Genetic Algorithm and Simulated Annealing Approach to Sensor Network Localization

Michał Marks[1] and Ewa Niewiadomska-Szynkiewicz[1,2]

[1] Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw, Poland, email: mmarks@elka.pw.edu.pl
[2] Research Academic Computer Network (NASK), Warsaw, Poland, e-mail: ens@ia.pw.edu.pl

Abstract. The paper[1] provides an overview of measurement techniques in sensor networks localization and optimization algorithms based on these measurements for estimation the physical location of nodes with unknown location. A novel localization methods, i.e., two phase algorithms based on simulated annealing and genetic algorithm are described. The numerical results presented and discussed in the final part of the paper show that these novel schemes give accurate and consistent location estimates of the nodes in the network.

1 Introduction to Sensor Network Localization

Recently, wireless sensor networks (WSNs) are deployed in various environments and are used in large number of practical applications, such as environmental information (light, pollution, temperature, sound levels, etc.), traffic or health monitoring, intrusion detection, etc. Typical sensor network consists of a large number of nodes – densely deployed sensor devices. Nodes networked through wireless must gather local data and communicate with other nodes. The information sent by a given sensor is relevant only if we know what location it refers to. Location estimation allows applying the geographic-aware routing, multicasting and energy conservation algorithms. It makes self-organization and localization capabilities one of the most important requirement in sensor networks.

The simplest way to determine a node location is to equip this node with a global positioning system (GPS) or install it at point with known coordinates. Because of the cost, size of sensors and constraints on energy consumption most sensors usually do not know their locations, only a few nodes, called *"anchors"* are equipped with GPS adapters. Location of other nodes, called *"non-anchors"*, are unknown. In such model the techniques that estimate the locations of *"non-anchors"* based on information about positions of *"anchors"* are utilized.

The paper is organized as follows. An overview of localization methods is presented in section 2, starting from measurement techniques, simple single-hop algorithms up to multi-hop distance-based methods. The formulation of the distance based localization problem is provided in section 3. Section 4 describes a novel method proposed by authors

[1] This work is partially supported by Warsaw University of Technology Research Program grant.

– two phase stochastic approach to localization problem solution. The simulation results are presented and discussed in section 5.

2 Sensor Network Localization Techniques

The objective of the location estimation method is to estimate the position (coordinate) of sensor nodes with respect to a set of nodes with known global location information. Wireless sensor network localization is a complex problem that can be solved in different ways [5]. Generally, the proposed solutions are based on signal processing and algorithms transforming measurements into the coordinates of the nodes in the network.

2.1 Measurement Techniques

A number of measurement techniques are available in wireless sensor networks. They can be classified into the following categories: Angle of Arrival measurements (AOA), Distance related measurements, Received Signal Strength (RSS) profiling techniques.

AOA, is a technique for determining the direction of propagation of a radio-frequency wave incident on an antenna array. Two subclasses can be distinguished: making use of the receiver antenna's amplitude response and those making use of the receiver antenna's phase response. AoA calculates the direction by measuring the Time Difference of Arrival (TDOA) at individual elements of the array. TDOA measurement is made by measuring the difference in received phase at each element in the antenna array. The delay of arrival at each element is measured directly and converted to an AoA measurement. Propagation time based measurements such as: one-way propagation time, roundtrip propagation time, RSS and TDOA are included in distance related measurements techniques. The idea of RSS profiling is to construct a form of map of the signal strength behavior in the coverage spatial domain. The map can be obtain based on off-line measurements or on-line sniffing devices.

2.2 One-hop and Multi-hop Localization Techniques

In one-hop localization techniques the non-anchor to be localized has to be one-hop neighbor of a sufficient number of anchors. In the literature different approaches based on AOA, TDOA, RSS measurements and hybrid techniques are considered. The most popular technique based on distance measurement uses GPS. The GPS space segment consists of 24 satellites in the medium earth orbit with an orbital inclination of 55 degrees. In the case of methods based on RSS profiling each non-anchor node uses the signal strength measurements to determine its own RSS vector and sends it to the central station. The central station estimates the location of the non-anchor node based on the obtained data, using probabilistic methods or some kind of nearest neighbor-based technique. RSS-profiling based methods produce relatively small location estimation errors in comparison to distance-based approaches [1].

The other group of localization algorithms are multi-hop techniques, in which the non-anchor nodes do not have to be one-hop neighbors of the anchors. They can be considered into two main classes: *connectivity based* and *distance based*. The connectivity based algorithms use only connectivity information to locate the entire sensor network. The distributed algorithm – Ad Hoc Positioning System (APS) developed by Niculescu [8] is an example of this approach. Centralized connectivity based localization algorithm was proposed by Doherty [7] and Shang [9].

Recently, the most popular are distance based multi-hop localization algorithms. They use inter-sensor distance measurements in the sensor network to estimate the locations of the non-anchor nodes. Similarly to connective based, centralized and distributed variants are provided. In the case of centralized ones a single central processor is used to collect all local distance data provided by all nodes in the network. A map of nodes location in the entire network is generated based on available information. Distributed algorithms rely on self-organization of nodes in a sensor network. Each non-anchor node estimates its location based on measured distances and local data gathered from its neighbors.

In our paper we will focus on centralized methods. Even though they are less complicated than distributed ones they are likely to provide more accurate location estimates. Three main approaches for designing centralized distance based algorithms are provided in the literature: multidimensional scaling (MDS) [9, 4], semi-definite programming (SDP) [2, 10] and stochastic optimization [5, 6].

3 Distance based localization problem formulation

The mathematical model of the distance based localization is as follows. Let us consider the network of N nodes (sensors), among them there are M anchor nodes with known location. Our aim is to estimate the coordinates $(\hat{x}_i, \hat{y}_i), i = M+1, \ldots, N$ of $N - M$ non-anchors. We can formulate optimization problem with the performance measure J considering estimated and measured Euclidean distances of all neighbor nodes

$$\min_{\hat{x},\hat{y}} \Big\{ J = \sum_{i=M+1}^{N} \sum_{j \in N_i} (\hat{d}_{ij} - d_{ij})^2 \Big\}, \qquad d_{ij} \leq R, \qquad j \in N_i \tag{1}$$

where $\hat{d}_{ij} = \sqrt{(\hat{x}_i - \hat{x}_j)^2 + (\hat{y}_i - \hat{y}_j)^2}$, $(\hat{x}_i, \hat{y}_i)$ are estimated coordinates of node i, $(\hat{x}_j, \hat{y}_j)$ estimated coordinates of one hop neighbor j of node i, $\hat{d}_{ij}$ estimated distance between nodes i and j, d_{ij} measured distance between nodes i and j, N_i a set of neighbors of node i, R a fixed parameter called radio range.

The measured distance between two neighbor nodes is produced by measurement methods described in section 2.1. These methods involve measurement uncertainty; each distance value d_{ij} represents the true physical distance corrupted with a noise describing the uncertainty of the distance measurement. Figure 1 shows the influence of measurement uncertainty on the accuracy of a given node localization. Consider the example presented in Fig.1. Our goal is to estimate the coordinates of node 4 based on distance measurements d_1, d_2 and d_3. In the case of poor quality measurements the calculated set of expected positions of the non-anchor node is different than the true ones. Due to measurement uncertainty it is difficult to find a good metric to compare the results obtained using different localization methods. To compare the performance of the tested algorithms we used the mean error between the computed and the actual unknown location of the nodes in the network, defined as follows

$$LE = \frac{1}{N-M} \cdot \frac{\sum_{i=M+1}^{N} ((\tilde{x}_i - \hat{x}_i)^2 + (\tilde{y}_i - \hat{y}_i)^2)}{R^2} \cdot 100\% \tag{2}$$

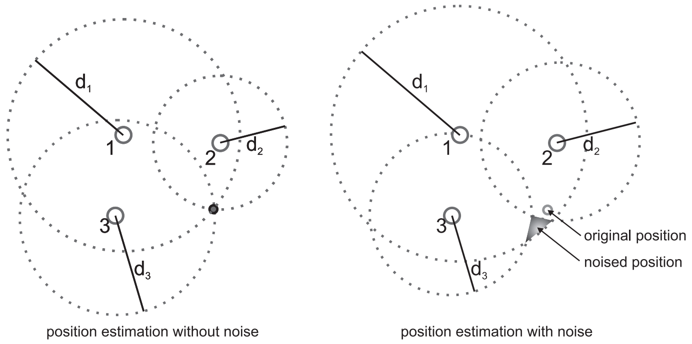

Figure 1. Influence of measurement error on quality of node's estimated location

where $(\tilde{x}_i, \tilde{y}_i)$ is true location of sensor node i, $(\hat{x}_i, \hat{y}_i)$ estimated location of sensor node i and R radio range. The location error LE is expressed as a percentage error. It is normalized with respect to the radio range to allow comparison of results obtained for different size and range networks.

4 Two phase localization method

We propose two variants of the two phase multi-hop localization method. The proposed technique is based on distance measurement and multi-hop localization. The algorithm operates in two phases. In the first phase the auxiliary solution (initial localization) is produced. The second phase is the crucial one. The solution of the first phase is modified by applying stochastic global optimization methods. In this paper we consider two well-known techniques – simulated annealing (SA) and genetic algorithm (GA). Both these heuristics are implemented as a computer simulation of stochastic proces. The differences between them are such that SA is based on a point-to-point transformation and GA transforms a population of points.

4.1 Phase I

The first phase of the algorithm is similar to single-hop distance based techniques. Only the nodes with three anchor neighbors are localized. All nodes are divided into two groups: group A – M nodes with known location (in the beginning only the anchor nodes) and group B – nodes with unknown location. In each step of the algorithm node i, where $i = M + 1, \ldots, N$ from the group B is chosen. Next, the three nodes from the group A that are within node i radio range are randomly selected. If such nodes exist the location of node i is calculated based on true inter-nodes distances between three nodes selected from group A and the measured distances between node i and these three nodes. The localized node i is moved to the group A. Otherwise, another node from the group B is selected and the operation is repeated. The algorithm stops when there are no more nodes that can be localized based on the available information about all nodes localization.

4.2 Phase II

Due to distance measurement uncertainty the coordinates are estimated with non-zero errors as defined in (2) (see Fig. 1). Such calculated solution is improved in phase II. The stochastic optimization algorithms: SA and GA are applied to increase the accurancy of location estimation.

Simulated annealing (SA)

Simulated annealing method was implemented according to the algorithm described in [5]. It is a classical version of SA with one modification – the cooling process is slowed down. At each value of the coordinating parameter T (temperature), not one but $q \cdot N$ non-anchor nodes are randomly selected for modification (where N denotes the number of sensors in the network and q is a reasonably large number to make the system into thermal equilibrium). Coordinate estimations of chosen nodes are perturbed with a small displacement of the distance Δd in a random direction. The structure of the SA algorithm is presented in the Fig. 2.

T = initial temperature
(Δd) = initial move distance
WHILE (final temperature not met)
{
FOR $i = 1$ to $(q \cdot N)$
{
pick a node to perturb
DO p times
{
generate a random perturbation to a node's estimated location
evaluate the change in cost function, $\Delta(CF)$
if $(\Delta(CF) \leq 0)$
//downhill move $\Rightarrow$ accept it
accept this perturbation and update the configuration system
else
//uphill move $\Rightarrow$ accept with probability
pick a random probability rp = uniform(0,1)
if $(rp \leq exp(-\Delta(CF)/T))$
accept this perturbation and update the configuration system
else
reject this perturbation and keep the old configuration system
}
}
$T_{new} = \alpha \cdot T_{old}$
$(\Delta d)_{new} = \beta \cdot (\Delta d)_{old}$

Figure 2. Simulated annealing algorithm, Mao [5].

Task configuration. The goal of this task is to localize $N - M$ non-anchor nodes with coordinates (x_i, y_i), $i = M + 1, \ldots, N$, placed in the domain. The initial location of all nodes is determined in phase I of the algorithm.

Moving operation. In each iteration of the algorithm a new solution is calculated. The node is randomly selected and is moved in random direction at distance Δd. The value of Δd depends on the control parameter T, for small value of T the distance Δd is restricted by shrinking factor $\beta < 1$, $(\Delta d)_{new} = \beta \cdot (\Delta d)_{old}$.

Performance measure. The performance measure describes the quantitative measure of estimation quality and is defined in (1).

Cooling scheme. The simple cooling scheme is proposed: $T_{new} = \alpha \cdot T_{old}$.

Genetic algorithm (GA)

Task configuration. The goal of this task is to localize $N - M$ non-anchor nodes placed in the domain. The abstract representations of candidate solutions called chromosomes are vectors of random variable – coordinates of all non-anchor nodes: $[x_M, y_M, x_{M+1}, y_{M+1}, \ldots, x_N, y_N]$, $x_i, y_i \in \Re$.

Initial population. The initial population consists of 200 chromosomes, the genes of which (initial coordinates of all nodes) were determined in the first phase of the algorithm.

Performance measure. Similarly to SA algorithm the performance measure (fitness function) is defined in (1).

Selection. The tournament selection of size $q = 2$ is used.

Crossover. Discrete recombination similar to elements exchanging applied to binary vectors is used with one modification – both coordinates of a given node are recombined simultaneously.

Mutation. The simple mutation operator is used. The components of chromosome are modified by adding a vector of generated $2 \cdot (N - M)$ Gaussian random variables.

4.3 Correction of points location

From the experiments it was observed that in case of both methods applied in the second phase of the algorithm, the increased value of the location error (2) is usually driven by incorrect location estimates calculated for a few nodes. This phenomena is depicted in the figure 3. Let us consider the nodes A, B, C with known locations and

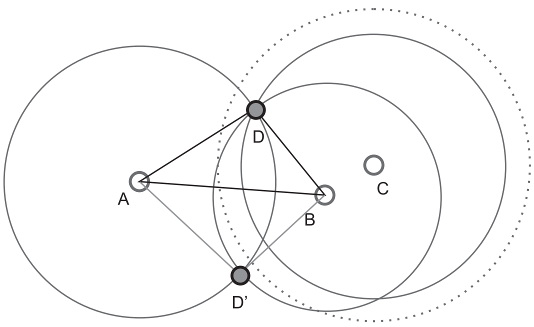

Figure 3. Correct and incorrect node's locations

the measured distances between the node D to be localized and nodes A, B, C. We can determine the coordinates of the nodes D and D' based on the distances d_{AD} and d_{BD}. As a final result the location D will be selected, but in the case of non-zero measurement errors the incorrect location, i.e. D' will be chosen. The low level measurement error may involve the inaccurate location estimates.

To alleviate this phenomena a new functionality was added to the proposed localization algorithm. Its objective is to make correction in the location estimates when the location error (2) exceeds the assumed threshold value. Additional constraints concerned with this threshold value are introduced to the optimization problem and the following algorithm is performed. Three nodes from the group of neighbors of a given node i that violates the less number of constraints are randomly selected. Next, the location of the node i is determined based on the selected nodes locations. If the new location is more accurate, i.e. violates less number of constraints, it replaces the previous one. The correction operation is repeated until all constraints are fulfilled or the assumed number of iterations is achieved.

5 Numerical results

In order to evaluate the proposed approaches to sensor network localization many numerical tests were performed. Sensor networks with 200 and more nodes with randomly generated positions were considered. The calculations were carried out on the machine Intel Core2 Duo E6600 - 2.4GHz, 2GB RAM. The average results provided by different localization algorithms and obtained during five runs of each task are presented in tables and figures.

In our numerical experiments the measured distance between neighboring nodes i and j was disturbed by introducing Gaussian noise with a mean of 0 and a standard deviation of 1 added to the true distance $\tilde{d}_{ij}$.

$$d_{ij} = \tilde{d}_{ij} * (1.0 + randn() * nf) \tag{3}$$

where nf denotes a noise factor.

Two stochastic optimization algorithms: SA and GA were applied in the second phase of the algorithm to improve the nodes' location estimation. The results obtained while using SA method are collected in Table 1 and in Fig. 4. The calculation time was equal 4 seconds.

Figure 4 depicts the two phases of SA based algorithm. The presented sensor network consists of 200 nodes, with 20 anchor nodes (marked by rhombus) and 180 non-anchor nodes (marked by circle). The estimated positions of non-anchor nodes are marked by stars. The localization error is denoted by lines connecting the true and estimated locations. The assumed transmission range was equal 0.18, noise factor of measurement in (3) was taken as 10%.

Next, results obtained for two phase SA and GA based localization techniques were compared with those obtained applying Semidefinite Programming (SDP) proposed by Ye in [10]. The numerical results of GA and SDP methods are presented in Fig. 5. From this figure we can see that the GA method gives more accurate location estimates w.r.t. SDP, but with longer computation time.

Table 2 presents the results obtained for three methods SA, GA and SDP. The location errors and computation times are compared. From this table we can see that SA and GA estimate the location of nodes quite accurately, with the location error less than 1%. The location error for the SDP algorithm and the same network was about 15%.

In comparison to the results obtained for SDP and the results for one phase simulated annealing based localization algorithm presented in [3] , atwo phase simulated annealing based algorithm seems to be very promising. The location estimates are very accurate and the computation time is three times smaller than in the case of SDP.

In the next series of experiments the simulations of sensor networks with different number of nodes were performed. From the results presented in Table 3 we can see that

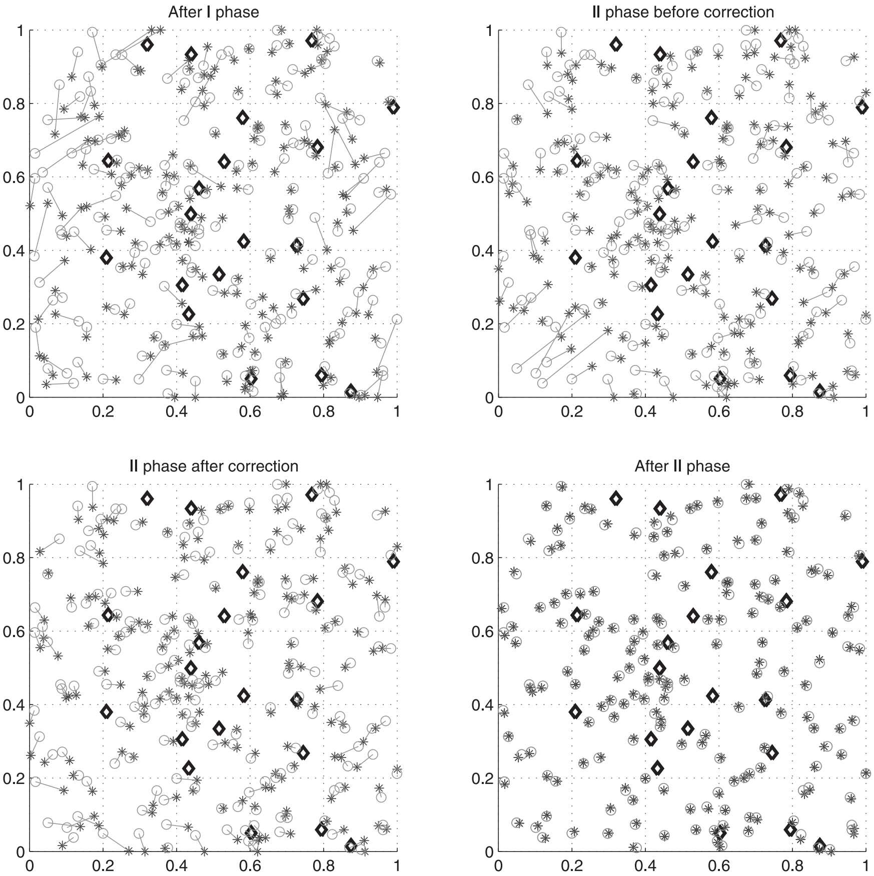

Figure 4. Simulated annealing in network nodes localization

Table 1. Location errors in two phases of the algorithm

Phase		Performance value	Location error
I	phase I final solution	7.26	10.64
II	before correction	2.00	6.25
	after correction	2.00	3.04
	phase II final solution	0.43	0.14

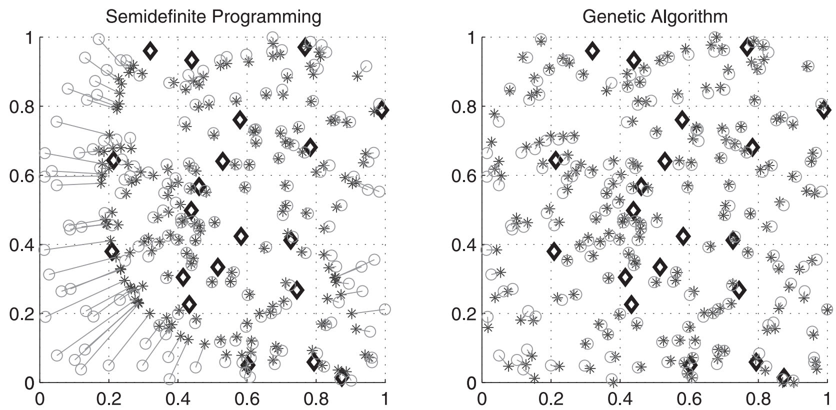

Figure 5. Results comparison: genetic algorithm and SDP method

Table 2. Computing time and Location errors. Different methods

Method	Location error	Computation time [s]
Semidefinite Programming	15.16	13.86
Genetic algorithm	0.62	34.00
Simulated annealing	0.11	4.00

the computation time increases proportionally to the square of the size of the network.

Table 3. Location error and computation times for different network sizes

Number of nodes	Location error	Computation time [s]
200	0.11	4.00
500	0.15	20.00
1000	0.29	97.00

It should be pointed here that all considered optimization methods were evaluated based on the value of the location error defined in (2) and assuming the knowledge about the true location of a given node. The performance measure that is minimized in the optimization problem is different, and defined in (1). So, it is obvious that in some experiments the results which give better value of (1) give less accurate location estimates, i.e. bigger location error (2), see table 4.

Table 4. Several runs of genetic algorithm

Run	Localization error	Performance measure
1	0.27	0.807810
2	0.40	0.807342
3	0.53	0.791686
4	0.65	0.797158
5	0.74	0.823304

In summary, the criterion based on location error let us evaluate the quality of the coordinates estimation but not the efficiency of the applied optimization algorithm.

6 Summary and conclusions

In this paper we described the application of stochastic global optimization techniques to wireless sensor network localization problem. We demonstrated that the proposed two phase simulated annealing and genetic algorithm based methods provide quite accurate location estimates in the sensible computing time. Finally, it should be pointed that accurate self-organization and localization are fundamental requirements in high performance ad hoc networks.

Bibliography

[1] Paramvir Bahl and Venkata N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *INFOCOM (2)*, pages 775–784, 2000.

[2] Pratik Biswas and Yinyu Ye. Semidefinite programming for ad hoc wireless sensor network localization. In *IPSN '04: Proceedings of the third international symposium on Information processing in sensor networks*, pages 46–54, New York, NY, USA, 2004. ACM Press.

[3] Brian D. O. Anderson Guoqiang Mao, Baris Fidan. *Sensor Network and Configuration: Fundamentals, Techniques, Platforms and Experiments*, pages 281–316. Springer-Verlag, November 2006.

[4] X. Ji. Sensor positioning in wireless ad-hoc sensor networks with multidimensional scaling, 2004.

[5] Anushiya A. Kannan, Guoqiang Mao, and Branka Vucetic. Simulated annealing based localization in wireless sensor network. In *LCN '05: Proceedings of the The IEEE Conference on Local Computer Networks 30th Anniversary*, pages 513–514, Washington, DC, USA, 2005. IEEE Computer Society.

[6] Anushiya A. Kannan, Guoqiang Mao, and Branka Vucetic. Simulated annealing based wireless sensor network localization with flip ambiguity mitigation. In *63rd IEEE Vehicular Technology Conference*, page 1022– 1026, 2006.

[7] K. Pister L. Doherty and L. El Ghaoui. Convex postion estimation in wireless sensor networks. In *IEEE INFOCOM*, pages 1655–1663, 2001.

[8] D. Niculescu and B. Nath. Ad hoc positioning system (aps), 2001.

[9] Yi Shang, Wheeler Ruml, Ying Zhang, and Markus Fromherz. Localization from connectivity in sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 15(11):961–974, 2004.

[10] Stephen Boydy Zizhuo Wang, Song Zheng and Yinyu Ye. Further relaxations of the sdp approach to sensor network localization. September 2006.