was introduced in Knowles and Corne (2000) and probably represents the simplest possible non-trivial algorithm capable of generating diverse solutions in the Pareto optimal set. Further, PAES is characterized by a lower

computational complexity than traditional niching methods (Knowles and Corne, 2000; Coello et al., 2006). Briefly, it consists of three steps: (1) generation of a candidate solution  $c$ ; (2) mutation of  $c$  so as to produce a new candidate solution  $m$ ; and (3) replacement of  $c$  with  $m$  if  $m$  dominates  $c$ , or inclusion of  $m$  in the archive of non-dominated solutions if it is not dominated by any solution already included in the archive. The archive of non-dominated solutions is divided into a number of equally sized regions for which a crowding degree value is determined by counting the solutions associated to each region. This approach prioritizes solutions associated with scarcely crowded regions, so as to increase the diversity in the Pareto front. The algorithm stops after a maximum number of iterations, and the final archive of non-dominated solutions represents the Pareto front approximation. For more details on PAES the reader should refer to Knowles and Corne (2000) and Coello et al. (2006).

Regarding the specific PAES-based approach to the node localization problem presented in Vecchio et al. (in press), two ad-hoc mutation operators, namely the node mutation and the neighborhood mutation operators, are used. The first mutation operator performs a uniform-like mutation with probability  $P_M$ , in which the mutated position of a randomly selected non-anchor node is generated within its topological constraints. If not applied, then the second mutation operator is executed. It also mutates the position of a given non-anchor node considering its topological constraints, but it further applies the same translation to the neighbors of the node with a certain probability  $P_N$ . In addition to the topological constraints imposed by the connectivity matrix  $C$ , the PAES-based approach in Vecchio et al. (in press) defines a new

**Table 1**  
Parameter setup of the compared techniques.

PAES	NSHS
Archive size: 50	$K$ : 50
Number of regions: 25	HMCR: 0.9
$P_M$ : 0.9	PAR: 0.01
$P_N$ : 0.3	RSR: 0.01
$\mathcal{I}_P$ : 100 000	$\mathcal{I}_{HS}$ : 2000

type of node which is second-level neighbor to at least one anchor node. Thus, its position is forced to lie within the intersection of the annuli with inner and outer radii  $R$  and  $2R$ , respectively, and centered in the anchor nodes which it is second-level neighbor to.

#### 4.2. Simulation setup

The simulation framework consists of 12 different network topologies built by uniformly placing  $n=200$  nodes in  $T \triangleq [0, 1] \times [0, 1]$ . Out of these nodes, 10% are assumed to be the anchor nodes (i.e.  $m=20$ ), while the remaining  $n-m=180$  nodes represent the non-anchor nodes to be localized. Moreover, three different connectivity ranges  $R \in \{0.13, 0.15, 0.17\}$  are considered to model different degrees of network sparsity. We assume that the inter-node distances are estimated from RSS measurements, which are generally affected by log-normal shadowing (Liu et al., 1998). Therefore, the errors  $e_{ij}$  follow a zero-mean distribution with variance  $\sigma^2$ . Note that the variance of  $e_{ij}$  in expression (1) is given by  $\lambda^2 r_{ij}^2$  where, without loss of generality,  $\lambda$  is set to 0.1 in all the scenarios.

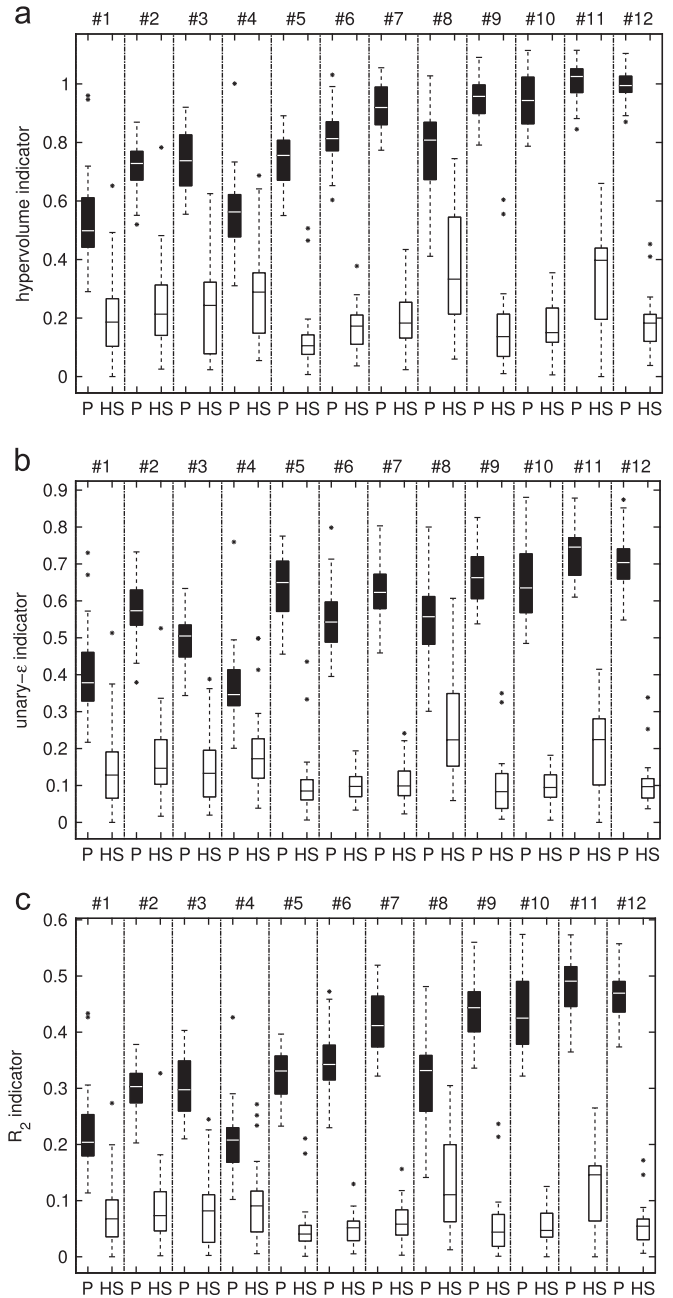
Table 1 summarizes the values of the parameters used in the execution of both algorithms.<sup>2</sup> It should be pointed out that a fair comparison in terms of computational complexity has been targeted by imposing an equal number of fitness evaluations between the compared algorithms. In this context, the number of fitness evaluations for both algorithms equals  $10^5$  in all the simulated cases (i.e.  $K \cdot \mathcal{I}_{HS}$  for NSHS and  $\mathcal{I}_P$  for the PAES-based approach).

Finally, the simulation results have been obtained by performing 30 Monte Carlo experiments for each algorithm and network scenario.

#### 4.3. Simulation results

As a preliminary step, in order to quantitatively compare the approximated fronts obtained by the two multi-objective stochastic solvers on each simulated scenario, we have computed three unary Pareto-compliant quality indicators. In particular, to compare the NSHS and the PAES approaches we have used the hypervolume indicator (Zitzler and Thiele, 2009), the unary- $\epsilon$  indicator (Zitzler et al., 2003) and the  $R_2$  indicator (Hansen and Jaskiewicz, 1998), computed in the performance assessment package provided in the PISA toolkit (Bleuler et al., 2003), and whose features and properties have been discussed in detail in Zitzler et al. (2003). Here, the lower the value of the indicator is, the higher the quality of the corresponding algorithm is.

Fig. 5 shows the box plots of the distributions of three indicators obtained by the two multi-objective solvers, in all the considered scenarios, with the following notation:  $P$  and  $HS$  stand for PAES and NSHS, respectively; 1, ..., 12 denote the network topologies; whiskers are used to represent the lowest and largest



**Fig. 5.** Box plots of the three unary quality indicators obtained by PAES and NSHS ( $P$  and  $HS$ , respectively) in the 12 network topologies. (a) Hypervolume indicator. (b) Unary- $\epsilon$  indicator. (c)  $R_2$  indicator.

values of the distribution; the boxes delimit the lower and the upper quartiles; the medians are depicted with a solid line and the outliers are eventually marked with asterisks. By analyzing the box plots we can observe that the proposed NSHS approach produce significantly better multi-objective performance indicators with respect to the PAES counterpart. For verifying whether the distributions of the different indicators are statistically different, we have performed a two-sided rank sum test (Wilcoxon test). The different colors of the box plots have been used to represent the results of the non-parametric test applied to each distribution of the indicators, for each scenario. Briefly, the test compares the medians of two samples, and returns the  $p$ -value for the null hypothesis that the samples are drawn from the same population. If the  $p$ -value is lower than 0.05, we deduce that the null hypothesis does not hold, that is, the sample medians are

<sup>2</sup> These values are the same for all the simulated scenarios and have been obtained by means of a previous optimization process, not shown for the sake of brevity.

significantly different from each other. A white box plot denotes the best distribution of the indicator values obtained for a specific topology. A black box plot identifies a distribution of the indicator values whose median is larger than the median of the corresponding white box plot with a statistical significance. A gray box plot denotes a distribution of the indicator values whose median is larger than the median of the corresponding white box plot, but the difference between the medians is not statistically significant. We observe that NSHS was always able to produce, for all the considered scenarios, better distributions of the three indicators, with statistical significance. Furthermore, the coverage rate (C-metric) has been calculated based on the approximated dominant front for all scenarios obtaining a value of  $C(NSHS,PAES)=1$ , hence supporting the claim that NSHS dominates PAES with respect to Pareto dominance.

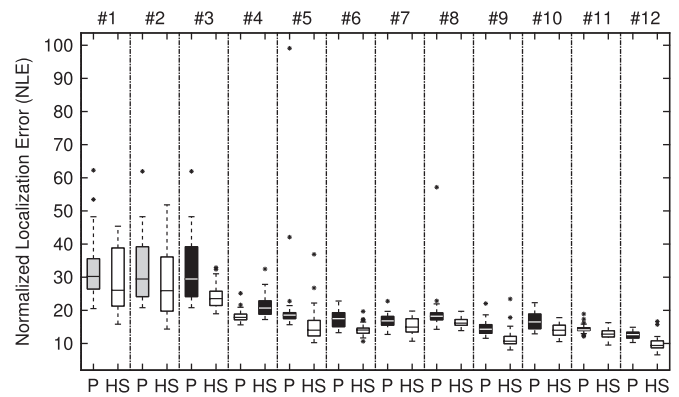
Moreover, in order to evaluate and compare the localization accuracy of the two approaches, we use the normalized localization error (NLE) formally posed in expression (7). Table 2 lists, for every scenario, algorithm and averaged over 30 Monte Carlo experiments, the mean, minimum and standard deviation of the NLE values associated to the topologies belonging to the approximated Pareto front produced after  $\mathcal{I}_P$  (correspondingly,  $\mathcal{I}_{HS}$ ) iterations. Note that in light of the obtained results, NSHS in general outperforms the previously published PAES approach in terms of minimum, mean and standard deviation of the NLE values associated to the topologies embodying their estimated fronts.

To shed more light on the obtained NLE results, the distribution of the best NLE values is computed by selecting, within the corresponding 30 Monte Carlo experiments, the minimum NLE values associated to each approximated non-dominated front for each algorithm and scenario. Fig. 6 shows the results, again in box plot fashion. The results elucidate a quite stable behavior of both algorithms, as shown by the equally balanced lower and upper quartiles, whiskers close to the quartiles boundaries and a relatively low number of outliers. However, it can be observed that the outliers of PAES are in general more distant from the medians, being lower for NSHS. Again, for verifying whether the distributions are statistically different, we have performed a two-sided rank sum test (Wilcoxon test) with a confidence interval of 95%, and exploited the same color notation as before for the boxes, to highlight the best median (white box) and whether the worse median is statistically significant (black box) or not (gray box). We observe that only for topology 4 PAES outperforms NSHS with statistical significance. For all the remaining topologies NSHS outperforms PAES: in particular NSHS gave better results with statistical significance in all the remaining scenarios, except for

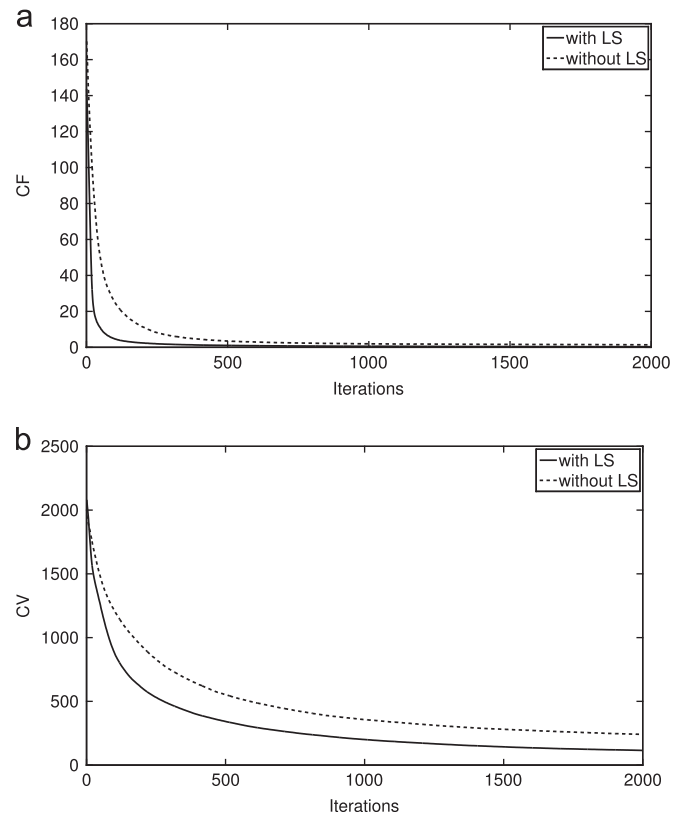
**Table 2**

Mean/minimum/standard deviation statistics for the NLE values associated to the topologies of the dominated fronts of the PAES scheme in Vecchio et al. (in press) and the NSHS approach herein proposed.

#	R	PAES	Proposed NSHS
1	0.13	33.05/19.67/13.54	<b>29.11/15.62/8.82</b>
2		29.86/20.93/5.54	<b>28.45/14.34/10.57</b>
3		24.08/18.68/2.42	<b>24.00/18.49/3.75</b>
4		<b>18.43/14.30/3.85</b>	20.47/14.75/3.46
5	0.15	19.65/14.46/4.17	<b>15.80/9.94/6.58</b>
6		18.27/11.82/4.20	<b>14.14/10.29/1.84</b>
7		17.85/14.18/2.49	<b>15.27/10.67/2.56</b>
8		18.08/14.59/2.14	<b>15.95/13.93/1.45</b>
9	0.17	14.95/11.75/2.01	<b>11.58/8.04/3.11</b>
10		17.22/13.02/2.16	<b>13.92/10.01/1.94</b>
11		14.31/11.28/1.33	<b>12.79/9.26/1.31</b>
12		13.09/10.70/0.94	<b>9.88/6.51/2.17</b>



**Fig. 6.** Box plot of the distributions of the best NLE values obtained by PAES and NSHS (P and HS, respectively) in the 12 network topologies.



**Fig. 7.** Evolution of the fitness functions with and without the application of the local search procedure.

scenarios 1 and 2 where the null hypothesis of equal medians could not be rejected.

Finally, in order to evaluate the benefits introduced by the local search procedure, Fig. 7 depicts the behavior of the CF and CV fitness functions, averaged over the harmony memory at each iteration, for two different runs of NSHS executed on scenario 1, with and without the application of the local search procedure, respectively. It can be observed that the introduction of the local search procedure positively affects the convergence of both the metrics. We have verified that this positive effect does not depend on the specific topology, and is directly reflected also in the NLE metric: for scenario 1, the statistical (mean/min/std) NLE performance degrades from 29.11/15.62/8.82 to 88.70/56.19/19.09 when the local search procedure is removed from the NSHS algorithm.

## 5. Concluding remarks

In this paper we have presented a novel multi-objective heuristic localization technique for wireless sensor networks based on the harmony search algorithm. The proposed approach is further aided by a local search procedure that aims at alleviating the so-called flip ambiguity phenomenon. The proposed algorithm exploits the information on the node connectivity by imposing geometrical constraints so as to bound the areas where nodes can be placed. Through extensive computer simulations, it is shown that our approach embodies a cost-effective localization scheme that outperforms, in terms of multi-objective quality and localization accuracy, a state-of-the-art localization approach based on the PAES algorithm, for the majority of the simulated scenarios.

## Acknowledgments

This work has been supported in part by the Spanish Ministry of Science and Innovation through the CONSOLIDER-INGENIO 2010 (CSD200800010) funding program and the projects ECO2010-22065-C03-02 and TEC2011-28250-C02-02 (ACORDY).

## References

- Akyildiz, I.F., Su, W., Sankarasubramanian, Y., Cayirci, E., 2002. Wireless sensor networks: a survey. *Comput. Networks* 38, 393–422.
- Alippi, C., Vanini, G., 2006. A RSSI-based and calibrated centralized localization technique for wireless sensor networks. In: *Proceedings of Fourth IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 301–305.
- Aspnes, J., Goldenberg, D., Yang, Y., 2004. On the computational complexity of sensor network localization. In: *Algorithmic Aspects of Wireless Sensor Networks*. Lecture Notes in Computer Science, vol. 3121. Springer-Verlag, pp. 32–44.
- Biswas, P., Liang, T.C., Toh, K.C., Ye, Y., Wang, T.C., 2006. Semidefinite programming approaches for sensor network localization with noisy distance measurements. *IEEE Trans. Autom. Sci. Eng.* 3 (4), 360–371.
- Biswas, P., Ye, Y., 2004. Semidefinite programming for ad-hoc wireless sensor network localization. In: *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*. ACM Press, pp. 46–54.
- Bleuler, S., Laumanns, M., Thiele, L., Zitzler, E., 2003. PISA—a platform and programming language independent interface for search algorithms. *Evolutionary Multi-Criterion Optimization*. Lecture Notes in Computer Science, vol. 2632. Springer-Verlag, pp. 494–508.
- Coello, C.A., Lamont, G.B., Van Veldhuizen, D.A., 2006. *Evolutionary Algorithms for Solving Multi-Objective Problems* (Genetic and Evolutionary Computation). Springer-Verlag, New York Inc.
- Costa, J.A., Patwari, N., Hero, A.O., 2006. Distributed weighted-multidimensional scaling for node localization in sensor networks. *ACM Trans. Sensor Networks* 2 (1).
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6, 182–197.
- Del Ser, J., Matinmikko, M., Gil-Lopez, S., Mustonen, M., 2012. Centralized and distributed spectrum channel assignment in cognitive wireless networks: a harmony search approach. *Appl. Soft Comput.* 12 (2), 921–930.
- Del Ser, J., Bilbao, M.N., Gil-Lopez, S., Matinmikko, M., Salcedo-Sanz, S., 2011. Iterative power and subcarrier allocation in rate-constrained orthogonal multicarrier downlink systems based on hybrid harmony search heuristics. *Eng. Appl. Artif. Intell.* 24 (5), 748–756.
- Forsati, R., Haghghat, A.T., Mahdavi, M., 2008. Harmony search based algorithms for bandwidth-delay-constrained least-cost multicast routing. *Comput. Commun.* 31 (10), 2505–2519.
- Geem, Z.W., Kim, J.H., Loganathan, G.V., 2001. A new heuristic optimization algorithm: harmony search. *Simulation* 76 (2), 60–68.
- Geem, Z.W., Lee, K.S., Park, Y., 2005. Application of harmony search to vehicle routing. *Am. J. Appl. Sci.* 2, 1552–1557.
- Geem, Z.W. (Ed.), 2009. *Music-Inspired Harmony Search Algorithm: Theory and Applications*. Studies in Computational Intelligence, vol. 191. Springer.
- Gil-Lopez, S., Del Ser, J., Olabarrieta, I., 2009. A novel heuristic algorithm for multiuser detection in synchronous CDMA wireless sensor networks. *IEEE Int. Conf. Ultra Modern Commun.*, 1–6.
- Hansen, M.P., Jaskiewicz, A., 1998. Evaluating the Quality of Approximations to the Non-dominated Set. Technical Report IMM-REP1998-7, Technical University of Denmark.
- Hu, L., Evans, D., 2004. Localization for mobile sensor networks. In: *Proceedings of the 10th International Conference on Mobile Computing and Networking*, pp. 45–57.
- Jackson, B., Jordan, T., 2005. Connected rigidity matroids and unique realizations of graphs. *J. Combinatorial Theory Ser. B* 94-1, 1–29.
- Ji, X., Zha, H., 2004. Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling. In: *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies*, pp. 2652–2661.
- Kalyanmoy, D., Agrawal, S., Pratap, A., Meyarivan, T., 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197.
- Kannan, A.A., Fidan, B., Mao, G., 2010. Analysis of flip ambiguities for robust sensor network localization. *IEEE Trans. Veh. Technol.* 59 (4), 2057–2070.
- Kannan, A.A., Mao, G., Vucetic, B., 2005. Simulated annealing based localization in wireless sensor network. In: *Proceedings of the IEEE Conference on Local Computer Networks*. IEEE Computer Society, pp. 513–514.
- Kannan, A.A., Mao, G., Vucetic, B., 2006. Simulated annealing based wireless sensor network localization with flip ambiguity mitigation. In: *Proceedings of the 63rd IEEE Vehicular Technology Conference*, pp. 1022–1026.
- Knowles, J.D., Corne, D.W., 2000. Approximating the nondominated front using the pareto archived evolution strategy. *Evol. Comput.* 8 (2), 149–172.
- Ko, Y.B., Vaidya, N.H., 2000. Location-aided routing (LAR) in mobile ad hoc networks. *Wireless Networks* 38-4, 307–321.
- Landa-Torres, I., Gil-Lopez, S., Salcedo-Sanz, S., DelSer, J., Portilla-Figueras, J.A., 2011. A novel grouping harmony search algorithm for the multiple-type access node location problem. *Expert Syst. Appl.* 39 (5), 5262–5270, (2012).
- Liang, T.C., Wang, T.C., Ye, Y., 2004. A Gradient Search Method to Round the Semidefinite Programming Relaxation for Ad hoc Wireless Sensor Network Localization. Stanford University Technical Report.
- Liao, T.W., 2010. Two hybrid differential evolution algorithms for engineering design optimization. *Appl. Soft Comput.* 10 (4), 1188–1199.
- Liu, T., Bahl, P., Chlamtac, I., 1998. Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. *IEEE J. Sel. Areas Commun.* 16 (6), 922–936.
- Niculescu, D., Nath, B., 2003. Ad-hoc positioning system (APS) using AoA. In: *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1734–1743.
- Savvides, A., Han, C.C., Srivastava, M., 2001. Dynamic fine-grained localization in ad-hoc networks of sensors. In: *7th ACM International Conference on Mobile Computing and Networking*, pp. 166–179.
- Tseng, P., 2007. Second-order cone programming relaxation of sensor network localization. *SIAM J. Optim.* 18 (1), 156–185.
- Vecchio, M., Lopez-Valcarce, R., Marcelloni, F. A two-objective evolutionary approach based on topological constraints for node localization in wireless sensor networks. *Appl. Soft Comput.*, in press, <<http://dx.doi.org/10.1016/j.asoc.2011.03.012>>, 12 (7), 1891–1901.
- Wang, Z., Zheng, S., Ye, Y., Boyd, S., 2008. Further relaxations of the semidefinite programming approach to sensor network localization. *SIAM J. Optim.* 19 (2), 655–673.
- Yang, Z., Liu, Y., 2011. Understanding node localizability of wireless ad-hoc and sensor networks. *IEEE Trans. Mobile Comput.* 99 (1).
- Zhang, R., Hanzo, L., 2009. Iterative multiuser detection and channel decoding for DS-SSMA using harmony search. *IEEE Signal Process. Lett.* 16 (10), 917–920.
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Grunert da Fonseca, V., 2003. Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Trans. Evol. Comput.* 7 (2), 117–132.
- Zitzler, E., Thiele, L., 2009. *Multiobjective Optimization Using Evolutionary Algorithms—A Comparative Case Study*, vol. 1498. Springer, pp. 292–301.