

LOCALIZATION IN WIRELESS SENSOR NETWORKS: CLASSIFICATION AND EVALUATION OF TECHNIQUES

EWA NIEWIADOMSKA-SZYNKIEWICZ

Institute of Control and Computation Engineering
Warsaw University of Technology, ul. Nowowiejska 15/19, 00-665 Warsaw, Poland
e-mail: ens@ia.pw.edu.pl

Research and Academic Computer Network, ul. Wązowska 18, 02-796 Warsaw, Poland
e-mail: ewan@nask.pl

Recent advances in technology have enabled the development of low cost, low power and multi functional wireless sensing devices. These devices are networked through setting up a Wireless Sensor Network (WSN). Sensors that form a WSN are expected to be remotely deployed in large numbers and to self-organize to perform distributed sensing and acting tasks. WSNs are growing rapidly in both size and complexity, and it is becoming increasingly difficult to develop and investigate such large and complex systems. In this paper we provide a brief introduction to WSN applications, i.e., properties, limitations and basic issues related to WSN design and development. We focus on an important aspect of the design: accurate localization of devices that form the network. The paper presents an overview of localization strategies and attempts to classify different techniques. A set of properties by which localization systems are evaluated are examined. We then describe a number of existing localization systems, and discuss the results of performance evaluation of some of them through simulation and experiments using a testbed implementation.

Keywords: wireless sensor network, *ad hoc* network, localization, location systems, network simulation.

1. Introduction to wireless sensor networks

A Wireless Sensor Network (WSN) is a wireless decentralized structure network comprised of nodes, which autonomously set up a network. No external network infrastructure is necessary to transmit data—there is no central administration. The WSN is a distributed system typically composed of small-size, embedded devices grouped into network nodes deployed densely over a significant area. Each node is equipped with a processing unit, a radio transceiver, battery, sensors and/or actuators. In most cases WSNs are stationary or quasi-stationary, but recently Mobile Wireless Sensor Networks (MWSNs) have become more and more popular. The common communication pattern used is many-to-one, i.e., sensors send gathered data to designated nodes responsible for transmitting the information to the network's operator. The lack of fixed network infrastructure components allows creating unique topologies and enables the network dynamics. However, for the protocols to operate in this mode in practice, several basic issues must be solved. The most important ones include: limited energy resources, through-

put, computational power and memory, poor quality of connection, dynamically changing network topology, limitations in sensor accuracy, or problems with ensuring secure network operation. Therefore, design and development of WSNs is a non-trivial task. The main directions of current research in *ad hoc* networking include increasing the potential of hardware components in terms of a smaller size of devices and their cost, accurate location systems for calculating positions of devices, energy-aware communication, dedicated operating systems, or protocols and algorithms that provide scalable and security services (Akyildiz and Vuran, 2010; Beutel, 2005; Karakehayov, 2009; Marks, 2010; Rappaport, 2002; Santi, 2006; Verdone *et al.*, 2008).

An important direction of research is devoted to design and development of methods and tools for evaluation of the performance of novel protocols and algorithms for WSNs. Evaluation and comparison of WSN applications and technologies can be performed on various scales and using various metrics. Three techniques, i.e., analytical methods, computer simulation and practical

evaluation on a wireless testbed, are accepted and commonly used. Recently, formal methods have been used to design and evaluate many WSN applications, i.e., the placement of devices in the environment and density control algorithms for energy-aware sensing (Olveczky and Thorvaldsen, 2007), localization algorithms (Biswas and Ye, 2004; Mao and Fidan, 2009), prototyping of simple routing protocols (Bernardeschi *et al.*, 2008), etc. Unfortunately, the complexity of WSNs in many applications makes analytical methods to be unsuitable. The simulation is a relatively easy and highly available technique to validate the performance of WSN technologies, and it is commonly used to evaluate localization techniques. To apply a given programming approach in a real network, programmers must consider the specific hardware and software platforms. Recently a number of software systems for wireless networks simulation have been developed to aid programmers. A survey of open source and commercial platforms is presented by Di Caro (2003) and Kasch *et al.* (2008). Some are dedicated systems focused on a specific attribute of the behavior of a given network simulation, the others are general purpose tools that can be used for development and testing of various types of networks.

Unfortunately, there is a high risk of incorrect results of simulations (Pawlikowski *et al.*, 2002). It is obvious that simulation can only approximate real-world computations and depends on the accuracy of the description of the behavior of a target environment. In general, experiment setup needs values of input parameters that have to be estimated. Hence, it is important to calculate input values based on real life measurements. Moreover, different assumptions may be hidden in a given network simulator; in particular, radio propagation must be modeled using simplifying assumptions. These problems should be considered in evaluation of simulated outcomes. The solution is to evaluate the WSN performance through experiments using testbed implementation. TinyOS (www.tinyos.net) and Contiki (www.contiki-os.org) are open source, highly portable operating systems that can be used for on-line operation of a WSN formed of real low-power wireless devices. However, experiments in physical networks are often costly, time consuming, and in many applications impossible.

The remainder of this paper is focused on localization in WSNs. The main contribution is to point out the problems concerned with accurate localization of nodes in the network, and survey different localization strategies and systems. Strategies, their classification, and criteria for performance evaluation are presented in Section 2. In Section 3, we investigate some location systems, in section 4, we evaluate selected localization techniques through simulation and testbed implementation.

2. Localization strategies in WSNs

The goal of localization (Anderson *et al.*, 2007; Mao and Fidan, 2009) is to assign geographic coordinates to each node with an unknown position in the deployment area. Most applications of WSNs, e.g., monitoring, target tracking, search, etc., require the correlation of sensor readings with physical locations. Moreover, even if the accessible knowledge about positions of nodes is only approximate, there are great opportunities for using various network services, location-based routing, data aggregation, etc.

Information on the location of nodes can be obtained in two ways: (i) recording data on the location of nodes during their distribution, (ii) fitting nodes with the GPS. Both methods have significant defects. Manually recording and entering positions of each sensor node is impractical for very large sensor networks. In many applications, sensors are distributed randomly in an *ad hoc* fashion, which is a cheaper and in some cases the only possible solution. Moreover, this method cannot be used in mobile networks where nodes can travel. Another solution is to collect data on the location of sensors by means of GPS devices. This solution can be used in different types of networks, including mobile ones. Unfortunately, it is very costly, both due to the price of GPS receivers and the increased requirements related to power consumption, which may decrease the lifetime of a WSN. Moreover, adding an extra sensor increases the size and weight of the total device (network node). Due to the drawbacks of the above-mentioned solutions, many automated location systems for assigning geographic coordinates to each node have been developed. All these schemes should work with inexpensive off-the-shelf hardware, have minimal energy requirements, scale to large networks, and also achieve good accuracy in the presence of irregularities and give the solution in the short time. Various localization strategies for WSNs have been developed and described in the literature (Anderson *et al.*, 2007; Akyildiz and Vuran, 2010; Mao and Fidan, 2009; Sarigiannidis, 2007).

2.1. Classification. The localization techniques can be classified with respect to various criteria. They differ on the assumed localization precision, hardware capabilities, measurement and calculation methods, computing organization, the assumed network configuration, architecture, nodes properties and deployment, etc. The localization strategies can generally be divided into

- coarse-grained localization, and
- fine-grained localization.

Coarse-grained localization techniques consist in finding approximate coordinates of nodes in a network, and

hence provide lower precision estimates of these coordinates. In contrast, fine-grained strategies determine precisely the coordinates but require much more communication and computation efforts. Representatives of fine-grained strategies use costly distance estimations to achieve localization with high precision, and various techniques to transform the distances to coordinate vectors.

With regard to fitting all or selected nodes with a GPS device, we can distinguish

- GPS-free and anchor-free localization schemes, and
- anchor-based localization schemes.

The most recently proposed localization strategies use the available knowledge on the location of the set of nodes in a network. An anchor (also known as a beacon node) is defined as a node that is aware of its own location, either through the GPS or manual pre-programming during deployment. In anchor-based localization schemes, identification of other nodes is up to an algorithm locating non-anchors based on merely partial information on the location of anchor nodes. In contrast, anchor-free localization does not use any anchor or dedicated hardware like the GPS. Therefore, using anchor-free localization is possible even in environments that are out of a GPS signal.

With regard to hardware capabilities of devices that form a network and mechanisms used for the estimation of inter-node distances, we divide the localization protocols into two categories:

- range-based (distance-based) methods, and
- range-free (connectivity-based) methods.

The former are defined by protocols that use absolute point to point distance estimates (range) or angle estimates in location calculation. The latter make no assumption about the availability or validity of such information, and use only connectivity information to locate the entire sensor network.

Various techniques may be used for a location estimation. The coordinates of nodes in a network can be calculated using

- geometrical techniques,
- multidimensional scaling,
- stochastic proximity embedding,
- algorithms for convex and nonconvex optimization, and
- hybrid schemes that use two different techniques.

Geometrical range-based techniques give solutions to a set of nonlinear equations. The commonly used ones are triangulation, trilateration and multilateration. The

philosophy of localization techniques based on Multi-Dimensional Scaling (MDS) and Stochastic Proximity Embedding (SPE) is to transform a mathematical model to convert distance information into the coordinate vector. The common idea of other methods is formulating the localization problem as a nonlinear, nonconvex optimization task solved by global optimization solvers, or relaxing the resulting problem as a convex optimization problem solved by quadratic or linear solvers. A recently, popular group consists of methods that use more than one technique to estimate location, i.e., results of initial localization are refined using another localization method.

Considering the distribution of the calculation process, we can distinguish

- centralized location systems, and
- distributed location systems.

In centralized systems, data collected in the whole network are transmitted to the central unit that calculates the estimated location of each node in a network. In distributed systems each node estimates its own position based on the local data gathered from its neighbors.

With regard to nodes' mobility, we can distinguish

- localization for static WSNs,
- mobile assisted localization for static WSNs, and
- localization for mobile WSNs.

Most localization algorithms are focused on localization for static sensor networks, but not for mobile WSNs. Some location systems for static WSNs use a mobile device or devices to improve localization accuracy and scalability, and to decrease energy and computational costs. Mobility of nodes can be used to improve the accuracy of inter-node distance estimation, increase the density of a network, etc. Hence, it can improve results of the localization process. The requirement is that we can control the movement. Another problem is design and development of localization systems for mobile sensor networks. The localization algorithm for mobile networks has to consider the movement of sensor nodes, i.e., movement direction and speed. Traditional localization strategies and systems for static WSNs are not suitable for MWSNs. Different approaches to localization in MWSNs are proposed in the literature. Some of them utilize event-driven schemes (Zhang and Yu, 2010a) and the Monte Carlo method (Hu and Evans, 2004).

Some localization strategies take into account the characteristics of a network. A loosely connected WSN typically has different properties compared with a strictly hierarchical network formed by clusters of nodes. Hence we can consider two types of localization strategies that assume the following characteristics of a network:

- sensors loosely organized in a sensor field,
- network consists of clusters of devices.

Hierarchical cluster-based solutions are often proposed to improve scalability and efficiency of the location system. The network is divided into clusters, and cluster heads are selected. Different algorithms can be used to estimate positions of cluster heads and remaining nodes. For position estimation of cluster heads usually more complex but accurate protocols are used. The remaining nodes can use a simpler but less accurate method with cluster heads as reference nodes.

2.2. Performance metrics. Multiple metrics can be used to measure the performance of a localization technique. It is not enough to observe accuracy only. Referring to the literature and considering results of our research we provide the following performance measures: accuracy, coverage, complexity, scalability, robustness and cost. They are mainly connected with economical or technical constraints such as hardware cost, low battery power and limited computation capabilities.

Localization accuracy. Accuracy is the most important requirement of location systems. Usually, the mean error between the estimated and the true location of the non-anchor nodes in the network is adopted as the performance metric. It is defined as follows:

$$LE = \frac{1}{N} \cdot \sum_{i=1}^N \frac{\|\hat{x}_i - x_i\|^2}{r_i^2} \cdot 100\%, \quad (1)$$

where N denotes the number of nodes in a network whose location is estimated, LE denotes a localization error, x_i the true position of the node i in the network, \hat{x}_i is estimated location of the node i (solution of the location system) and r_i is the radio transmission range of the node i . The localization 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 sizes and ranges of networks. Usually, centralized location systems give more accurate position estimates than distributed ones. Distributed implementation may involve a loss of information due to an incomplete network map and parallel computations.

It is obvious that the higher the accuracy, the better the system. However, there is often a trade off between position estimation accuracy and other characteristics. Therefore a compromise between the required accuracy and other characteristics is needed.

Coverage. In general, coverage of localization techniques is related to the deployment area, network density, hardware equipment and resources of devices that form a

network. Sometimes, particularly in large, distributed networks when nodes do not have enough neighboring nodes, unevenly distributed anchor nodes, or in the case of poor equipped devices, problems with localization of the whole network may occur. In such a situation the question is how much of the network can be localized. In the case of poor results, the only option is to increase the number of anchor nodes in a network.

Complexity. The complexity of a location system can be attributed to hardware, software, and operation factors. In general, range-based methods are much complex than range-free techniques and involve hardware complexity. Software complexity depends on the computing complexity of the positioning algorithm. In centralized location systems a central unit calculates the estimated locations due to its powerful processing capability and sufficient power supply and memory. If calculations are carried out on the sensor node, the effects of complexity could be evident. Most of devices that form a network lack strong processing power, memory and power source, so techniques with low complexity are often preferred.

Scalability. The scalability of a location system ensures suitable estimation of localization when the network or deployment area gets larger. A location system should scale on the network size (number of nodes) and density, the size of a deployment area and dimensional space. In the case of range-based techniques the location performance degrades when the distance between the transmitter and receiver increases. On the other hand, in dense network wireless signal channels may become congested, and more complex communication infrastructure may be required. The location system can locate the nodes in 2-D or 3-D space; some of them can support both 2-D and 3-D spaces. Centralized systems usually aggregate all measurements and input data at a central unit to carry out processing. By contrast, distributed implementation of localization improves scalability.

Robustness and fault tolerance. A robust and fault-tolerant localization system should function normally even when some signals are noisy or even unavailable, or when some of the signal values or angle character have never been seen before. Some devices that form a WSN system could be damaged or out of over. Sometimes, particularly in indoor localization, the signal from a transmitter is blocked because of obstacles and signals from some nodes cannot be obtained. Hence, the only information to estimate the location is the signal from other measuring nodes. Location systems have to use this incomplete information to estimate the position of nodes. Moreover, the location system should be resistant to irregular WSN topology and unevenly distributed anchor nodes.

Cost. The cost of a location system may depend on many factors. The most important include money (hardware and software costs), weight, energy, and time. Each sensor node is equipped with a radio transceiver to communicate with other nodes. Some location systems, i.e., range-free or range-based using the RSSI for inter-node distance estimation, do not need additional hardware equipment. To obtain better resolution of measurements, additional hardware equipment is usually required that significantly increases the cost of each node. Moreover, it increases the weight of a device.

Energy is an important cost factor of a system. A localization process involves additional computation and communication. At each sensor node, energy is consumed for data recording, processing and transmission. Energy used for data processing depends on the quantity of data and the complexity of calculations. Centralized systems minimize the complexity of operations that are performed by nodes in a network, but involve communication overhead. On the other hand, in distributed implementations, inter-node communication is limited to neighboring nodes but more computational power and memory are necessary to provide calculations.

Just like energy also computation time is related to data processing and transmission. The computation time depends on the network size, the efficiency of multi-hop transmission, the complexity of the localization technique, and computational power.

3. Survey of location systems

A number of research and commercial location systems for WSNs have been developed. They implement the strategies described in Section 2. Recently proposed location systems usually consist of three components: identification and data exchange, measurement and data acquisition, and position calculation. Simpler but less accurate are coarse-grained location systems. The common representatives include: Centroid Localization (CL), Weighted Centroid Localization (WCL) and Adaptive Centroid Localization (AWCL) presented by Salzmann *et al.* (2011). In the basic CL-based localization, each unknown node estimates its location by calculating the center of the location of all received anchors. WCL and AWCL use additional information about inter-node distance measurements to calculate the centroid. An alternative approach is the Location Estimation Algorithm (LEA) described by Hu and Evans (2004). The LEA provides a probabilistic distribution of possible node locations. The observations from anchor nodes are used to improve localization accuracy and filter impossible locations. Heurtefeux and Valois (2008) propose the Qualitative Localization Protocol (QLoP) to approximate coordinates. Qualitative localization is a simple GPS-free and anchor-free scheme. The QLoP determines coarsely the position of the neighbors

of a given node based only on local information received from its 1-hop neighbors. Each node creates a qualitative distance table according to 1-hop and 2-hop neighborhood data. Finally, qualitative location of each neighbor is classified as *very closed*, *closed* or *far*. The disadvantage of coarse-grained location systems is that they are not able to estimate the exact position of an unknown node even with exact distance measurements.

Fine-grained location systems determine exact positions of unknown nodes for exact distance measurements but use costly computations. They usually calculate the location of nodes based merely on partial information on the location of the anchor nodes that are aware of their own location. The remainder of this paper is focused on fine-grained, anchor-based location systems.

Let us consider a network formed by $L = M + N$ sensors, M anchor nodes, each with a known position expressed as n -dimensional coordinates $a_k \in \mathbb{R}^n$, $k = 1, \dots, M$, and N non-anchor nodes $x_j \in \mathbb{R}^n$, $j = 1, \dots, N$ with unknown locations. The goal of a location system is to estimate coordinate vectors of all N non-anchor nodes. In general, the localization scheme operates in two stages:

- *Relative location calculation* (measurement and distances estimation stage): estimation of inter-node distances based on inter-node transmissions.
- *Absolute location calculation* (position calculation stage): calculation of geographic coordinates of nodes with unknown location.

The results of calculations have to be transmitted to a base station or to other nodes in a network. Hence, the important component of each location system is *identification and data exchange*.

3.1. Relative location calculation. As has been mentioned, the most recently proposed location systems compute approximate location of nodes based on the information on the location of anchor nodes. All nodes, regardless of their absolute coordinate, estimate the distance between themselves and their neighbors. Such location estimation is called *relative location*. In the next stage of the localization process the distances are converted into coordinate vectors.

We can divide relative location techniques into two categories: range-based and range-free methods. The popular range-free solutions are hop-counting techniques. Assume that each anchor node a_k , $k = 1, \dots, M$ exchanges messages with other nodes. Hence, the distances in hops h_{kl} between each pair (k, l) of anchors in the network are estimated. Next, each anchor computes an average size for one hop

$$c_k = \frac{\sum_{l \in S_k} \|a_k - a_l\|}{\sum_{l \in S_k} h_{kl}}, \quad k \neq l,$$

where S_k denotes a set of anchors located within a transmission range of r_k , $S_k = \{(k, l) : \|a_k - a_l\| \leq r_k\}$, $l = 1, \dots, M$. The calculated values are broadcasted into the network, and the inter-node distances are estimated.

Range-based methods require additional equipment but with that we can reach much better resolution than in the case of range-free approaches. In accordance with the available hardware, they exploit the angle of arrival, the time of arrival, the time difference of arrival of two different signals and the received signal strength. A survey and discussion of the most popular technologies is available in the works of Akyildiz and Vuran (2010), Barsocchi *et al.* (2009), Benkic *et al.* (2008), Beutel (2005), Karl and Willig (2005), and Motter *et al.* (2011). In this paper we provide a brief description.

The Angle of Arrival (AoA) is a technique for determining the direction of propagation of a radio-frequency wave. The traditional approach to measuring angles is to use directional antennas, rotating on their axis similar to a radar station. Another technique is determining the direction of propagation of a radio-frequency wave incident on an antenna array. This approach cannot be used in networks formed by small-size sensors with only a single antenna setting. The Time of Arrival (ToA) and Time Difference of Arrival (TDoA) methods exploit the relationship between the distance and transmission time when the propagation speed is known. If the sender and receiver know the time when the transmission starts, the time of signal propagation can be used to estimate the distance. ToA schemes typically use signals that move at a slower speed (such as ultrasound). The sender (node i) sends a signal to the receiver (node j) and, in return, the receiver j sends a signal back to the sender i . Next, the inter-node distance d_{ij} is estimated according to the formula (Beutel, 2005)

$$d_{ij} = 0.5[(T_i^r - T_i^t) - (T_j^r - T_j^t)]v, \quad (2)$$

where T_i^t denotes the time of transmission and T_i^r the time of reception of a signal at the node i , v is the predefined velocity of a signal (the same for all nodes). In order to provide results of acceptable accuracy, very high resolution accurately synchronized clocks are required for the sender and receiver in the case of the ToA method.

The technique that allows us to overcome the need of explicit synchronization is the TDoA method, which utilizes implicit synchronization by directly providing the start of transmission information to the receiver. This is possible if two transmission mediums of very different propagation speeds are used. For example, the sender can send an ultrasound and a radio signal simultaneously. When the receiver gets the radio signal, it can start measuring the time until arrival of the ultrasound transmission, safely ignoring the propagation time of the radio communication. The distance d_{ij} between the sender i and the

receiver j can be determined as follows:

$$d_{ij} = \frac{v_r v_u}{v_r - v_u} [(T_j^{ru} - T_j^{rr}) - (T_i^{tu} - T_i^{tr})], \quad (3)$$

where T_i^{tu} and T_i^{tr} denote respectively the time of transmission of an ultrasound and a radio signal at the node i , T_j^{ru} and T_j^{rr} the times of reception of these signals at the receiver j , v_u and v_r velocities of the ultrasound and the radio signal.

A combination of AoA and TDoA techniques is proposed by Magnani and Leung (2007) to improve estimation accuracy. It should be pointed that both the ToA and the TDoA are too demanding to be applied in sensor networks with limited computational and communication abilities. The alternative proposed by Liu and E (2010) is a method that combines the linear Frequency Modulation Continuous Wave (FMCW) and the Time Frequency Difference of Arrival (TFDA) technique. In this approach, instead of an estimate of an inter-node distance, a distance difference from one node to another is calculated.

The common technique based on a standard feature found in most wireless devices is the Received Signal Strength Indicator (RSSI). The advantage of this method is low cost (no additional hardware), easy configuration, calibration and deployment; the disadvantage—low level of measurement accuracy because of a high variability of the RSSI value. In real-world channels, multipath signals and shadowing are two major sources of environment dependence in the measured RSS.

The commonly used radio signal propagation models indicate that the received signal power decreases with a distance, both in outdoor and indoor environments. Therefore, the power of the signal received by a receiver P^r at a distance d is defined as

$$P^r(d)[\text{dBm}] = P^t[\text{dBm}] - PL(d)[\text{dB}], \quad (4)$$

where P^t denotes power used by a sender to transmit the signal and $PL(d)$ the average signal degradation (path loss) with a distance d . A path loss $PL(d)$ in (4) is modeled as a function of a distance d raised to an attenuation constant n that indicates the rate, at which the path loss increases with a distance:

$$PL(d)[\text{dB}] = PL(d_0)[\text{dB}] + 10n \log \left(\frac{d}{d_0} \right), \quad (5)$$

where d_0 is a close-in reference distance (for IEEE 802.15.4 usually $d_0 = 1$ m). The formula (5) was developed as a combination of analytical and empirical methods.

It should be remarked that the transmission area of a transmitter may be different at two different locations, which leads to measuring signals that are different than the average value calculated according to (5). In the extended model (6), a path loss at the distance d is expressed as a

random variable with log-normal distribution,

$$PL(d)[\text{dB}] = PL(d_0)[\text{dB}] + 10n \log \left(\frac{d}{d_0} \right) + X_\sigma, \quad (6)$$

where X_σ is a zero-mean Gaussian distributed random variable with standard deviation σ (all in dB).

The above channel models can be used to estimate the average distance \tilde{d}_{ij} between each pair of nodes (i, j) in a network. Let us consider the basic model (5). We can estimate the average distance between each pair of nodes as a function of the received signal strength:

$$\tilde{d}_{ij} = d_0 \cdot 10^{(P^t - PL(d_0))/10n} \cdot 10^{-P_{ij}^r/10n}, \quad (7)$$

where $PL(d_0)$ denotes the path loss at the reference distance d_0 , P^t is output power of the transmitter, P_{ij}^r is the received signal strength measured for each pair (i, j) of nodes.

In the formula (7) we have four parameters that should be tuned: n , d_0 , P^t , $PL(d_0)$. After substituting $\alpha = d_0 \cdot 10^{(P^t - PL(d_0))/10n}$ and $\beta = -1/10n$ in (7), the estimated distance is expressed as

$$\tilde{d}_{ij} = \alpha \cdot 10^{\beta P_{ij}^r(d)}. \quad (8)$$

The parameters α and β can be easily calculated based on known distances between pairs of anchor nodes. Finally, the following algorithm is proposed for inter-node distances calculation:

Step 1: Measure the RSSI for all pairs of nodes in a given WSN.

Step 2: Calculate the values of unknown parameters α and β in (8) solving a least squares problem for all pairs of anchor nodes,

$$\min_{\alpha, \beta} \sum_{k=1}^M \sum_{l \in S_k} (\alpha \cdot 10^{\beta P_{kl}^r} - d_{kl})^2, \quad (9)$$

where d_{kl} denotes a known true physical distance between nodes k and l , $S_k = \{(k, l) : \|a_k - a_l\| \leq r_k\}$, $l = 1, \dots, M$, is a set of neighboring anchor nodes of the node k .

Step 3: Calculate the average distances \tilde{d} to all nodes located within transmission ranges using the formula (8) and optimal values of α and β calculated in Step 2.

A practical evaluation of the radio signal strength for ranging-based localization is presented and discussed by Mao and Fidan (2009), Motter *et al.* (2011), and Whitehouse *et al.* (2007).

3.2. Absolute location calculation. The estimated inter-node distances \tilde{d} and known positions of anchor nodes a_k , $k = 1, \dots, M$ are used to calculate the positions of all non-anchor nodes in the network. As presented in Section 2, position estimation can be done by using triangulation, trilateration or multilateration, multidimensional scaling, stochastic proximity embedding, linear or nonlinear optimization solvers or hybrid methods. Triangulation, trilateration and multilateration are popular geometrical techniques reported in many works (Akyildiz and Vuran, 2010; Beutel, 2005). In the case of the *triangulation* method, the position of the unknown node is calculated using the locations of two anchor nodes and the AoA measurements at an unknown node. The *trilateration* technique requires the distance measurements between the node with unknown location and its neighbors (in 2-D space, three neighbors with known locations are sufficient). The minimization problem with the performance function calculated as a difference between the measured and estimated distances is formulated and solved. Several variants of *multilateration* method are proposed to reduce limitations of the typical trilateration scheme. *Atomic multilateration* incorporates distance measurements from multiple neighbors. It is used to improve the accuracy of location estimation if the distance measurements are noisy. The idea of *iterative multilateration* is to repeat trilateration for an increased number of anchor nodes (every iteration, each node with an estimated position becomes an anchor).

A number of protocols and location systems implementing geometrical techniques have been developed. A simple and often used protocol implementing the trilateration method is called the *Ad-hoc* Positioning System (APS) and was developed by Niculescu and Nate (2001). Two versions of the APS, range-free (DV-hop) and range-based (DV-distance), are most popular. In the case of DV-distance, the inter-node distance is estimated based on the RSSI method. The AHLoS (*Ad Hoc* Localization System) described by Savvides *et al.* (2001) uses the multilateration method to calculate the location of nodes in the network, and either the RSSI or the ToA for inter-node distance measurement. The usage of the multilateration technique and the dynamic spatial RSSI filter for distance estimation is proposed for dynamic indoor localization of slowly moving sensor devices and is described by Wessels *et al.* (2010). The TPS (Time-based Positioning Scheme) (Cheng *et al.*, 2004) provides localization service using the TDoA for range detecting, and trilateration for position calculation. This method is energy efficient but requires anchor nodes with powerful radio transceivers.

Several popular range-free and range-based location systems apply Multi-Dimensional Scaling (MDS) to node localization. In general, MDS is a set of data analysis techniques often used for exploring similarities or dissimilarities in data (Mao and Fidan, 2009). The basic MDS

model is defined by specifying how given similarity data between two entities i and j are mapped into distance in an Euclidean space. The matrix of distances for a configuration consisting of all entities is formed. Next, distances are converted into the coordinate vector.

There are many types of MDS techniques, i.e., metric and non-metric, deterministic and probabilistic, replicated, weighted, etc. A number of MDS-based localization algorithms have been developed and described in the literature. They differ in inter-node distances estimation and the type of MDS. The centralized and distributed implementations are proposed. Selected algorithms developed by various researchers are listed below:

- metric MDS: MDS-MAP(C), MDS-MAP(P), MDS-Hybrid, RangeQ-MDS;
- non-metric MDS: the algorithm developed by Vo, Challa and Lee;
- weighted MDS: the algorithm developed by Costa, Patwari and Hero.

MDS-MAP(C), described by Shang *et al.* (2004) is a centralized technique that simply applies MDS to estimate positions of nodes in a given deployment area based on a global map of the network. It operates in three steps: (i) the shortest path distances between all pairs of nodes are computed, and a distance matrix is created; (ii) the relative global map of the network is generated applying classical MDS to the distance matrix; (iii) the relative map is transformed to an absolute map based on the positions of anchor-nodes. The MDS-MAP(P) method is a distributed version of MDS-MAP, in which local maps are built and then merged together to form a global one. It is a good solution due to its ability to localize irregular networks, but it is complicated and costly. Moreover, centralized calculation is necessary to merge local maps. MDS-Hybrid is a modification of MDS-MAP(C). It operates in two steps: (i) MPS-MAP(C) is used to localize reference nodes, (ii) the reference nodes are used to localize the rest of nodes. Range-free and range-based versions of MDS-MAP(C), MDS-MAP(P) and MDS-Hybrid are available. The RangeQ-MDS algorithm is partially range-based. It uses a quantized RSSI-based distance estimation technique to calculate more accurate hop distances. In the case of the method described by Vo *et al.* (2008), instead of a network map created due to inter-node distances, the configuration of nodes maintains the rank of dissimilarities. In the work of Costa *et al.* (2006) a distributed algorithm higher weight assigns a to the distance measurements that are believed to be more accurate.

The idea of localization techniques based on Stochastic Proximity Embedding (SPE) is similar to those applying MDS, i.e., distance information is converted into the coordinate vector. Unlike MDS-based localization methods, the SPE-based algorithm described by Aloor and

Jacob (2010) does not require a complete matrix of distance. Only exact inter-node distances between neighboring nodes and lower bounds between remote nodes are required.

Cluster-based variants of MDS-based localization algorithms are described by Medidi *et al.* (2006) and Shu *et al.* (2009). The former propose two algorithms with different complexity for localization of the cluster head and other nodes. The latter focus on three-dimensional large scale network localization. The so-called CBLALS localization algorithm presented utilizes cluster topology to divide the whole network into small clusters. Next, the MDS algorithm is used to estimate the in-cluster relative coordinate in each cluster.

Another approach to position estimation is to formulate an optimization problem and solve it using linear or nonlinear solvers. The standard formulation is the optimization problem with nonlinear performance function J_N :

$$\min_{\hat{x}} \left\{ J_N = \sum_{k=1}^M \sum_{j \in S_k} (\hat{d}_{kj} - \tilde{d}_{kj})^2 + \sum_{i=1}^N \sum_{j \in S_i} (\hat{d}_{ij} - \tilde{d}_{ij})^2 \right\}, \quad (10)$$

where $\hat{d}_{kj} = \|a_k - \hat{x}_j\|$, $\hat{d}_{ij} = \|\hat{x}_i - \hat{x}_j\|$, a_k denotes the real position of the anchor node k , \hat{x}_i and \hat{x}_j denote respectively the estimated positions of nodes i and j , \tilde{d}_{kj} and \tilde{d}_{ij} are the estimated distances between pairs of nodes calculated based on measurements, and S_i, S_k are sets of neighboring nodes. Common for most sensor localization methods is selecting node neighborhoods for range measurements:

$$S_k = \{(k, j) : \|a_k - x_j\| \leq r_k\}, \quad j = 1, \dots, N, \quad (11)$$

$$S_i = \{(i, j) : \|x_i - x_j\| \leq r_i\}, \quad j = 1, \dots, N,$$

where x_i and x_j denote real positions of nodes with unknown locations, and r_i and r_k their transmission ranges. However, selecting node neighborhoods only for range measurements cannot be the best solution when ranges are measured with noise. An alternative approach—the two-stage neighbor selection process that can be used to make location estimates unbiased even in high-noise environments—is proposed by Costa *et al.* (2006).

Various optimization techniques are used to solve the optimization problem (10). The most popular approaches are quadratic programming, linear programming, nonlinear and nonconvex optimization techniques. The first class of methods relax the original nonconvex problem (10) in order to obtain a quadratic problem:

$$\min_{\hat{x}} \left\{ J_L = \sum_{k=1}^M \sum_{j \in S_k} e_{kj} + \sum_{i=1}^N \sum_{j \in S_i} e_{ij} \right\} \quad (12)$$

subject to

$$(\tilde{d}_{ij}^{\min})^2 - e_{ij} \leq \|\hat{x}_i - \hat{x}_j\|^2 \leq (\tilde{d}_{ij}^{\max})^2 + e_{ij}, \quad (13)$$

$$\forall i \neq j, \quad j \in S_i,$$

$$(\tilde{d}_{kj}^{\min})^2 - e_{kj} \leq \|a_k - \hat{x}_j\|^2 \leq (\tilde{d}_{kj}^{\max})^2 + e_{kj}, \quad (14)$$

$$\forall k, j, \quad j \in S_k,$$

where $e_{kj} \geq 0$ and $e_{ij} \geq 0$ denote errors in sensor position estimations. Quadratic Programming (QP) can be applied to solve (12)–(14). Range-free and range-based methods can be used for distance estimation. A range-free localization system OPDMQP using quadratic programming is described by Lee *et al.* (2010), who discuss its application to isotropic and anisotropic networks.

Another popular technique is to transform the quadratic problem (12)–(14) into a standard Semi-Definite Programming (SDP) or a Second-Order Cone Programming (SOCP) problem. Quadratic distance constraints may be converted into linear constraints by introducing a relaxation to remove the quadratic term in the formulation. The existing linear solvers (usually interior point methods) are used to solve the transformed problem. The formulation of the SDP localization problem is provided by Biswas and Ye (2004).

In the literature, one can find many semidefinite programming methods (Boyd *et al.*, 1994). SDP-based localization is very computationally expensive for large-scale networks, thus centralized computing is necessary. Cluster-based SDP is proposed to reduce the computational load. Another approach is to apply further relaxation to the range-based problem; node-based SDP relaxation and edge-based SDP relaxation are described by Wang *et al.* (2008). An alternative way to reduce the localization problem complexity is to apply second-order cone programming relaxation to node localization as described by Tseng (2007) and Shi *et al.* (2010). A totally asynchronous distributed range-based algorithm using SOCP is presented by Srirangarjan *et al.* (2008). SOCP relaxation, though weaker than the SDP one, produces a simpler optimization problem formulation and speeds up calculations. Finally, mixed SDP-SOCP relaxation is considered.

The results of comparison of SDP and SOCP discussed by Tseng (2007) demonstrate that in the case of large WSNs the preferred relaxations are SOCP or mixed SDP-SOCP because SDP is usually too expensive to use. Shi *et al.* (2010) propose to apply the Sequential Greedy Optimization (SGO) algorithm to the edge-based SDP relaxation formulation. The SGO algorithm is an extension of the nonlinear Gauss–Seidel algorithm and is more suitable for distributed localization. Shi *et al.* (2010) present that by using SGO the edge-based SDP relaxation formulation can be solved in a distributed way through solving a sequence of second order cone programming. The convergence property of the SGO algorithm is proved and

analyzed. The results of extensive simulations show that the distributed SGO-based location system can work almost as well as the centralized SDP-based system, and provides a substantial improvement in localization accuracy with respect to the SOCP method. The authors suggest to use SGO-based localization in networks in which anchor positions cannot be exactly known. Unfortunately, they do not present the results for networks with anchors randomly and unevenly placed in a network.

The application of the Lagrange multiplier method to the localization problem is described by Li *et al.* (2010), taking into account the cost function (12) and the constraints (13)–(14). They formulate an unconstrained optimization problem with the augmented objective function, and use an iterative multiplier method to solve it. Unfortunately, the authors do not present any comparative study for other methods.

A formulation of a range-based and a range-free localization problem as a non-differentiable optimization problem solved by the normalized incremental subgradient algorithm is proposed by Shi *et al.*, (2008; 2009). The simulation results presented in the mentioned papers show that this approach exhibits much better performance than SDP-based and MDS-MAP localization.

A recently popular strategy is to apply nonlinear, nonconvex optimization algorithms to solve the localization problem (10). Numerous approaches are proposed and described in the literature. Many researchers suggest to use popular heuristic methods, i.e., deterministic (Tabu Search, TS) and stochastic (Simulated Annealing, SA; Genetic Algorithm, GA; Evolutionary Algorithm, EA; Particle Swarm Optimization, PSO) to calculate location estimates. Referring to the literature and considering the results of our research, it seems that stochastic techniques (SA, GA, EA) considerably outperform QP, SDP or SOCP. These algorithms give an acceptable location accuracy in a acceptable computation time.

Results of simulated annealing in location estimation are provided in several papers (Kannan *et al.*, 2005; 2006; Niewiadomska-Szynkiewicz and Marks, 2009). The authors propose different modifications of the basic SA to improve the results and speed up calculations. Another approach is to use various versions of the genetic or evolutionary algorithm. A two-objective evolutionary algorithm taking into account topological constraints for location estimation is developed by Vecchio *et al.* (2012). Since the connectivity in a WSN is not sufficiently high, the authors propose some modifications to a basic EA. The algorithm takes into account both localization accuracy and certain topological constraints induced by connectivity considerations during location estimation.

A range-based localization system with the distances estimation based on the signal strength measurement and the Imperialist Competitive Algorithm (ICA) used to calculate coordinate vectors is presented by (Sayadnavard

et al., 2010). The ICA is a new evolutionary algorithm that is based on the simulation of a human's socio-political evolution. The simulation results presented by Sayadnavard and others highlight that the ICA-based approach considerably outperforms the APS. Moreover, it calculates estimates characterized by a higher accuracy than the ones obtained by the PSO-based localization scheme using the RSSI ranging technique (Chuang and Wu, 2008), but with more computational time.

Another range-based localization system TS&SA with distance estimation based on the ToA as well as application of tabu search and simulated annealing methods to location estimation is proposed by Shekofteh *et al.* (2010). The method was evaluated through extensive simulations. The authors claim that the TS&SA-based location system has better convergence characteristics compared with the SA-based system described by Kannan *et al.* (2005), but in the cited paper only the results of TS&SA system simulation are demonstrated and discussed without comparison to other solutions.

The last strategies presented are hybrid schemes that combine commonly used methods for computing the coordinate vector. In most approaches, trilateration or multilateration is used to calculate an initial solution, which is improved in the next step. Tam *et al.* (2006) developed a two-phase method. The APS based on the basic trilateration is used to calculate the initial localization. The Micro-Genetic Algorithm (MGA) is adopted to improve the accuracy of calculated estimates. The application of the APS and MDS-based algorithm is proposed by Ahmed *et al.* (2005).

We propose our scheme for location calculation that combines iterative multilateration along with nonconvex optimization and final correction. Two versions of this scheme were developed and tested: TSA (Trilateration & Simulated Annealing) and TGA (Trilateration & Genetic Algorithm), (Niewiadomska-Szynkiewicz and Marks, 2009). Both algorithms are range-based with the RSSI technique used for distance estimation. They operate in two phases. In the beginning of the first phase, all nodes in the network are divided into two sets: set $G1 = \{a_1, \dots, a_M\}$ containing anchor nodes, and set $G2 = \{x_1, \dots, x_N\}$ of nodes with unknown location. Next, iterative multilateration is used to determine the relative positions of nodes based on the known locations of nodes from $G1$, and the distances between pairs of nodes. In the second phase, the location estimation task is formulated as the optimization problem (10), and the SA algorithm or the GA is used to solve it. The goal of the second phase is to increase the accuracy of the location estimation calculated in the first phase, and estimate the position of nodes that cannot be calculated using iterative multilateration. The final step of our scheme is the correction of incorrect location estimates.

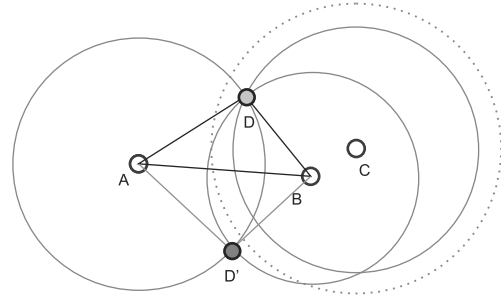


Fig. 1. Flip ambiguity phenomenon in WSN localization.

3.3. Correction of the localization error. In many WSN applications it can be observed that some nodes cannot be uniquely localizable. These location errors are often driven by the so-called flip ambiguity phenomenon, demonstrated in Fig 1. As the neighbors of node D are almost collinear and the inter-node distances are estimated with measurement errors, the localization algorithm usually calculates an incorrect location, i.e., D' instead of D in Fig. 1. It is obvious that the position of this node can be reflected with no change in the performance function in (10). This observation is discussed by many researchers, and different methods to solve this problem are proposed. From the simulations, it is observed that flip ambiguity is more common for networks with lower connectivity and a small number of anchor nodes. Hence, the simplest solution is to increase the connectivity and the number of anchors, but in many applications it is costly and impractical. A popular approach is to modify the basic localization algorithm or extend the localization process in the correction phase.

Kannan *et al.* (2006) and Shekofteh *et al.* (2010) propose methods in which localization is done in two steps, i.e., in the first phase the coordinate vectors are calculated, in the second phase the errors caused by the flip ambiguity are compensated. The algorithm described by Kannan *et al.* (2006) exploits two executions of the simulated annealing method. The goal of the first execution is to solve the optimization problem (10), and calculate the coordinates of the target nodes. The second phase is performed only on non-uniquely localizable nodes due to a flip ambiguity problem. The goal of this phase is to identify non-uniquely localizable nodes, and refine their location estimates calculated in the first phase. The SA algorithm is used again to solve the optimization problem with the modified objective function defined in (15). The function value is increased when a node is placed in a wrong neighborhood:

$$J_{FK} = \sum_{k=1}^M \left(\sum_{j \in S_k} (\hat{d}_{kj} - \tilde{d}_{kj})^2 + \sum_{\substack{j \in S_k \\ \tilde{d}_{kj} < r_k}} (\hat{d}_{kj} - r_k)^2 \right)$$

$$+ \sum_{i=1}^N \left(\sum_{j \in S_i} (\hat{d}_{ij} - \tilde{d}_{ij})^2 + \sum_{\substack{j \in S_i \\ \tilde{d}_{ij} < r_i}} (\hat{d}_{ij} - r_i)^2 \right), \quad (15)$$

where r_k and r_i denote the transmission ranges of the nodes k and i .

Shekofteh *et al.* (2010) propose a two-phase localization scheme TS&SA, in which two different optimization methods executed in cascade are used for network nodes' position (Shekofteh *et al.*, 2010). Tabu search is executed in the first phase to solve the optimization problem (10) and estimate initial locations of a node. In the second phase, the SA method is used to refine the location estimates of all non-uniquely localizable nodes. Similarly to Kannan's method, the optimization problem with the cost function J_{FK} (15) is solved.

Our contribution is to use nested optimization to solve a problem with non-uniquely localizable nodes. The idea is to introduce an additional functionality, correction operation, to the optimization solver. The correction is triggered every iteration in the optimization process whenever the value of the performance function J_N defined in (10) is lower than a threshold θ . Trilateration is executed to relocate all nodes placed in wrong neighborhoods by exploiting the nodes violating a smaller number of neighborhood constraints than the other randomly selected nodes. The threshold θ depends on the number of anchor nodes, network density and deployment, the power of radio devices and the expected noise measurement factor n_f . It is tuned according to the following formula:

$$\theta = \begin{cases} \mu \cdot n_f \cdot s^2 & \text{if } \frac{N+M}{M} < \gamma, \\ \lambda \cdot n_f \cdot s^2 & \text{if } \frac{N+M}{M} \geq \gamma, \end{cases} \quad (16)$$

where n_f is the noise measurement factor, μ , λ and γ experimentally tuned parameters. The variable s denotes an average number of neighbors of all nodes forming a network:

$$s = \frac{1}{N+M} \sum_{i=1}^{N+M} \sum_{j \in S_i} c_{ij}, \quad (17)$$

where

$$c_{ij} = \begin{cases} 1 & \text{if } j \in S_i, \\ 0 & \text{if } j \notin S_i, \end{cases}$$

with c_{ij} denoting the connectivity between i and j nodes, and S_i a set of neighbors of the node i . The correction algorithm is described in detail by Niewiadomska-Szynkiewicz and Marks (2009).

Vecchio *et al.* (2012) developed a location system in which a two-objective localization problem is formulated and an evolutionary algorithm is used to solve it. Two performance functions are concurrently minimized. The first one is defined in (10). The second cost function is defined

as follows:

$$J_{FV} = \sum_{k=1}^M \left(\sum_{j \in S_k} \delta_{kj} + \sum_{j \in \tilde{S}_k} (1 - \delta_{kj}) \right) + \sum_{i=1}^N \left(\sum_{j \in S_i} \delta_{ij} + \sum_{j \in \tilde{S}_i} (1 - \delta_{ij}) \right), \quad (18)$$

where $\delta_{ij} = 1$ if $\tilde{d}_{ij} > r_i$ and 0 otherwise, and $\tilde{S}_i = \{(i, j) : \|x_i - x_j\| > r_i\}$. Hence, the goal of the J_{FV} function is to count the number of connectivity constraints that are not satisfied by the current estimated locations of target nodes. The authors claim that their approach outperforms the SA-based localization algorithm proposed by Kannan *et al.* (2006). The simulation results presented by Vecchio *et al.* (2012) confirm the good performance of the algorithm.

3.4. Mobile assisted localization. Recently, some localization systems have used mobile devices to facilitate the localization and improve accuracy and precision of the results. It is assumed that these mobile nodes are equipped with a GPS receiver. Different numbers of mobile nodes with different hardware capabilities can be assumed. It depends on the localization algorithm, the size of the network and the environment. Zhang and Yu (2010) propose a localization scheme using a single mobile anchor node equipped with a directional antenna. Each non-anchor has to be equipped with an omni-antenna. The goal of the anchor is to move through the network formed by stationary sensors and broadcast its location information. The triangulation technique is used to estimate coordinate vectors of non-anchors in a network. The system was verified through simulations. Zhang and Yu (2010b) claim that their method gives high accuracy solutions and is energy efficient. Another system with a single mobile node is described by Sichertiu and Ramadurai (2004). In this approach, the RSSI technique is used for inter-node distance estimation. The application of the RSSI for estimation of the distance between stationary non-anchors and mobile anchor is proposed by Yu *et al.* (2008). The location system with a mobile anchor node moving in a sensor field and periodically broadcasting a beacon message with its current location is described by Su *et al.* (2005). Many other systems are presented in the literature.

4. Performance evaluation

We validated the selected location systems through simulation and experiments using testbed implementation. All simulations were performed on an Intel Core2 Duo E6600 2.4 GHz PC with 2GB RAM using our simulator, which employs Link Layer Model for MATLAB described by Zuniga and Krishnamachari (2004) for network model generation. The experiments in our WSN laboratory were

Table 1. Localization errors: DV-hop and DV-distance methods (varying number of anchor nodes).

Method	Number of anchor nodes		
	20	40	60
DV-hop	71.53 %	64.82 %	56.39 %
DV-distance	50.68 %	43.54 %	28.53 %

performed using testbed implementation involving MicaZ motes manufactured by Crossbow. The goal of the first series of experiments was to compare the accuracy of two techniques: range-free and range-based using the RSSI for inter-node distance estimation. The goal of the second series of tests was to compare various approaches to coordinate vector calculation. The key metric for evaluating location systems was the accuracy of the location estimates versus deployment, equipment, communication and computational cost. To evaluate the accuracy of the tested location systems we used the mean error LE between the estimated and the true physical location of the non-anchor nodes in the network defined in (1).

4.1. Range-free vs. range-based techniques. The goal of the first series of experiments was to compare the efficiency of range-free and range-based techniques. Two variants of the APS, i.e., DV-hop and DV-distance, were considered. We performed simulations for a WSN formed by 200 nodes randomly distributed in a square region $[0, 1] \times [0, 1]$. The number of anchor nodes was 10%–30% of all nodes. We assumed that all RSSI measurement errors were independent zero-mean Gaussian random variables of unit variance. Hence, in our simulations with the DV-distance method each inter-node distance \tilde{d}_{ij} was assumed to be

$$\tilde{d}_{ij} = d_{ij}(1 + \text{randn}() \cdot n_f), \quad (19)$$

where d_{ij} denotes the true physical distances between pairs of nodes and n_f the assumed noise factor.

Selected results, i.e., the values of localization errors calculated due to (1) in the case of two variants of the APS, a varying number of evenly distributed anchor nodes, and the noise factor $n_f = 20\%$, are collected in Table 1. The presented results show that the DV-distance variant of the APS produces much more accurate localization with respect to DV-hop. The localization error decreases as the number of anchors increases.

The localization accuracy strongly depends on anchor nodes deployment in the network. The results for evenly distributed 160 non-anchors and evenly and unevenly distributed 40 anchors were compared. Two scenarios were tested:

A: anchors and non-anchors evenly distributed in the deployment region;

Table 2. Localization errors: DV-hop and DV-distance (varying deployment of anchor nodes).

Method	Anchor deployment	
	A	B
DV-hop	64.82 %	345.43 %
DV-distance	43.54 %	278.90 %

Table 3. Localization errors. Comparison of results obtained via simulation and testbed implementation.

Method	Simulation	Testbed network
DV-hop	55.34 %	62.72 %
DV-distance	16.67 %	49.87 %

B: anchors distributed in the quarter of the deployment region, non-anchors evenly distributed.

DV-distance, similarly to DV-hop, gives an unsatisfactory solution for networks with unevenly distributed anchor nodes. The differences between estimated and true physical distances for some pairs of nodes are bigger than the radio range (see Table 2). This is a result of a lack of enough neighboring nodes with known location and/or the flip ambiguity phenomenon. In our experiments we did not execute any correction procedure.

The next step was to apply the DV-hop and DV-distance protocols to the testbed network of sensors and compare localization errors obtained in simulations with those obtained for practical evaluation in the laboratory. The experiment was performed in a very large hall. The testbed WSN was formed by 8 MicaZ motes and one base station, all operating under the TinyOS system. Four nodes (anchors) were equipped with the GPS. In the case of simulations, the noise factor n_f in the distance calculation (19) was equal to 10%. The results collected in Table 3 confirm the guess that due to distance measurement errors and some weakness of the APS method it is really difficult to obtain satisfactory accuracy in location estimation of nodes in a real network. The coordinates were estimated with evident errors. When comparing the simulation and testbed results, we can observe significant deterioration in the case of the DV-distance method. It is involved by measurement errors. In most works the performance of location systems is evaluated through simulation with a noise factor usually assumed to be equal to 10–20%. Our experiments with the testbed network were performed in realistic conditions. The accuracy of RSS measurements is limited mainly in in-door environments and a noise factor can be much higher than 10–20%. Range-based techniques are very sensitive to the accuracy of inter-node distance estimation. Unfortunately, the presented results demonstrate that application of location systems to estimate the position of nodes in a real network can give results much worse than those calculated through simulations.

Table 4. Localization errors: SDP, TGA and TSA methods (varying deployment of anchor nodes).

Method	Anchor deployment		
	A	B	C
SDP	0.18 %	174.91 %	330.56 %
TGA	3.80 %	20.61 %	56.06 %
TSA	0.13 %	1.78 %	1.81 %

Table 5. Computation times in seconds: SDP, TGA and TSA methods (varying deployment of anchor nodes).

Method	Anchor deployment		
	A	B	C
SDP	6.95	5.51	6.25
TGA	2.85	2.23	2.90
TSA	0.43	0.44	0.47

4.2. Deterministic vs. stochastic programming. The next series of experiments was to compare the performance of location systems based on linear and nonlinear optimization. A detailed analysis of the properties and performance of TSA, the TGA, SA and SDP based on the results of numerical tests performed for a wide range of network system configurations including the size of the network (number of nodes), the number of anchor nodes, the radio range, the distance estimation error and the computation time is presented by Niewiadomska-Szynkiewicz and Marks (2009). Selected results, i.e., values of localization errors calculated due to (1) for TSA, the TGA and SDP, are collected in Tables 4–6. The simulations were performed for a WSN consisting of 200–10000 nodes with randomly generated positions in a square region $[0, 1] \times [0, 1]$. In all simulations the percentage of anchor nodes was fixed to 10%. The value of the noise factor n_f in distance estimation was set to 10%. Three simulation scenarios were considered:

- A: anchors and non-anchors evenly distributed in the deployment region;
- B: anchors distributed only in the half of the deployment region, non-anchors evenly distributed;
- C: anchors distributed in the quarter of the deployment region, non-anchors evenly distributed.

The simulation results for the WSN with 200 nodes, the radio range r equal to 0.18, and three topologies: A, B and C are presented in Tables 4 and 5. The above results vary when we change the position of sensor nodes. The same number of anchor nodes gives different results if placed differently. Table 4 demonstrates that SDP-based location systems are very sensitive to uneven deployment of anchor nodes. Nonconvex methods produced much more accurate localization with respect to SDP. Moreover, the TSA method was 15 times faster than SDP (see Table 5).

In contrast to the SDP method, TSA gives satisfactory accuracy in location estimation for WSNs with anchor nodes deployed only in a part of the region to be covered by sensors (scenarios B and C).

From the experiments it was observed that all discussed methods give unsatisfactory solutions for networks with both anchors and non-anchors located only in some parts of the domain. The results presented in Table 6 demonstrate that the TSA method yields the location estimation accuracy even in the case of a large size of the network. Unfortunately, the computation time increases proportionally to the square of the network dimension. The

Table 6. Localization errors and computation times for different sizes of a network with evenly distributed anchor nodes.

Number of nodes	LE [%]	t [s]
200	0.1275	0.4
1000	0.1387	8.0
2000	0.1081	33.6
5000	0.1581	189.8
10000	0.1193	790.4

above results indicate that metaheuristics are versatile and attractive techniques for WSN localization. However, it is commonly known that efficiency and robustness of these methods strongly depend on different control parameters of the algorithm. To obtain a general-purpose algorithm to solve the localization problem, the parameters should be tuned for various network sizes and topologies. Unfortunately, it is a time-consuming process with no guarantee of success. The tuning of control parameters to obtain a fast and robust TSA algorithm to solve localization problems in wide range of WSN applications is described in detail by Marks and Niewiadomska-Szynkiewicz (2009).

4.3. Centralized vs. distributed systems. Tables 4–6 present the results of simulations performed for localization methods implemented in a centralized manner. Centralized localization algorithms depend on the nodes transmitting data to a base station (or a central computer), where computations are performed to determine the estimated positions of all nodes in the network. The results collected in Table 6 point out that in the case of large scale networks, the computation time can limit the application of centralized methods. Therefore, distributed techniques where each node calculates its position using only information about its neighbor have been developed and evaluated by many researchers. A distributed implementation offers a significant reduction in computation and communication requirements and is fault tolerant. A disadvantage is the loss of information due to parallel computation and an incomplete network map. The results of multiple simulations of the distributed version of the TSA method

confirmed that the accuracy of location estimation is much worse for the distributed TSA than the centralized one.

It is obvious that algorithms for nonconvex optimization have to perform many iterations to achieve a stable solution. The cost of each iteration, in energy terms, is different for a centralized and a distributed scheme. A centralized algorithm in large networks requires all sensor measurements to be sent over multiple hops to a central processor, while a distributed algorithm requires only local information exchanges between neighboring nodes, but many such local exchanges may be required. The number of additional transmissions increases with the number of iterations necessary to reach a global solution. In the case of centralized implementation, energy consumption is asymmetric—the multi-hop transmission stresses nodes close to the central station more than other nodes. Fortunately, the size of transmitted messages is very small. In contrast, distributed algorithms require many local information exchanges between neighboring nodes. Each iteration of the optimization algorithm involves inter-node data transmission. Moreover, each node performs calculations to estimate its coordinate vector. Hence, in the case of complex algorithms used for position estimation, distributed implementation may result in higher energy consumption.

5. Summary and conclusion

This paper provides a short overview of localization strategies and systems. We outline the main properties and criteria that should be considered while estimating the location of nodes with unknown positions in a network. The results of experimental evaluation of localization techniques described in the literature and our results presented in this paper show pros and cons of various methods and challenges in localization. The performance of location systems strongly depends on the accuracy of inter-node distance estimation. The results of practical evaluation of range-free and range-based methods show that localization techniques based on inter-node distance estimation produce much more accurate position estimates with respect to range-free techniques. Unfortunately, distance measurement techniques require special hardware support, and the accuracy of distance estimation varies for each method.

Measurements of the signal strength are relatively cheap and simple to implement in hardware. Moreover, nodes in the network communicate with their neighbors. Hence the RSS signal can be measured by each receiver during normal data communication without additional bandwidth or energy requirements. However, the accuracy of RSS measurements is limited mainly in indoor environments (Elnahrawy *et al.*, 2004). If they are to be a part of a robust localization system, the sources of the measurement error have to be well understood. Moreover, to

convert the measurements into inter-node distances, calibration is required to estimate the parameters for the channel model.

Most range-based algorithms require the existence of anchor nodes that have accurate location information. The localization error decreases as the number of anchors increases. Increasing the density of anchor nodes makes localization easier but increases the deployment costs. Problems occur in the case of unevenly distributed anchor and/or non-anchor nodes. Simple geometrical techniques give accurate location estimates for network topologies with a sufficient number of anchor nodes evenly distributed in the deployment region. Techniques based on MDS can achieve quite accurate location estimates but with high computational cost $\mathcal{O}(L^3)$, where L is the number of nodes.

The results of simulation studies performed for some MDS-based algorithms and various network topologies are discussed by Mao and Fidan (2009). The conclusion is that they are not a good solution for sparse networks. A comparative study of SPE-based and MDS-based methods shows that, using SPE, we can obtain localization results comparable with distributed MDS-based methods with much reduced computational cost. The SPE technique allows reducing computational effort, and scales linearly with the number of nodes in the network. In general, in the case of noisy distance measurements, application of stochastic nonconvex optimization significantly outperforms simple geometrical techniques and convex programming.

The improvement of solution accuracy is mainly observed in the case of unevenly distributed network nodes. Systems give better results with less computational cost. Our experimental results demonstrate that hybrid techniques are competitive to the other solutions. Systems that combine geometrical and nonconvex optimization techniques extended with correction of temporary solutions provide significant robustness and improve the accuracy compared with simple trilateration, as well as convex and nonconvex optimization. Hence, from the perspective of the location estimation accuracy, the suggestion is to use centralized range-based hybrid location systems with measurement techniques according to the available hardware (if possible, simultaneously more than one technique can be used to improve the measurement accuracy) and additional correction of localization errors. Unfortunately, complex location systems involve high communication and computational costs, and then a high energy consumption. In many practical applications, the limited resources of real-life networks can constrain us to use low complexity, energy aware and distributed localization strategies. Although numerous localization strategies and systems have been proposed and described in the literature, the development of a robust and scalable technique to estimate the location of nodes forming a WSN with high ac-

curacy, minimal hardware cost and computational burdens is still a challenging task.

Acknowledgment

This work was partially supported by the National Science Centre grant NN514 672940. The author would like to thank M. Marks and J. Skomial for their assistance with the experiments.

References

- Ahmed, A.A., Shi, H. and Shang, Y. (2005). Sharp: A new approach to relative localization in wireless sensor networks, *Proceedings of IEEE ICDCSW'05, Columbus, OH, USA*, Vol. 9, pp. 892–898.
- Akyildiz, I. and Vuran, M. (2010). *Wireless Sensor Networks*, John Wiley & Sons, West Sussex.
- Aloor, G. and Jacob, L. (2010). Distributed wireless sensor network localization using stochastic proximity embedding, *Computer Communications* **33**(6): 745–755.
- Anderson, B., Mao, G. and Fida, B. (2007). Wireless sensor network localization techniques, *Computer Networks* **51**(10): 2529–2553.
- Barsocchi, P., Lenzi, S., Chessa, S. and Giuntau, G. (2009). Virtual calibration for RSSI-based indoor localization with ieee 802.15.4, *Proceedings of the IEEE International Conference on Communications (ICC), Dresden, Germany*, pp. 1–5.
- Benkic, K., Malajner, M., Planinsic, P. and Cucej, Z. (2008). Using RSSI value for distance estimation in wireless sensor networks based on ZigBee, *Proceedings of the 15th International Conference on Systems, Signals and Image Processing, IWSSIP 2008, Bratislava, Slovakia*, pp. 303–306.
- Bernardeschi, C., Masci, P. and Pfeifer, H. (2008). Early prototyping of wireless sensor network algorithms in PVS, in M.D. Harrison and M.-A. Sujan (Eds.), *Computer Safety, Reliability and Security*, Lecture Notes in Computer Science, Vol. 5219, Springer-Verlag, Berlin/Heidelberg/New York, NY, pp. 346–359.
- Beutel, J. (2005). *Handbook of Sensor Networks Compact Wireless and Wired Sensing Systems*, CRC Press, Boca Raton, FL.
- Biswas, P. and Ye, Y. (2004). Semidefinite programming for ad hoc wireless sensor network localization, *IPSN '04: Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks, Berkeley, CA, USA*, pp. 46–54.
- Boyd, S., Ghaoui, L. E., Feron, E. and Balakrishnan, V. (1994). *Linear Matrix Inequalities in System and Control Theory*, Studies in Applied Mathematics, Vol. 15, SIAM, Philadelphia, PA.
- Cheng, X., Thaeler, A., Xue, G. and Chen, D. (2004). TPS: A time-based positioning scheme for outdoor wireless sensor networks, *Proceedings of IEEE INFOCOM'04, Hong Kong, China*, Vol. 4, pp. 2685–2696.
- Chuang, P. and Wu, C. (2008). An effective PSO-based node localization scheme for wireless sensor networks, *Proceedings of the 9th International Conference on Parallel and Distributed Computing Applications and Technologies, Dunedin, New Zealand*, pp. 187–194.
- Costa, J., Patwari, N. and Hero, A. (2006). Distributed weighted-multidimensional scaling for node localization in sensor networks, *ACM Transactions on Sensor Networks* **2**(1): 39–64.
- Di Caro, G. (2003). Analysis of simulation environments for mobile ad hoc networks, *Technical Report IDSIA-24-03*, IDSIA, Manno.
- Elnahrawy, E., Li, X. and Martin, R. (2004). The limits of localization using signal strength: A comparative study, *Proceedings of the IEEE Sensor and Ad hoc Communications and Networks Conference (SECON 2004), Santa Clara, CA, USA*, pp. 333–341.
- Heurtefeux, K. and Valois, F. (2008). Distributed qualitative localization for wireless sensor networks, in D. Coudert, D. Simplot-Ry and I. Stojmenovic (Eds.), *Ad-hoc, Mobile and Wireless Networks*, Lecture Notes in Computer Science, Vol. 5198, Springer-Verlag, Berlin/Heidelberg, pp. 218–229.
- Hu, L. and Evans, D. (2004). Localization for mobile sensor networks, *ACM MobiCom, Philadelphia, PA, USA*, pp. 45–57.
- Kannan, A., Mao, G. and Vucetic, B. (2005). Simulated annealing based localization in wireless sensor network, *LCN'05: Proceedings of the IEEE Conference on Local Computer Networks, 30th Anniversary, Washington, DC, USA*, pp. 513–514.
- Kannan, A., Mao, G. and Vucetic, B. (2006). Simulated annealing based wireless sensor network localization with flip ambiguity mitigation, *63rd IEEE Vehicular Technology Conference, Melbourne, Australia*, pp. 1022–1026.
- Karakehayov, Z. (2009). Multobjective design of wireless ad hoc networks: Security, real-time and lifetime, *Journal of Telecommunications and Information Technology* (2): 13–21.
- Karl, H. and Willig, A. (2005). *Protocols and Architectures for Wireless Sensor Networks*, John Wiley & Sons, West Sussex.
- Kasch, W., Ward, J. and Andrusenko, J. (2008). Wireless network modeling and simulation tools for designers and developers, *IEEE Communications Magazine* **47**(3): 120–127.
- Lee, J., Chung, W. and Kim, E. (2010). A new range-free localization method using quadratic programming, *Computer Communications* **34**(8): 998–1010.
- Li, C., Li, Y., Shen, Y., Liu, L. and Cao, Q. (2010). An optimization algorithm for wireless sensor networks localization using multiplier method, *Proceedings of the 3rd International Joint Conference on Computational Science and Optimization, Huangshan, Anhui, China*, pp. 337–341.

- Liu, L. and E, M. (2010). Localization for wireless sensor networks by combining TFDA and FMCW, *Proceedings of the IEEE International Conference on Mechatronics and Automation, Xian, China*, pp. 945–950.
- Magnani, A. and Leung, K. (2007). Self-organized, scalable GPS-free localization of wireless sensors, *IEEE WCNC, Hong Kong, China*, pp. 3798–3803.
- Mao, G. and Fidan, B. (2009). *Localization Algorithms and Strategies for Wireless Sensor Networks*, Information Science Reference, Hershey, PA.
- Marks, M. (2010). A survey of multi-objective deployment in wireless sensor networks, *Journal of Telecommunications and Information Technology* (3): 36–41.
- Marks, M. and Niewiadomska-Szynkiewicz, E. (2009). Multiobjective approach to localization in wireless sensor networks, *Journal of Telecommunications and Information Technology* (3): 59–67.
- Medidi, M., Slaaen, R., Zhou, Y., Mallery, C. and Medidi, S. (2006). Cluster-based localization in wireless sensor networks, *Proceedings of SPIE, Wireless Sensing and Processing* 6248(62480J): 1–8.
- Motter, P., Allgayer, R., Müller, I. and de Freitas, E. (2011). Practical issues in wireless sensor network localization systems using received signal strength indication, *Proceedings of the Sensors Applications Symposium (SAS), San Antonio, TX, USA*, pp. 227–232.
- Niculescu, D. and Nath, B. (2001). Ad hoc positioning system (APS), *GLOBECOM: Global Telecommunications Conference, San Antonio, TX, USA*, Vol. 5, pp. 2926–2931.
- Niewiadomska-Szynkiewicz, E. and Marks, M. (2009). Optimization schemes for wireless sensor network localization, *International Journal of Applied Mathematics and Computer Science* 19(2): 291–302, DOI: 10.2478/v10006-009-0025-3.
- Olveczky, P. and Thorvaldsen, S. (2007). Formal modeling and analysis of the OGDC wireless sensor network algorithm in real-time maude, *Proceedings of the 9th IFIP International Conference on Formal Methods for Open Object-Based Distributed Systems, Paphos, Cyprus*, pp. 122–140.
- Pawlikowski, K., Jeong, H. and Lee, J. (2002). On credibility of simulation studies of telecommunication networks, *IEEE Communications Magazine* 40(1): 132–139.
- Rappaport, T. (2002). *Wireless Communications: Principles and Practice*, Communications Engineering and Emerging Technologies Series, Prentice Hall, Upper Saddle River, NJ.
- Salzmann, J., Behnke, R., Gorski, P. and Timmermann, D. (2011). HyPAERLoc: Plausible hybrid localization for wireless sensor networks, *Proceedings of SENSORCOMM 2011, Nice, France*, pp. 51–57.
- Santi, P. (2006). *Topology Control in Wireless Ad Hoc and Sensor Networks*, John Wiley & Sons, West Sussex.
- Sarigiannidis, G. (2007). *Localization For Ad Hoc Wireless Sensor Networks*, LL, Delft.
- Savvides, A., Han, C. and Strivstava, M. (2001). Dynamic fine-grained localization in ad-hoc networks of sensors, *Proceedings of ACM MobiCom, Rome, Italy*, pp. 166–179.
- Sayadnavard, M., Haghghat, A. and Abdechiri, M. (2010). Wireless sensor network localization using imperialist competitive algorithm, *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology, Chengdu, China*, Vol. 9, pp. 818–822.
- Shang, Y., Ruml, W., Zhang, Y. and Fromherz, M. (2004). Distributed weighted-multidimensional scaling for node localization in sensor networks, *IEEE Transactions on Parallel and Distributed Systems* 15(11): 961–974.
- Shekofteh, S., Yaghmaee, M., Khalkhali, M. and Deldari, H. (2010). Localization in wireless sensor networks using tabu search and simulated annealing, *Proceedings of the 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore*, pp. 752–757.
- Shi, Q., He, C., Chen, H. and Jiang, L. (2010). Distributed wireless sensor network localization via sequential greedy optimization algorithm, *IEEE Transactions on Signal Processing* 58(6): 3328–3340.
- Shi, Q., He, C., Ljang, L. and Luo, J. (2008). Sensor network localization via nondifferentiable optimization, *Proceedings of IEEE GLOBECOM, New Orleans, LA, USA*, pp. 1–5.
- Shi, Q., He, C., Ljang, L. and Luo, J. (2009). Normalized incremental subgradient algorithm and its application, *IEEE Transactions on Signal Processing* 57(10): 3759–3774.
- Shu, J., Zhang, R., Liu, L., Wu, Z. and Zhou, Z. (2009). Cluster-based three-dimensional localization algorithm for large scale wireless sensor networks, *Journal of Computers* 4(7): 585–592.
- Sichitiu, M.L. and Ramadurai, V. (2004). Localization of wireless sensor networks with a mobile beacon, *Proceedings of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems, Fort Lauderdale, FL, USA*, pp. 174–183.
- Srirangarajan, A., Tewfik, H. and Luo, Z. Q. (2008). Distributed sensor network localization using SOCP relaxation, *IEEE Transactions on Wireless Communication* 7(12): 4886–4895.
- Su, K. F., Ou, C. H. and Jiau, H. C. (2005). Localization with mobile anchor points in wireless sensor networks, *IEEE Transactions on Vehicular Technology* 54(3): 1187–1197.
- Tam, V., Cheng, K. and Lui, K. (2006). A descent-based evolutionary approach to enhance position in wireless sensor networks, *Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06), Washington, DC, USA*, pp. 568–574.
- Tseng, P. (2007). Second-order cone programming relaxation of sensor network localization, *SIAM Journal on Optimization* 18(1): 156–185.
- Vecchio, M., Lopez-Valcarce, R. and Marcelloni, F. (2012). A two-objective evolutionary approach based on topological constraints for node localization in wireless sensor networks, *Applied Soft Computing* 12(7): 1891–1901.

- Verdone, R., Dardari, D., Mazzini, G. and Conti, A. (2008). *Wireless Sensor Networks and Actuator Networks. Technologies, Analysis and Design*, Elsevier, London.
- Vo, N., Vo, D., Challa, S. and Lee, S. (2008). Nonmetric MDS for sensor localization, *Proceedings of the International Symposium on Wireless Pervasive Computing, Karlsruhe, Germany*, pp. 396–400.
- Wang, Z., Zheng, S., Ye, Y. and Boyd, S. (2008). Further relaxation of the semidefinite programming approach to sensor network localization, *SIAM Journal on Optimization* **19**(2): 655–673.
- Wessels, A., Wangb, X., Laurb, R. and Langa, W. (2010). Dynamic indoor localization using multilateration with RSSI in wireless sensor networks for transport logistics, *Procedia Engineering* **5**: 220–223.
- Whitehouse, K., Karlof, C. and Culler, D. (2007). A practical evaluation of radio signal strength for ranging-based localization, *Mobile Computing and Communications Review* **11**(1): 41–52.
- Yu, G., Yu, F. and Feng, L. (2008). A three dimensional localization algorithm using a mobile anchor node under wireless channel, *Proceedings of the IEEE International Conference on Neural Networks, Hong Kong, China*, pp. 477–483.
- Zhang, B. and Yu, F. (2010a). An event-triggered localization algorithm for mobile wireless sensor networks, *Proceedings of the 2nd IEEE International Conference on Future Computer and Communication, Melbourne, Australia, Vol. 1*, pp. 250–253.
- Zhang, B. and Yu, F. (2010b). A feasible localization algorithm for wireless sensor networks using directional antenna, *Proceedings of the 12th IEEE International Conference on High Performance Computing and Communications, Melbourne, Australia*, pp. 354–361.
- Zuniga, M. and Krishnamachari, B. (2004). Analyzing the transitional region in low power wireless links, *Proceedings of the 1st International Conference on Sensor and Ad Hoc Communications and Networks, SECON, Santa Clara, CA, USA*, pp. 517–526.



Ewa Niewiadomska-Szynkiewicz, Ph.D. (1995), D.Sc. (2005), a professor of control and computer engineering at the Warsaw University of Technology, head of the Complex Systems Group. She is also the director for research of the Research and Academic Computer Network (NASK), and the author and co-author of three books as well as over 100 journal and conference papers. Her research interests focus on complex systems modeling, optimization and control, computer simulation, parallel computations, computer networks and *ad-hoc* networks. She has been involved in a number of research projects (including four EU projects), coordinated group activities, and managed organization of a number of national-level and international conferences.

Received: 18 February 2011

Revised: 16 September 2011