JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 34, 451-468 (2018)
DOI: 10.6688/JISE.201803_34(2).0009

Building Coordinate System of Sensor Nodes Using Self-configurable Grid-based Approach

PEI-HSUAN TSAI
Institute of Manufacturing Information and Systems
National Cheng Kung University
Tainan, 701 Taiwan
E-mail: phtsai@mail.ncku.edu.tw

Establishing relative coordinate system is essential for many applications of sensor networks. This paper proposes a computation scalable scheme for establishing relative coordinate system of sensor networks using distance measurements between sensors. A distributed selection algorithm is proposed to select a small number of nodes from all sensors to serve as virtual grid point and establish the grid relative coordinate system. The grid coordinate system forms the backbone of localization for static and mobile nodes, whereby the non-grid nodes can compute their positions via message exchange with their neighboring grid nodes. This paper shows by both mathematical analysis and simulations that the localization error of the scalable grid approach is upper bounded by half the grid width (*i.e.*, the length of two adjacent grid points). The results also indicate that the precision of grid coordinate system can be adaptively adjusted by the grid width and the density of grid nodes to meet target applications. For those non-grid nodes, the results show that the localization precision of non-grid nodes would be bounded by one grid-cell error.

Keywords: sensor networks, virtual coordinate system, distributed localization, error-bounded, low complexity

1. INTRODUCTION

In recent years, wireless sensor networks have become an active research topic because of its wide variety of applications, such as battlefield surveillance, smart environments, healthcare and disaster relief [1, 2]. Wireless sensor networks also draw a lot of research challenges, such as sensing coverage controls, energy-efficient communication protocols and data aggregations. The essential assumption in many applications is that the gathered data of a sensor node has to combine with location information to instantly let control center know where the data coming from or where an event has occurred. Geographic information can be obtained by installing a Global Positioning System (GPS) receiver on each node in the network, but the weakness of high cost makes it infeasible to install in every node. Feasible solutions are to equip only a few nodes with GPS, called *anchor nodes* or *beacon nodes*, and the other nodes can use the estimated distances between anchor nodes and them to infer their position. However, the computational complexity of establishing the coordinate system increases exponentially with the number of distances used in calculating the coordinates of nodes.

Although location information is crucial in wireless sensor networks, most applications only require coarse localization accuracy. For example, for many location-based

Received July 13, 2016; revised March 17, 2016 & May 11, 2017; accepted June 1, 2017.
Communicated by Jan-Jan Wu.

routing protocols [3-6], grid-based coordinate system provides sufficient geographic information. Another example is that, in many event-based applications, we are interested in the region where an event takes places rather than the precise location of the node issuing the event. In this case, region-based coordinates are accurate enough for supporting target applications.

In this paper, we propose a computation scalable localization scheme for establishing relative coordinate system of sensor networks using distance measurements between sensors. The basic idea of the proposed scheme is to select a small number of nodes from sensors which will serve as virtual grid points to establish the grid coordinate system. In the paper, the chosen nodes are called *grid-nodes* and the other non-selected nodes are referred to as *non-grid nodes*. The grid coordinate system acts as the backbone of node localization whereby the non-grid nodes can compute their positions via information exchange with their adjacent grid nodes. This paper conducts mathematical analysis and extensive simulations to show that the precision of grid coordinate system can be bounded by the grid width, *i.e.*, the grid length between two adjacent grid points. Furthermore, the precision of node localization is highly dependent on the number of grid nodes. The more the number of grid nodes, the higher the accuracy of localization. The localization scheme can be adaptively adjusted to balance the tradeoff between computation complexity and precision of the established coordinate system according to target applications. In summary, the proposed localization scheme has the following features:

(1) **Adaptive:** The precision of the proposed localization technique can be adjusted adaptively according to the requirements of target applications so that all nodes in the network could be positioned efficiently.
(2) **Self-configurable:** The relative coordinate system can be established without the existence of anchor nodes (the sensor nodes that know their global coordinates). The coordinates, however, can also be transformed into the global coordinates if there are at least three nodes knowing their global position.
(3) **Scalable:** The proposed scheme could be computation scalable for large-scale sensor networks.

The rest of this paper is organized as follows. Section 2 introduces the related works. Section 3 presents the proposed localization scheme. The simulation results are given in Section 4. Finally, we draw the conclusion of this paper in Section 5.

2. RELATED WORK

The localization schemes can be divided into two categories: *range-based* and *range-free*. The range-based localization schemes use the measured distances and/or angles between sensor nodes to estimate the location. There are several technologies that can be used for this purpose such as Time of Arrival (TOA), Time Difference of Arrival (TDOA), or Received Signal Strength Indicator (RSSI). Generally, RSSI is easier to implement, while ToA may have higher accuracy. The work [18] presented a probabilistic and constraint-based approach to perform node localization by using RSSI-based distance measurements. In [19], a class of algorithms for fine-grained localization called

Sweeps was proposed to estimate the location by triangulation or rigidity.

On the contrary, the range-free schemes perform node localization by network connectivity. Each sensor estimates its location by connectivity information with its neighbors that can be gathered by message exchange with neighboring nodes. It only achieves coarse localization accuracy in most cases, but it is easy to implement and applicable to large network. As range-free positioning system, DV-Hop is the typical representation. DV-Hop approach assumes that the network comprises a small number of anchor nodes [7]. The other non-anchor nodes calculate their locations based on the averaged one-hop distance to every anchor node in the network. If two sensor nodes are not within the transmission range of each other, a simple and effective scheme, called multi-hop distance estimation model, was proposed to estimate the distance between two remote nodes [14]. This model utilizes the correlation of the Euclidean distance and the corresponding shortest path length between two sensor nodes in the network for a given node distribution. Based on this model, a node can estimate the Euclidean distance to another node by sending a control packet.

Many localization schemes have been proposed to solve data delivery problems and improve localization performance, including power saving, communication costs, localization accuracy and so on, in the past few years [7-16, 18-27]. The work [11] addressed the problem of sensor localization in wireless networks in a multipath environment and proposed a distributed and cooperative algorithm based on belief propagation, which allows sensors to cooperatively self-localize with respect to a single anchor node in the network, using range and direction of arrival measurements. The research study [8] investigated neighborhood collaboration based distributed cooperative localization of all sensors in a particular network with the convex hull constraint. Three iterative self-positioning algorithms were present for independent implementation at all individual sensors of the considered network. The authors in [9] addressed the localization with incompletely paired mutual distances and proposed a Partially Paired Locality Correlation Analysis (PPLCA) algorithm. The work [10] described a novel SDP-based formulation for analyzing node localizability and providing a deterministic upper bound of localization error. The SDP-based formulation gives a sufficient condition for unique node localizability for any frameworks. A compressive sensing (CS) based approach for node localization in wireless sensor networks was presented in [12]. The study [13] proposed a localization system in which localization information is obtained through a probability based algorithm that requires the solving of a nonlinear optimization problem. The paper [27] proposes a grid-based working node (WN) selection approach for wireless sensor networks to save power and reduce interference in wireless channel.

One of the important issues in wireless sensor network is routing. Instead of geographic routing which uses the physical location as the node address, many research studies proposed virtual coordinate system and geography-free routing protocols to guarantee packet delivery without the computation and storage of the global topological features [2, 24, 25]. Different from traditional geography routing approaches, they do not attempt to approximate physical coordinates based on hop distances. They construct a virtual topology which is possibly unrelated to the physical topology of the network.

Compared to these previous studies, this paper provides an adaptive, self-configurable and scalable localization scheme which improves the localization accuracy and to lower the cost of previous works. This paper also conducts mathematical analysis and extensive simulations to show that the precision of grid coordinate system can be bounded by the grid width. DRLS proposed in [26] is a grid-scan algorithm with vector-based refinement scheme to estimate the locations of nodes. In this paper, DRLS is simulated to compare with our method.

3. THE PROPOSED LOCALIZATION ALGORITHM

Given a set of sensor nodes S deployed in an interested region F, the coordinate system is established in two steps. First, a distributed algorithm is proposed to select a small number of nodes serving as grid nodes to establish the grid coordinate system. The grid nodes exchange information with the other grid nodes to obtain distance measurements whereby the grid coordinate system can be established by minimizing an error objective function. Second, the grid coordinate system then serves as the backbone of the coordinate system whereby the non-grid nodes can compute their own position based on the grid coordinate system.

3.1 Grid Coordinate System Establishment

3.1.1 Distributed grid node selection algorithm

To create the grid coordinate system, all the nodes in the network need to estimate the distances between themselves and their communication neighbors. There are several technologies that can be used for this purpose such as Received Signal Strength Indicator (RSSI) and Time-of-Arrival (ToA). Generally, RSSI is easier to implement, while ToA may have higher accuracy. If two grid nodes are adjacent, ToA or RSSI can be employed to estimate their distances. If two grid nodes are not within the transmission range of each other, a simple and effective scheme, called multi-hop distance estimation model, can be utilized to estimate the distance between two remote grid nodes [14]. The model utilizes the correlation of the Euclidean distance and the corresponding shortest path length between two nodes in the network for a given node distribution. Based on this model, a node s_i can estimate the Euclidean distance to another node s_j by sending a control packet that includes a *route length field* with initial value of zero. When an intermediate node receives the control packet, it adds the one-hop distance between itself and the previous node, obtained by using ToA or RSSI, to the route length field. Upon receiving the control packet, node s_j sends it back to node s_i. After node s_i receives the return control packet from s_j, it can read the route length field and estimates the distance between s_i and s_j according to the multi-hop distance estimation model.

Since the problem of finding an optimal subset of nodes to be grid nodes for establishing the grid coordinate system is computationally hard, this paper proposes a heuristic method for selecting an appropriate set of grid nodes and shows its effectiveness by experiments. The method is designed based on the following observations. The process of building grid coordinate system is to determine a coordinate of grid point, referred as grid coordinate, for each of the chosen nodes. This can be viewed as virtually mapping the locations of chosen nodes onto the grid-points of grid coordinate system. Therefore, for a group of nodes that are closely adjacent to each other, it is proposed to select only one of them as grid node. This criterion of proximity is defined in this paper as the distance of grid-width, and the problem can then be more formally described as finding a subset of nodes in which the distance between any two of them are larger than the grid-width. In addition, since the non-grid nodes will need to estimate their location by exchanging messages with their nearby grid nodes once the grid coordinate system has been created, the number of grid nodes is supposed to be large enough to cover the entire coordinate system. With the two objectives, the problem of selecting grid nodes here can be described as the problem of finding the maximal independent set in a graph as follows.

Let $\tilde{d}_{i,j}$ be the estimated distance between two adjacent nodes s_i and s_j. Let w be the grid width of the grid coordinate system, *i.e.*, the distance between two adjacent grid points. We assume, without loss of generality, sensor nodes are deployed in an $L \times L$ square region, where L is multiple of w. Let $G = (V, E)$ be a graph. Each node $s_i \in S$ is represented by a vertex $v_i \in V$. There is an edge $e \in E$ between two vertices v_i and v_j if and only if $\tilde{d} < w$. An independent set of graph G is a subset $\breve{V} \subseteq V$ of vertices such that each edge in E is incident on at most one vertex in $\breve{V}$. A maximal independent set is an independent set $\breve{V}$ such that for all vertices $v \in V - \breve{V}$, the set $\breve{V} \cup \{v\}$ is not independent. Consequently, finding a maximum set of grid nodes that meets the criterion of proximity is equivalent to finding a maximal independent set of G, which is a NP-complete problem.

In this paper, we propose a distributed grid node selection algorithm, similar to the one proposed by Luby [17]. Let N_i be the set of nodes that are within the transmission range of node s_i. Let $NG_i = \{v_j \| \forall v_j \in V \text{ and } \tilde{d}_{i,j} < w\}$ be the set of nodes that their distances to s_i is smaller than the grid width w. The distributed grid node selection algorithm proceeds in rounds.

Distributed grid node selection algorithm
Initial round: Every node broadcasts a message containing its identity.
Subsequent rounds: When a node s_i collects all such messages from its neighbors, it becomes a grid node (winner) if either of the following two conditions is satisfied:

- NG_i is nonempty (there exists at least one communication neighbor s_j such that $\tilde{d}_{i,j} < w$). Moreover, s_i has the smallest Node ID when comparing its Node ID with the received Node IDs of the nodes in NG_i.
- NG_i is empty

The grid nodes (winners) send to all their neighbors a message specifying their winner state. A sensor that receives such a message from one of its neighbors becomes a non-grid node (loser). Winners and losers of a round do not participate in subsequent rounds. Repeat this procedure until all of the nodes have determined their roles. ❑

Message complexity It can be seen that the resulting set of winners, denoted by W of winners is independent since all neighbors of the winners in each round have changed their roles as losers (*i.e.*, they are removed from the graph). Besides, W is also maximal; otherwise, there exists another independent set $\hat{W}$ that strictly contains W. It implies that there is one additional winner node $j \in \hat{W} \setminus W$ such that $(i, j) \in E$ for some $i \in W$. This contradicts that $\hat{W}$ is independent. In addition, assume that the average number of nodes that decide their roles (winner or loser) in each round is k. Then the average number of broadcasting messages, denoted by B, needed is given by

$$B \le \frac{2}{n}\sum_{r=0}^{\frac{n}{k}}[n - r\cdot k] = \frac{2}{n}\cdot\left[n\cdot\frac{n+k}{k} - \sum_{r=0}^{\frac{n}{k}}(r\cdot k)\right] = \frac{2}{n}\cdot\left[n\cdot\frac{n+k}{k} - \frac{n^2+n\cdot k}{2k}\right] = \frac{(n+k)}{k}. \quad (1)$$

For example, if G is fully connected, then $k = n$ and $B \le 2$. By using the result shown by Luby in [17] that $k \ge \frac{|E|}{8}$, where $|E|$ is the number of edges in graph G, $B \le \frac{8n+|E|}{2|E|}$. Accordingly, it can be concluded that $B \le \min\left(\frac{(n+k)}{k}, \frac{8n+|E|}{2|E|}\right)$. In terms of worst-case complexity, B is upper bounded by $O(\frac{|V|}{|E|})$. In addition, it was also shown by Luby that the average number of execution rounds is $O(\log n)$.

3.1.2 Computation of grid coordinates

Let S' be the set of grid nodes obtained by using the proposed distributed grid node selection algorithm. Here we use $\tilde{d}_{i,j}$ to denote the estimated distance between node s_i and node s_j by using ToA or RSSI if s_i and s_j are adjacent; otherwise, $\tilde{d}_{i,j}$ denotes the estimated distance obtained by using the multi-hop estimation model between two remote grid nodes s_i and s_j. We assume, without loss of generality, the grid node with the lowest identify number, say s_1, will perform the calculation for creating the grid coordinate system. Once a grid node s_i has obtained a set K_i of distances to all other grid nodes, *i.e.*, $K_i = \{\tilde{d}_{i,j} \| \forall s_j \in S' \text{ and } j \ne i\}$, it sends K_i to s_1. Thus s_1 has the distance information between any two grid nodes in the network. Our objective is to calculate the grid coordinate system that minimizes the sum of the errors of the distances between any two grid nodes. Let (g_{x_i}, g_{y_i}) be the grid coordinate of node $s_i \in S'$, *i.e.*, the coordinates of grid points. Then the distance between two grid nodes s_i and s_j in the established grid coordinate system is

$$d_{ij} = \sqrt{\left(g_{x_i} - g_{x_j}\right)^2 + \left(g_{y_i} - g_{y_j}\right)^2}. \quad (2)$$

Without loss of generality, assuming that $|S'| = M$ and $S' = \{s_i \| i = 1, 2, \ldots, M\}$. The error function is defined as

$$\varepsilon(X_1, X_2, \ldots, X_M) = \sum_{j=0}^{M-1}\sum_{k=j+1}^{M}\left(d_{j,k} - \tilde{d}_{j,k}\right)^2. \quad (3)$$

Where $X_i = (g_{x_i}, g_{y_i})$, for $i = 1, 2, \ldots, M$. Then, the establishment of grid coordinate system can be formulated as the following optimization problem.

$$\{X_i^*\}_{i=1}^M = \arg\min_{\{X_i^*\}_{i=1}^M}\left\{\varepsilon\left(X_1, X_2, \ldots, X_M\right)\right\} \quad (4)$$

One of the advantages of the grid-based localization approach is that Eq. (3) can be solved efficiently since the number of grid nodes used to create grid coordinate system is usually small relative to the full set of sensors. Particularly, the number of grid nodes can be adjusted adaptively by setting the grid width in the proposed distributed selection method. In addition, the grid-based coordinate system digitizes the solution space of the optimization. Let $K = (\frac{L}{w})^2$ be the number of grid points in the established grid coordinate system. The upper bound of position vectors needed to be checked is

$$C_M^K \cdot M! = \prod_{i=1}^{M} \left(K - i + 1\right). \tag{5}$$

Example: Figs. 1 (a)-(d) illustrate an example of establishing the grid coordinate system for the network consisting of twenty nodes.

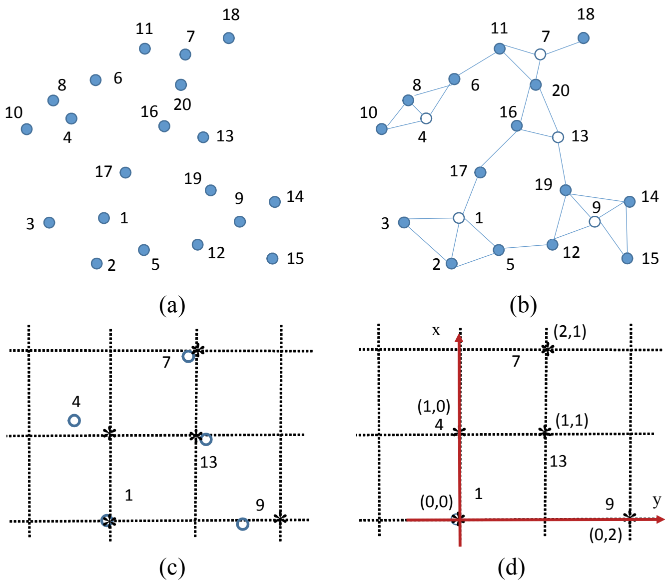


Fig. 1. Grid coordinate system establishment. The small circles ("*o*") denote the real coordinates of the nodes and the stars ("∗") represent the calculated grid coordinates after coordinate transformation; (a) Distribution of nodes; (b) The constructed graph $G(V, E)$ and the set of grid nodes for establishing the grid coordinate system; (c) The local grid coordinate system; (d) The transformed grid coordinate system.

- Fig. 1 (a) shows the distribution of nodes. Initially, each node broadcasts a message containing their ID.
- As long as the distance between two nodes is less than grid width w, $\tilde{d}_{i,j} < w$, there is an edge. Fig. 2 (b) shows the constructed graph $G(V, E)$, where $V = \{01, 02, ..., 20\}$ and $E = \{(01, 02), (01, 03), (01, 05), (01, 017), (02, 03), (02, 05), (04, 06), (04, 08), (04, 10), (05, 12), (06, 08), (06, 11), (07, 11), (07, 18), (07, 20), (08, 10) (09, 12), (09, 14), (09, 15), (09, 19),(11, 20), (12, 19), (13, 16), (13, 19), (13, 20), (14, 15), (14, 19), (16, 17), (16, 20)\}$. Based on the proposed distributed grid node selection algorithm, the set of grid nodes $S' = \{01, 04, 07, 09, 13\}$, which are with the smallest Node ID

compared to their neighbor nodes, are selected for establishing the grid coordinate system.

- Fig. 1 (c) shows the established local grid coordinate system based on Eqs. (3) and (4).

$$\varepsilon\left(X_1, X_4, X_7, X_9, X_{13}\right)=\left(d_{01,04}-\tilde{d}_{01,04}\right)^2+\left(d_{01,07}-\tilde{d}_{01,07}\right)^2+\left(d_{01,09}-\tilde{d}_{01,09}\right)^2$$
$$+\left(d_{01,13}-\tilde{d}_{01,13}\right)^2+\left(d_{04,07}-\tilde{d}_{04,07}\right)^2+\left(d_{04,09}-\tilde{d}_{04,09}\right)^2+\left(d_{04,13}-\tilde{d}_{04,13}\right)^2$$
$$+\left(d_{07,09}-\tilde{d}_{07,09}\right)^2+\left(d_{07,13}-\tilde{d}_{07,13}\right)^2+\left(d_{09,13}-\tilde{d}_{09,13}\right)^2,$$
$$\{X_i^*\}_{i=1}^M=\underset{\{X_i^*\}_{i=1}^M}{\arg\min}\left\{\varepsilon\left(X_1, X_2, \ldots, X_M\right)\right\}.$$

The small circles ("o") denote the real coordinates of the nodes and the stars ("$*$") represent the grid coordinates after coordinate transformation. Note that the coordinate system is relative coordinate system and it can be arbitrary rotated or translated as long as the relative distances between nodes remain unchanged.

- To facilitate the comparison between the real positions of grid nodes and the established grid coordinates, the coordinate system in Fig. 1 (c) is transformed to obtain Fig. 1 (d). The node with lowest ID is set to be the origin (0, 0), the node with the second lowest ID is set to be on positive X-axis with coordinate (x, 0) where $x > 0$. Y-axis is set to be perpendicular to the X-axis. The node with the third lowest ID is used to decide the positive direction of Y -axis. In the example, node 1 is set to be the origin (0, 0). Node 4 is set to be on positive X-axis with coordinate (1, 0). Y-axis is set to be perpendicular to the X-axis. Node 7 is used only to decide the positive direction of Y-axis and therefore its coordinate is (2, 1). ❑

3.2 Error Bound Analysis of the Grid Coordinate System

This subsection mathematically shows that the averaged localization error of the grid relative coordinate system is bounded by $\frac{w}{2}$, where w is the grid width. Here, we assume all the nodes in the network are able to estimate the distances between themselves and their communication neighbors by using ToA or RSSI range measurement. Measurements have shown that if one subtracts the ensemble mean, the time-of-arrival error measurements can be roughly modeled as a zero-mean Gaussian random variable. To simplify the analysis, we assume that the time-of-arrival errors on different links are independent and identically distributed (i.i.d.) random variables. Assume that

$$\tilde{d}_{i,j}=d_{i,j}+x \tag{6}$$

where $\tilde{d}_{i,j}$ is the distance measurement between s_1 and s_2, and x is a random variable. For simplicity, let us consider the localization of $M + 1$ grid nodes, $\{s_i\}_{i=0}^{M}$, which are arranged in one-dimension line. We will show in Section 4 that the result holds in the case of two-dimensional plane. Let x_i and $\tilde{x}_i$ denote the coordinate and the estimated coordinate of node s_i, respectively. Without loss of generality, let us assume that x_0 is the origin of the grid coordinate system, *i.e.*, $x_0 = 0$. The averaged error of the grid coordinate system is defined as

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\left|round\left(\tilde{x}_l, w\right) - x_i\right| \tag{7}$$

where $round(\tilde{x}_i, w)$ denotes rounding $\tilde{x}_i$ to a multiple of increment w. Taking the partial derivative of Eq. (2) with respect to x_i, i = 1, 2, ..., M, we can obtain

$$\begin{pmatrix} M & -1 & \cdots & -1 \\ -1 & M & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & M \end{pmatrix}\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_M \end{pmatrix} = \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_M \end{pmatrix} \tag{8}$$

where $D_k = \sum_{j=0}^{k-1}\tilde{d}_{k,j} - \sum_{i=k+1}^{M}\tilde{d}\_\left(k,i\right)$, k = 1, 2, …, M. The optimal solution to the optimization problem (4) is given by

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_M \end{pmatrix} = \frac{1}{M+1}\begin{pmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 2 \end{pmatrix}\begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_M \end{pmatrix}. \tag{9}$$

By some manipulations, we can obtain

$$D_k = \sum_{j=0}^{k-1}\tilde{d}_{k,j} - \sum_{i=k+1}^{M}\tilde{d}_{k,i} = \left(M+1\right)\tilde{x}_k - \sum_{i=1}^{M}\tilde{x}_i + \left(2k - M\right)\chi. \tag{10}$$

Inserting Eq. (9) into Eq. (8), one obtains

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_M \end{pmatrix} = \frac{1}{M+1}\begin{pmatrix} \left(M+1\right)x_1 + 2\chi \\ \left(M+1\right)x_2 + 4\chi \\ \vdots \\ \left(M+1\right)x_M + 2M\chi \end{pmatrix}. \tag{11}$$

Inserting Eq. (11) into Eq. (7), one obtains

$$\begin{aligned} \Psi &= \frac{1}{M}\sum_{i=1}^{M}\left|round\left(x_i + \frac{2i\cdot\chi}{M+1}, W\right) - X_i\right| \\ &\le \frac{1}{M}\sum_{i=1}^{M}\left(\left|x_i + \frac{2i\cdot\chi}{M+1} - x_i\right| + \frac{w}{2}\right) \\ &= \frac{1}{M}\sum_{i=1}^{M}\left(\left|x_i + \frac{2i\cdot\chi}{M+1} - x_i\right|\right) + \frac{w}{2} \end{aligned}$$

$$= |\chi| + \frac{w}{2}. \tag{12}$$

If the measurement error in the estimation of distances between adjacent nodes is modeled as zero-mean additive white Gaussian noise $\sim n(0,\sigma^2)$, then

$$E[\Psi] \le \sqrt{\frac{2}{\pi}} \cdot \sigma + \frac{w}{2}. \tag{13}$$

❑

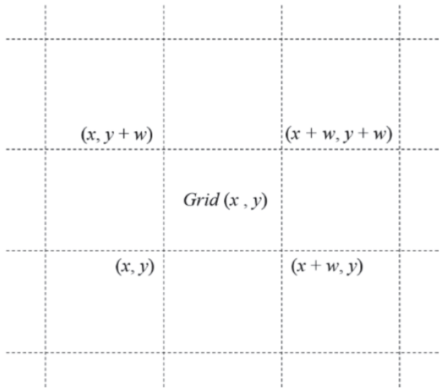


Fig. 2. *Grid*(x, y).

3.3 Localization of the Non-grid Nodes

Rather than computing the accurate coordinates of the non-grid nodes, this paper proposes a distributed scheme, *localization by grid optimization*, whereby a non-grid node can estimate the grid cell containing it, *i.e.*, the grid cell it is located in, by sending a request message to its nearby grid nodes (within the transmission range R) on demand. Let *Grid*(x, y) denotes the grid shown in Fig. 2. Then $(x + \frac{w}{2}, y + \frac{w}{2})$ is the center coordinate of *Grid*(x, y). *Grid*(x, y) is said to be within the $R + \in'$ range of the grid node s_i if $\sqrt{\left(x + \frac{w}{2} - g_{x_i}\right)^2 + \left(w + \frac{w}{2} - g_{y_i}\right)^2} \le R + \in'$, where $\in'$ is a tunable parameter used to tolerate possible measurement errors in estimation of distance. Let G_i be the set of grid nodes that receive the request message from non-grid node s_i. Upon receiving the request, every grid node $s_j \in G_i$ responds a message containing its identity number, the grid coordinate (g_{x_i}, g_{y_i}), and the set H_j of grid coordinates that are within the $R + \in'$ range of grid node s_j. Let $H = \bigcap_{s_j \in G_i}$ be the intersection set of grid coordinates. With the received information from nearby grid nodes, s_i calculates the center coordinates (p, q) of the grid containing s_i by calculating the following formula.

$$(p,\ q) = \arg\min_{(x,\ y) \in H} \{ f(x,\ y) \}$$

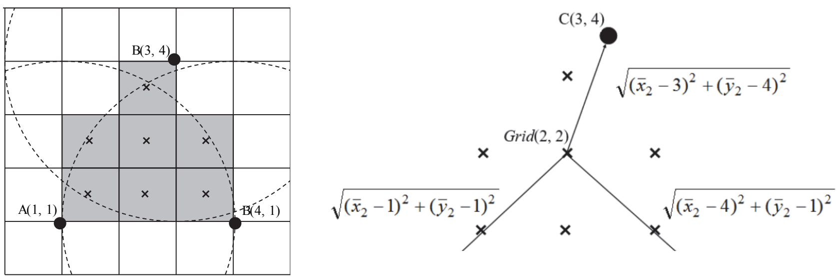
where

$$f(x, y) = \sum_{S_k \in G_i} \left(\sqrt{\left(x + \frac{w}{2} - g_{x_k}\right)^2 + \left(y + \frac{w}{2} - g_{y_k}\right)^2} - \tilde{d}_{i,k} \right)^2 \tag{14}$$

The proposed localization scheme for the non-grid nodes is a digitized triangulation approach. The advantage of this scheme is that every non-grid node can be positioned efficiently because the total number of grids needed to be evaluated in Eq. (13) is bounded by $\left(\frac{(R+\varepsilon')^2 \cdot \pi}{w^2}\right)$. Since R and w are usually constant in most wireless sensor networks, the computation complexity of the localization approach for non-grid nodes is $O(1)$. For example, assume that $R = 40$, $w = 15$ and $\in' = 5$. Each non-grid node only requires evaluating at most 28 grids to determine the best grid in which they reside.

Example: Fig. 3 shows an example to illustrate this scheme more clearly. Assume $G_i = \{A, B, C\}$, $R + \varepsilon' = 3$, $w = 1$, $\tilde{d}_{i,A} = 2$, $\tilde{d}_{i,B} = 3$, and $\tilde{d}_{i,C} = 2$. Among all the possible position candidates, $H = \{(1, 1), (2, 1), (3, 1), (1, 2), (2, 2), (3, 2), (2, 3)\}$, (2, 2) is selected as the final position because of its minimum difference value, which is 0.96.

$$f(2, 2) = \left(\sqrt{\left(\frac{5}{2}-1\right)^2 + \left(\frac{5}{2}-1\right)^2} - 2\right)^2 + \left(\sqrt{\left(\frac{5}{2}-4\right)^2 + \left(\frac{5}{2}-1\right)^2} - 3\right)^2$$
$$+ \left(\sqrt{\left(\frac{5}{2}-3\right)^2 + \left(\frac{5}{2}-4\right)^2} - 2\right)^2 \approx 0.96$$



(a) (b)

Fig. 3. An illustration of positioning non-grid nodes; (a) The intersection set of grid cells; (b) The distances between grid *Grid*(2, 2) and grid nodes $\{A, B, C\}$.

4. SIMULATION

In this section, we evaluate the performance of the proposed localization technique via extensive simulations. First, we study the impact of several parameters, such as grid width, density of grid nodes, node density and power consumption. We also compare our work with the DRLS proposed in [26]. We developed our own simulator to evaluate the proposed algorithms. Measurement error in the estimation of distances between adjacent nodes were modeled as zero-mean additive white Gaussian noise $x \sim n(0, \sigma^2)$ with $\sigma^2 = 1$, which was also adopted in [14-16]. In the following evaluations, the network topologies were generated by randomly placing sensor nodes in a 100m × 100m area. The simulator was executed 100 times to obtain the average results. The estimated error of each sensor node is the distance between the estimated position and actual physical location. We evaluate the localization schemes of grid coordinate system by the localization error (denoted by ψ), which is the average value of estimated error of all grid nodes.

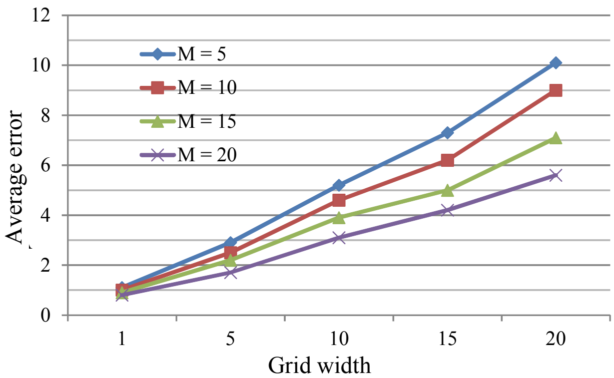


Fig. 4. Average grid coordinate error versus grid width.

4.1 Localization Error of Grid Coordinate System

4.1.1 The impact of grid width and density of grid nodes

In this subsection, we evaluate the impact of grid width (w) and the density of grid nodes (M) on the localization error of grid coordinate system. Extensive simulations have shown that the more distance measurements used to locate nodes, the higher accuracy of the established grid coordinate system, since the errors in the estimation of distances could be smoothed out more effectively [15, 16]. However, since the coordinate of every grid node is restricted by grid coordinate (see Eq. (3)), this would bring additional errors slightly, depending on the grid width and the total number of grid nodes used in the establishment of grid coordinate system. To evaluate the accuracy of the established grid coordinate system, the grid coordinate system is transformed by setting the node with lowest ID to be the origin (0, 0), the node with the second lowest ID to be on positive X-axis with coordinate $(x, 0)$ where $x > 0$, and the node with the third lowest ID to be on positive Y-axis. Fig. 4 shows the average error of the grid coordinate system where the X-axis represents grid width w and the Y-axis denotes the average error ψ. As demonstrated in Fig. 5, ψ increases (approximately) linearly with grid width w and the average error is no more than $\frac{w}{2}$ when $M \leq 20$ and $1 \leq w \leq 20$. (Note that $M \leq \left(\frac{100}{w}\right)^2$ in a 100×100 area.) This confirms the theoretical results in Eq. (13) that the overall localization error of grid coordinate system is upper bounded by $\sqrt{\frac{2}{\pi}} \cdot \sigma + \frac{w}{2}$. This result would become more clearly when more grid nodes are used.

4.1.2 Grid coordinate system

Figs. 6 (a)-(d) show one example of the established grid coordinate system with M = 15. As shown in Fig. 5, the real coordinates are very close to the estimated coordinates when w = 5 and 10. As w increases, the differences between them become slightly obvious. However, the simulation results in Fig. 4 demonstrate that the average error is no more than $\frac{w}{2}$. That is, the maximum localization error of a grid node is within one grid cell. In our simulations, the overall error is far less than this bound in most situations. For example, the localization error is nearly 4.8 units in the case of w = 15. The results also indicate that the precision of grid coordinate system can be adaptively adjusted by the grid width and the density of grid nodes to meet target applications. When a target application requires higher localization accuracy, a smaller grid width and/or more grid nodes can be utilized. Finally, the results also show that the grid coordinate system can be tuned to the global coordinates successfully by coordinate transformation.

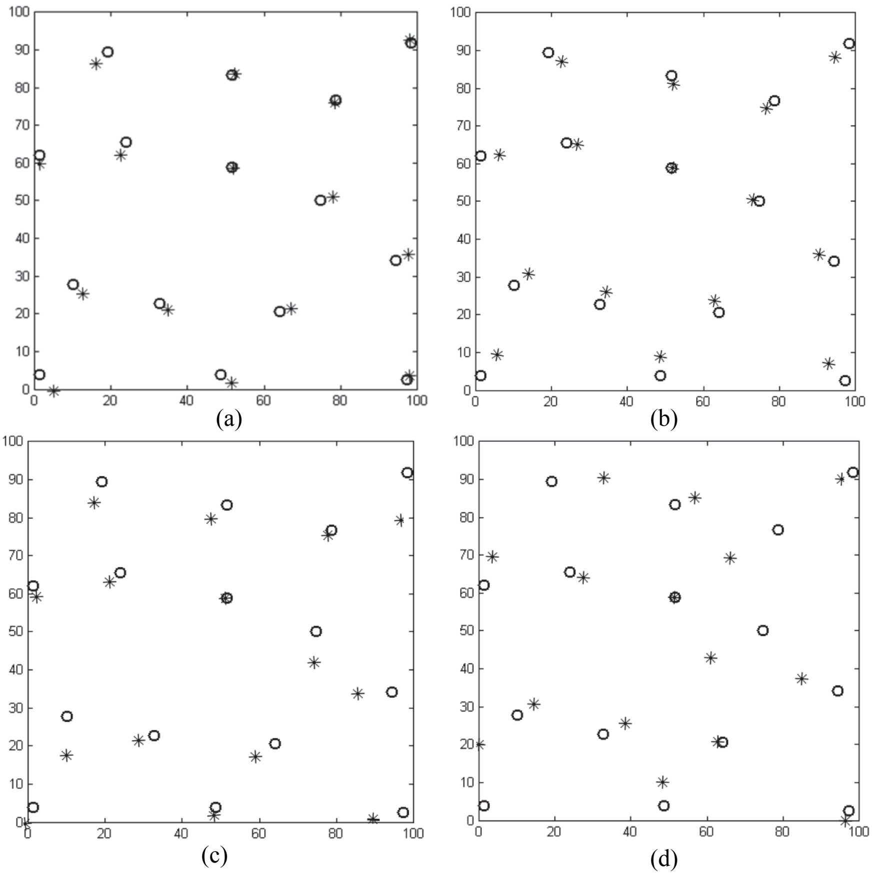


Fig. 5. The established grid coordinate system with M = 15. The small circles (“o”) denote the real coordinates of the nodes and the stars (“*”) represent the calculated grid coordinates after coordinate transformation; (a) ψ = 1.7804; w = 5; (b) ψ = 4.0752, w = 10; (c) ψ = 4.8085, w = 15; (d) ψ = 8.4885, w = 20.

4.2 Localization of Non-grid Nodes

4.2.1 Localization error

We simulate networks that consist of 115 randomly distributed sensor nodes. Fifteen of them are selected to serve as grid nodes, *i.e.*, $M = 15$, and the other nodes serve as non-grid nodes that locate themselves by communicating with the nearby grid nodes to obtain necessary information. Then they determine which grid they belong to using Eq. (14). The communication range of a non-grid node is set to 40 units so that each non-grid node has an average of about five neighboring grid nodes. Since the non-grid nodes locate themselves by the coordinates of grid center, we will use a grid-based error measurement to evaluate the localization error of non-grid nodes instead of the localization error used in grid coordinate system. Let (x_i, y_i) and $(\tilde{x}_i, \tilde{y}_i)$ be the coordinates of the actual physical location of s_i and the coordinates of estimated grid center, respectively. Then, a non-grid node s_i is said to be correctly located if (x_i, y_i) is within the grid cell of $(\tilde{x}_i, \tilde{y}_i)$, *i.e.*, $|x_i, \tilde{x}_i| \leq \frac{w}{2}$ and $|y_i, \tilde{y}_i| \leq \frac{w}{2}$. We use Exact($P_1$) to represent the rate of non-grid grid nodes that locate themselves *correctly*. On the other hand, the localization result of a non-grid node is said to be *partially correct* if (x_i, y_i) falls into one of eight grid cells surrounding the estimated grid $(\tilde{x}_i, \tilde{y}_i)$. We use Partial($P_2$) to represent the rate of non-grid nodes that locate themselves partially correct. Fig. 6 depicts the simulation results, where Total = Exact(P_1) + Partial(P_2). The result shows that Exact(P_1) increases with grid width w. About sixty percent of non-grid nodes locate themselves at correct grid. Based our observations, a non-grid node might be located partially correct if its real position is near grid boundary; however, almost all of these nodes would be located partially correct. The results indicate that the proposed localization scheme is applicable in many applications, such as region-based routing [3-6].

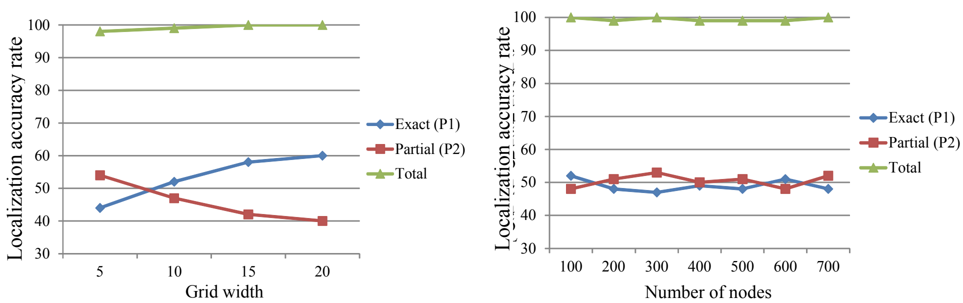


Fig. 6. Localization accuracy of non-grid nodes versus grid width and node density.

4.2.2 Scalability

Next, we demonstrate the scalability of the proposed method in the network consisting N randomly distributed sensor nodes, where N changes from 100 to 700. The grid width w is set to be 10 units, *i.e.*, $w = 10$. As shown in Fig. 6, although there is no obvious trend related to node density, we could also observe that near half nodes are located correctly and the remaining nodes are located partially correct even though N increases from 100 to 700 nodes.

4.2.3 Localization cost

In this subsection, we evaluate the averaged power consumption for locating non-grid nodes. The energy consumption is simulated by the energy model of telosb [23]. The grid width is set to be 10 units. As depicted in Fig. 7, the averaged power consumption per non-grid node would increase with the transmission range. This is because the larger the transmission range, the more neighboring grid nodes will participate in the localization process Eq. (14). In our simulations, each non-grid node would have an average of five neighboring grid nodes when the transmission range is set to be 40 units. As discussed previously (see Fig. 6), the proposed localization scheme could obtain acceptable localization accuracy with negligible computation cost.

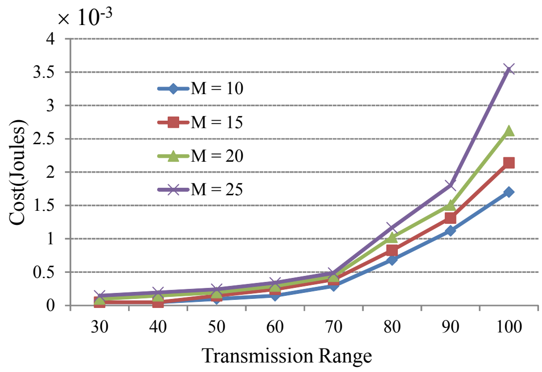


Fig. 7. Energy cost versus transmission range.

4.3 Comparison with DRLS

In this section, we compare our methods with DRLS [26] by evaluating the impact of the number of anchor nodes on the localization error and coverage. The localization error is the difference between real position and the estimated position of the normal node which is not equipped with GPS, and the mean error is calculated as the average localization errors of all normal nodes. The coverage is the ratio of successfully positioning nodes over total nodes. We here define a node is successfully positioned as long as there are sufficient reference nodes in its communication range for trilateral positioning. In this simulation, 115 nodes including normal nodes and anchor nodes are randomly distributed in 100m × 100m area. The communication range is set as 10 m and the grid width is set as 3 and 5.

From Fig. 8, we can see that the performance of DRLS highly depends on the number and distribution of anchor nodes. When the number of anchor nodes is five or less, almost all nodes cannot be successfully positioned due to lack of sufficient reference nodes. As the ratio of anchor nodes increases up to 40%, the coverage increases but still lower than 25% because the distribution of anchors are not well-planned. In contrast with DRLS, as long as there are three anchor nodes and no matter where the anchor nodes are arranged, over 90% nodes are successfully positioned by our method and the mean error is less than 10 which also meets our error bound analysis in Section 3.2. As the ratio of anchors increases to 40%, our method still provides lower errors and higher coverage than DRLS. In general, anchor nodes are nodes equipped with GPS to get their precise locations which are costly when the number of anchor nodes increases. Obviously, compared to DRLS and other previous works, our method improves the localization accuracy and reduces the cost which is the main contribution of this paper.

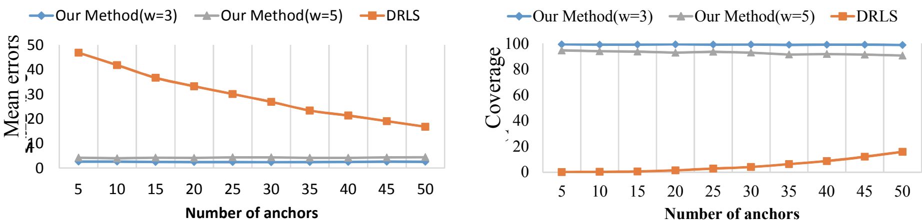


Fig. 8. Mean errors and coverage versus number of anchor nodes.

5. CONCLUSION

In this paper, we propose a scalable grid-based localization technique to balance the tradeoff between computation complexity and precision of localization. The simulation results show that a small number of nodes, *e.g.*, five to fifteen, are usually enough for create the coordinate system with acceptable localization accuracy for many applications. The average error of the grid coordinate system is no more than half the grid width. About sixty percent of non-grid nodes could compute in which grids they are located correctly. Significantly, although some of non-grid nodes (near 40%) locate themselves incorrectly, almost all of them are located correctly in one of eight grids surrounding their real positions. Compared to previous studies, this paper provides an adaptive, self-configurable and scalable localization scheme which improves the localization accuracy and to lower the cost of previous works.

REFERENCES

1. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” *Computer Networks*, Vol. 38, 2002, pp. 393-422.
2. A. Caruso, S. Chessa, S. De, and A. Urpi, “GPS free coordinate assignment and routing in wireless networks,” in *Proceedings of IEEE INFOCOM*, Vol. 1, 2005, pp. 150-160.
3. X. Li, J. Yang, A. Nayak, and I. Stojmenovic, “Localized geographic routing to a mobile sink with guaranteed delivery in sensor networks,” *IEEE Journal on Selected Areas in Communications*, Vol. 30, 2012, pp. 1719-1729.
4. A. Awad, R. German, and F. Dressler, “Exploiting virtual coordinates for improved routing performance in sensor networks,” *IEEE Transactions on Mobile Computing*, Vol. 10, 2011, pp. 1214-1226.
5. B. Li, W. Wang, Q. Yin, H. Li, and H.-M. Wang, “Energy-efficient cooperative geographic routing in wireless sensor networks,” in *Proceedings of IEEE International*

Conference on Communications, 2012, pp. 152-156.
6. G. Tan and A.-M. Kermarrec, "Greedy geographic routing in large-scale sensor networks: a minimum network decomposition approach," *IEEE/ACM Transactions on Networking*, Vol. 20, 2012, pp. 864-877.
7. D. Nicolescu and B. Nath, "DV-based positioning in ad hoc networks," *Journal of Telecommunication Systems*, Vol. 22, 2003, pp. 267-280.
8. S. Zhu and Z. Ding, "Distributed cooperative localization of wireless sensor networks with convex hull constraint," *IEEE Transactions on Wireless Communications*, Vol. 10, 2011, pp. 2150-2161.
9. J. Gu, S. Chen, and T. Sun, "Localization with incompletely paired data in complex wireless sensor network," *IEEE Transactions on Wireless Communications*, Vol. 10, 2011, pp. 2841-2849.
10. R. Sugihara and R. K. Gupta, "Sensor localization with deterministic accuracy guarantee," *IEEE INFOCOM*, 2011, pp. 1772-1780.
11. M. Leng, W.-P. Tay, and T. Q. S. Quek, "Cooperative and distributed localization for wireless sensor networks in multipath environments," in *Proceedings of IEEE International Conference on Acoustics*, *Speech and Signal Processing*, 2012, pp. 3125-3128.
12. B. Zhang, X. Cheng, N. Zhang, Y. Cui, Y. Li, and Q. Liang, "Sparse target counting and localization in sensor networks based on compressive sensing," in *Proceedings of IEEE INFOCOM*, 2011, pp. 2255-2263.
13. K.-S. Low, H. A. Nguyen, and H. Guo, "A particle swarm optimization approach for the localization of a wireless sensor network," in *Proceedings of IEEE International Symposium on Industrial Electronics*, 2008, pp. 1820-1825.
14. H. Wu, C. Wang, and N. F. Tzeng, "Novel self-configurable positioning technique for multi-hop wireless networks," *IEEE/ACM Transactions on Networking*, Vol. 13, 2005, pp. 609-621.
15. Q. Shi, S. Kyperountas, N. S. Correal, and F. Niu, "Performance analysis of relative location estimation for multihop wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, Vol. 23, 2005, pp. 830-838.
16. K. Whitehouse, A. Woo, C. Karlof, F. Jiang, and D. Culler, "The effects of ranging noise on multi-hop localization: an empirical study," in *Proceedings of International Conference on Information Processing in Sensor Networks*, 2005, pp. 73-80.
17. M. Luby, "A simple parallel algorithm for the maximal independent set problem," *SIAM Journal on Computing*, Vol. 15, 1986, pp. 1036-1055.
18. R. Peng and M. L. Sichitiu, "Robust, probabilistic, constraint-based localization for wireless sensor networks," in *Proceedings of IEEE Annual Conference on Sensor and Ad Hoc Communications and Networks*, 2005, pp. 541-550.
19. D. K. Goldenberg, P. Bihler, M. Cao, and J. Fang, "Localization in sparse networks using sweeps," in *Proceedings of the 25th ACM Annual International Conference on Mobile Computing and Networking*, 2006, pp. 110-121.
20. S.-P. Kuo, H.-J. Kuo, and Y.-C. Tseng, "The beacon movement detection problem in wireless sensor networks for localization applications," *IEEE Transactions on Mobile Computing*, Vol. 8, 2009, pp. 1326-1338.

21. J.-P. Sheu, W.-K. Hu, and J.-C. Lin, "Distributed localization scheme for mobile sensor networks," *IEEE Transactions on Mobile Computing*, Vol. 9, 2010, pp. 516-526.
22. N. B. Priyantha, H. Balakrishnan, E. D. Demaine, and S. Teller, "Mobile-assisted localization in wireless sensor networks," in *Proceedings of IEEE INFOCOM*, Vol. 1, 2005, pp. 172-183.
23. J. Polastre, R. Szewczyk, and D. Culler, "Telos: enabling ultra-low power wireless research," *IEEE/ACM Information Processing in Sensor Networks*, 2005, pp. 364-369.
24. Y. Liu, L. M. Ni, and M. Li, "A geography-free routing protocol for wireless sensor networks," in *Proceedings of IEEE International Conference on High Performance Switching and Routing*, 2005, pp. 351-355.
25. C. Lin, B. Liu, H. Yang, C. Kao, and M. Tsai, "Virtual-coordinate-based delivery-guaranteed routing protocol in wireless sensor networks," *IEEE/ACM Transactions on Networking*, Vol. 17, 2009, pp. 1228-1241.
26. J.-P. Sheu, P.-C. Chen, and C.-S. Hsu, "A distributed localization scheme for wireless sensor networks with improved grid-scan and vector-based refinement," *IEEE Transactions on Mobile Computing*, Vol. 7, 2008, pp. 1110-1123.
27. H. Chen, H. Wu, and N.-F. Tzeng, "Grid-based approach for working node selection in wireless sensor networks," in *Proceedings of IEEE International Conference on Communications*, Vol. 6, 2004, pp. 3673-3678.

Pei-Hsuan Tsai (蔡佩璇) received her B.S. degree in Computer Science and minor B.S. degree in Materials Science and Engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2003. She received her M.Eng degree in Computer Science from Cornell University, Ithaca, NY, in 2004, and her Ph.D. in Computer Science from National Tsing Hua University, Hsinchu, Taiwan, in 2010. In August 2011, she joined the Institute of Manufacturing Information and Systems at National Chen Kung University as an Assistant Professor. Before joining NCKU, she was a Postdoctoral Fellow at the Institute of Information Science, Academia Sinica, Taipei.

Her research interests have been in the areas of sensor network, localization, path planning, task scheduling, data fusion and decision analysis. Her recent research focuses on technologies for building medical information system and disaster response system.