Distance-based Localization Scheme for Ad Hoc Network

Michał Marks and Ewa Niewiadomska-Szynkiewicz

Warsaw University of Technology, Institute of Control and Computation Engineering, Warsaw
Research Academic Computer Network (NASK), Warsaw
email: {mmarks, ens}@ia.pw.edu.pl

Abstract. Recently, we proposed a centralized distance-based technique TSA (*Trilateration* & *Simulated Annealing*) that uses a combination of the trilateration method, along with the simulated annealing optimization algorithm for performing localization of nodes in an ad hoc network with static nodes (Wireless Sensor Networks). Our location scheme depends on network nodes transmitting data to a central computer, where calculation is performed to determine the estimated location of each node. Although it achieves high accuracy in estimating sensors' locations, speed of the method can not be satisfactory for large scale practical applications. It strongly depends on the values of the parameters specific to the algorithm. In this paper[1] we report the results of numerical tests performed for various values of these parameters. On the other hand we propose the distributed version of our technique where each network node estimates its position based on only local data gathered from its neighbors. Finally, we present the comparative study of centralized and distributed schemes.

1 Introduction and localization problem formulation

Wireless Sensor Networks (WSN) used to monitoring and controlling the environment surrounding them are useful only if it is possible to correlate data transmitted to a base station in time and in space. This is the reason why network nodes' localization is a fundamental tool for ad hoc networks. Because of constraints on the cost and size of network nodes, energy consumption and the randomly deployment of sensors, most nodes do not know their position. Recently some techniques [2] for assigning geographic coordinates to each node in an ad hoc network system have been proposed. They allow to estimate nodes location using information transmitted by a set of anchor nodes. An anchor is defined as a node that is aware of its own location, either through GPS adapter or manual deployment at the point with known position. It should be pointed that localization schemes should give the solution in the short time, achieve good accuracy even in the case of unevenly distributed nodes, and scale to large networks.

Two classes of localization methods can be distinguished: connectivity-based and distance-based. Distance-based methods use inter-sensor distance or angle measurements in location calculation. The techniques proposed in literature are based on applying

[1] This work is supported by Ministry of Science and Higher Education grant N N514 416934.

popular optimization methods to nodes localization, such as semidefinite programming (SDP) [1], simulated annealing [3, 4] and genetic algorithm [7]. The connectivity-based methods use only contents of the received messages to locate the entire sensor network. They are based on hop-counting. A survey of these approaches can be found in [5, 9, 10]. Both, connectivity-based and distance-based techniques have some weaknesses, i.e., distance-based algorithms require the additional costly equipment, connectivity-based algorithms are cost effective but their performance is usually worse. The method described in this paper belongs to the class of centralized distance-based algorithms.

Hence, let us formulate the mathematical model of the localization problem for distance-based approaches. There is a network of N nodes (sensors) in $\Re^k$ with bidirectional communication constraints as the edges. Positions of M nodes (anchors) are known. The Euclidean physical distance d_{ij} between the ith and jth nodes can be measured if $(i, j) \in N_i$, where $N_i = \{(i, j) : ||x_i - x_j|| = d_{ij} \leq r\}$ denotes a set of neighbors of node i, $x_i \in \Re^k$ and $x_j \in \Re^k$ true locations of nodes i and j, r is a fixed parameter called transmission range (radio range). Assuming that we have the measurements of distances between all pairs of nodes we can formulate the model of the localization problem that minimizes the sum of squares of errors in sensor positions for fitting the distance measurements.

$$\min_{\hat{x}} \left\{ J = \sum_{i=M+1}^{N} \sum_{j \in N_i} (\hat{d}_{ij} - \tilde{d}_{ij})^2 \right\} \tag{1}$$

where $\hat{d}_{ij}$ denotes an estimated distance between nodes i and j, $\hat{d}_{ij} = ||\hat{x}_i - \hat{x}_j||$, $\hat{x}_i \in \Re^k$ an estimated position of node i and $\hat{x}_j \in \Re^k$ an estimated position of a neighbor of node i, $\tilde{d}_{ij}$ a measured distance between nodes i and j. The measured distance $\tilde{d}_{ij}$ between two neighbor nodes is produced by measurement methods described in literature [6]. These methods involve measurement uncertainty; each distance value $\tilde{d}_{ij}$ represents the true physical distance d_{ij} corrupted with a noise describing the uncertainty of the distance measurement. For the purpose of numerical experiments we supposed that this disturbance is described by introducing Gaussian noise with a mean of 0 and a standard deviation of 1 added to the true physical distance d_{ij}.

$$\tilde{d}_{ij} = d_{ij} \cdot (1.0 + randn() \cdot nf) \tag{2}$$

where nf denotes a noise factor.

To evaluate the performance of tested algorithms we used the mean error between the estimated and the true location of the non-anchor nodes in the network, defined as follows

$$LE = \frac{1}{N - M} \cdot \frac{\sum_{i=M+1}^{N} (||\hat{x}_i - x_i||)^2}{r^2} \cdot 100\% \tag{3}$$

where x_i denotes the true position of the sensor node i in the network, $\hat{x}_i$ estimated location of the sensor node i and r the radio transmission 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.

2 TSA scheme description

In [7] we proposed the localization technique that uses a combination of the geometry of triangles (trilateration), along with the stochastic optimization. This algorithm operates in two phases. In the first phase the initial localization is provided. Trilateration uses the known locations of a few anchor nodes, and the measured distance between a given non-anchor and each anchor node. To accurately and uniquely determine the relative location of a non-anchor on a 2D plane using trilateration alone, generally at least three neighbors with known positions are needed. Hence, all nodes are divided into two groups: group A containing M nodes with known location (in the beginning only the anchor nodes) and group B of 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, 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 inter-nodes distances between three nodes selected from the 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 first phase stops when there are no more nodes that can be localized based on the available information about all nodes localization. It switches to the second phase.

Due to the distance measurement uncertainty the coordinates calculated in the first phase are estimated with non-zero errors. Hence, the solution of the first phase is modified by applying stochastic optimization methods. Two techniques, i.e., simulated annealing and genetic algorithm were considered. The numerical results obtained for simulated annealing SA were much more promising (see [7, 8]) w.r.t. calculated location accuracy and speed of convergence. So, we decided to focus on this approach. We called this method TSA (*Trilateration* & *Simulated Annealing*).

From the numerical experiments it was observed that the increased value of the location error is usually driven by incorrect location estimates calculated for a few nodes. The additional functionality (correction) was introduced to the second phase to remove incorrect solutions involved by the distances measurement errors. The additional constraints were introduced to the optimization problem. The detail description of the correction algorithm can be found in [7].

3 TSA scheme evaluation

We performed many numerical tests to cover a wide range of network system configurations including: size of the network (200 − 10000 nodes), anchor nodes deployment (evenly and unevenly distributed), distance measurement error and computation time. The key metric for evaluating a localization method was the accuracy of the location estimates versus the computation costs. A detailed analysis of performance of the TSA method for networks consisting of 1000 randomly generated nodes is presented in [8]. The results of TSA were compared with those obtained using semidefinite programming (SDP). The summary results for network with 200 evenly and unevenly deployed nodes, TSA and SDP methods are collected in the table 1. We have to point that although we achieved high accuracy in estimating sensors' locations, speed of the method was not satisfactory for large scale networks. From the table 2 we can observe that the computation time increases proportionally to the network dimension. The compromise should

Table 1. Computing time and localization errors; different deployment of anchors

Task	Method	Localization error (LE) [%]	Computation time [s]
anchors evenly distributed	SDP	1.32	17.0
	TSA	0.14	3.0
anchors unevenly distributed	SDP	98.34	9.5
	TSA	0.24	3.0

Table 2. Localization error and computation times for different network sizes

Number of nodes	Localization error (LE) [%]	Computation time [s]
200	0.11	1.4
500	0.15	7.6
1000	0.29	29.4

be made between efficiency and accuracy. To decrease the calculation effort the optimal values of method's parameters have to be estimated. In the simulated annealing implementation used in the second phase of the TSA scheme at each value of the coordinating parameter T (temperature), $P \cdot (N - M)$ non-anchor nodes are randomly selected for modification (where N denotes the number of sensors in the network, M the number of anchors, and P a reasonably large number to make the system into thermal equilibrium). The parameter P plays the important role – it influences the estimated location accuracy and calculation time. The figures 1 and 2 present the results of numerical tests performed for the network with 200 nodes and various values of P.

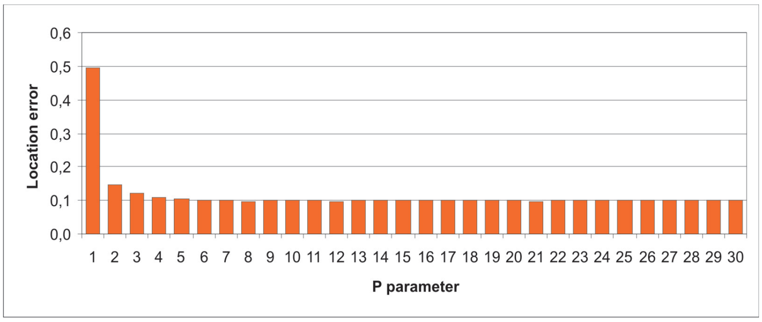

Figure 1. Localization error for various values of P

To calculate the optimal value of the parameter P for a given network we can solve the two-criterion optimization problem

$$\min_{\Delta t, LE} \delta = \left(\frac{\Delta t}{\Delta t_e} - \frac{\Delta t_t}{\Delta t_e}\right)^2 + \left(\frac{LE}{LE_t} - \frac{LE_e}{LE_t}\right)^2 \tag{4}$$

where Δt denotes a calculation time, LE a localization error defined in (3), Δt_t and LE_t are, respectively, a time of calculations for $P = 1$ and corresponding localization error, LE_e and Δt_e acceptable localization error and corresponding calculation time.

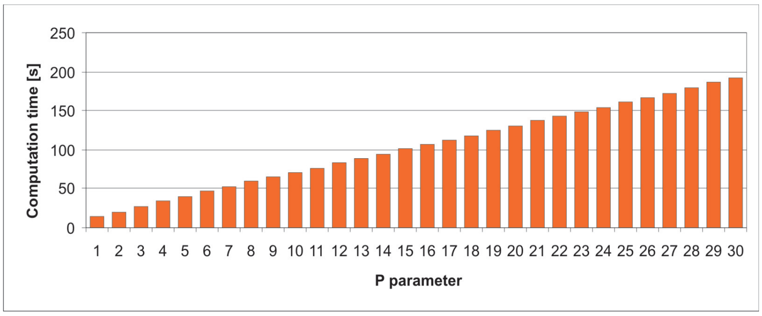

Figure 2. Computation times for various values of P

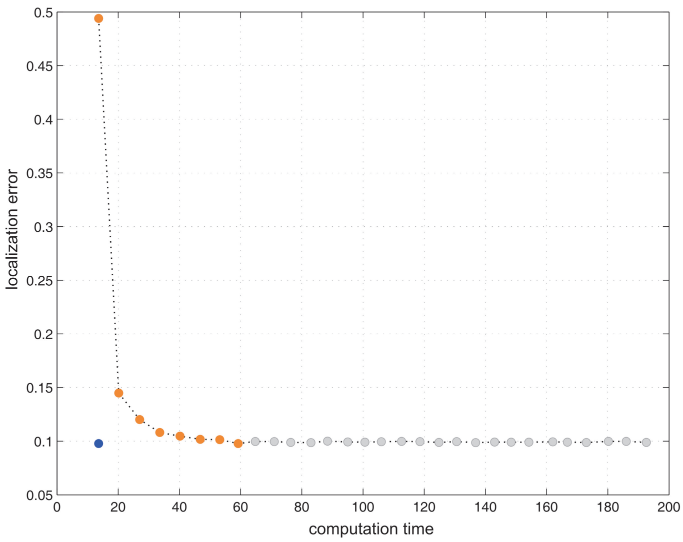

Figure 3. The solution of the problem (4) for the network with 200 nodes

The figure 3 presents the Pareto frontier for the network of 200 nodes. We can see the utopia point (13.6; 0.0978) and eight undominated points. The solution closest to the utopia point (in the Euclidean metric) is equal (33.6;0.1081) and calculated for $P = 4$. The optimal values of the parameter P corresponding to the solutions of the task (4) for different networks are presented in the table 3. Because TSA should be the general

Table 3. Optimal values of the parameter P for different size of network

Number of nodes	200	500	1000	2000	4000
Calculated P	4	4	4	2	3

purpose localization scheme that can be used to solve different dimension problems we suggest to choose the parameter value $P = 4$. The results of calculations performed for network with 200 to 10000 nodes and $P = 4$ are presented in the table 4.

Table 4. Localization errors and computation times for different sizes of network

Number of nodes	Localization error (LE) [%]	Computation time [s]
200	0.1275	0.4
500	0.4124	2.2
1000	0.1387	8.0
2000	0.1081	33.6
4000	0.1086	125.8
5000	0.1581	189.8
10000	0.1193	790.4

4 Distributed version of TSA scheme

The centralized TSA method provides quite accurate location estimates even in the case of unevenly distributed nodes with known positions. However we have to gather the measurements of distances between all pairs of network nodes in a single computer to solve the optimization problem (1). The data transmission to the central station involves time delays, high communication cost and high energy consumption. Because of these disadvantages the centralized techniques can not be acceptable in many applications (e.g. mobile networks and large scale networks).

In contrast to the centralized method we can propose a fully distributed method where computation takes place at every node. Each node is responsible for determining its position using information about neighbors. It offers a significant reduction in computation requirements because the number of neighbors is usually between ten and twenty, so the number of connections is usually a few orders of magnitude less. The use of a fully distributed computation model is also tolerant to node failures, and distributes the communication cost evenly across the sensor nodes. On the other hand, distributed algorithms implementation is often connected with the loss of information. There are two reasons of that: loss of information due to parallel computation and loss of information due to the incomplete network map.

The loss of information due to parallel computation is connected with the way how the optimization in the second phase of TSA method is performed. At each value of the coordinating parameter T (temperature), P times non-anchor nodes ($N - M$ nodes) are randomly selected for modification, i.e., coordinate estimates of chosen nodes are perturbed with small distance Δd in a random direction. Modifications are done in order – the location of current node is determined based on the previous transformations. In distributed algorithm each node have to do P small displacements. This is done in parallel and the information is exchanged between neighboring nodes every P iterations. An example is presented in the figure 4. The considered node (marked by square) updates the location using the information about previous positions of neighbor nodes (light-gray circles). Because all movements are done independently there is no guarantee that the performance value is better after the iteration. The loss of information due to an incomplete network map has the major impact on the correction operation in the TSA method. The correction operation depends on nonconvex neighborhood constraints involved by the transmission range. In the centralized approach, a complete network state (location estimates of all nodes) is available. Due to this information it is possible to detect the situation when the estimated distance between two neighbor nodes is greater than the

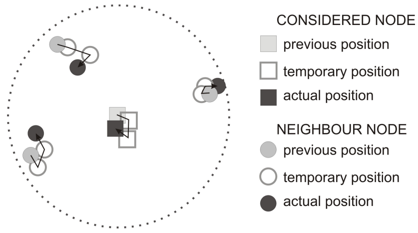

Figure 4. Parallel nodes displacements

transmission range. In the same way it can be detected that the estimated distance between two nodes is less than the transmission range, but these nodes are not in neighborhood. In the distributed approach situation is different. Certainly each node can detect the situation when the estimated distance between it and its neighbor is greater than the transmission range, but it is impossible to find out that from the estimated location appears that the node is closed to another node which is not a true neighbor. In order to overcome this problem we propose to modify the correction operation. In the centralized scheme all neighbors are considered. In distributed we distinguish two versions: basic – only 1-hop neighbors are considered, and extended – both, 1-hop and 2-hop neighbors are considered. In the table 5 the results obtained for networks with evenly and unevenly distributed anchors with 1-hop neighbor and 2-hop neighbor correction are presented. Figure 5 presents the solution quality difference between centralized

Table 5. Localization errors for 1-hop and 2-hop correction. Different deployment of anchors

Task	Localization error (LE) [%]	
	1-hop neighbor correction	2-hop neighbor correction
anchors evenly distributed	0.34	0.31
anchors unevenly distributed	12.84	3.01

and distributed algorithms (using two-hop correction). The obtained results confirm that from the perspective of location estimation accuracy, centralized algorithm provides more accurate location estimates than distributed one. As a final result we can say that for evenly distributed anchors we obtain quite accurate solution using both methods, otherwise the results of location estimation are much worse in case of distributed version of our scheme.

5 Summary and conclusions

We presented and evaluated the localization scheme that combines simple geometry of triangles and simulate annealing technique to determine the location of nodes with unknown positions in the sensor network. We demonstrated that TSA method provides quite accurate location estimates in the sensible computing time. The proposed technique was implemented in a centralized and distributed variants. We discussed the advantages and drawbacks of both approaches. In our future research, we plan to perform experiments using software environments for sensor networks simulation, and finally apply our scheme to the testbed network of sensors in the laboratory.

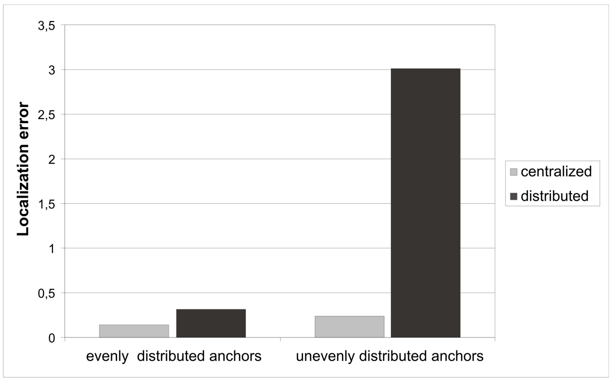

Figure 5. Localization error for centralized and distributed algorithms

Bibliography

[1] P. Biswas and Y. 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.

[2] J. Hightower and G. Borriello. Localization systems for ubiquitous computing. *IEEE Computer*, 34(8), 2001.

[3] A. A. Kannan, G. Mao, and B. 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.

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

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

[6] G. Mao, B. Fidan, and B. D. O. Anderson. Wireless sensor network localization techniques. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 51(10):2529–2553, 2007.

[7] M. Marks and E. Niewiadomska-Szynkiewicz. Genetic algorithm and simulated annealing approach to sensor network localization. In *Proceedings of the KAEiOG'07 conference*, pages 193–202. WUT press, 2007.

[8] M. Marks and E. Niewiadomska-Szynkiewicz. Two-phase stochastic optimization to sensor network localization. In *SENSORCOMM 2007: Proceedings of the international conference on Sensor Technologies and Applications*, pages 134–139. IEEE Computer Society, 2007.

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

[10] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz. Localization from connectivity in sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 15(11):961–974, 2004.