A Distributed Localization Scheme for Wireless Sensor Networks

Prasan Kumar Sahoo  
Dept. of Information Management  
Vanung University  
Chungli, 32026, Taiwan  
pksahoo@msa.vnu.edu.tw

I-Shyan Hwang  
Dept. of Computer Engineering and Science  
Yuan-Ze University  
Chungli, 32026, Taiwan  
iswang@saturn.yzu.edu.tw

Shi-Yao Lin  
Dept. of Computer Engineering and Science  
Yuan-Ze University  
Chungli, 32026, Taiwan  
lin12832@gmail.com

ABSTRACT
Localization of the nodes in Wireless sensor networks (WSNs) is an important research issue, since it can enhance the efficiency in computation and minimize the power consumption of the nodes. In this paper, a novel localization algorithm is proposed to estimate the location information of the normal nodes with help of few beacon nodes and angle information of the anchor nodes. Our localization scheme can use at most three beacon nodes to find location information of any normal node in a distributed manner. Besides, we give the theoretical basis for determining the localization error using probability distribution function. Our performance analysis shows that there is a tradeoff between deployed number of beacon nodes and localization error and average localization time of the network can be increased with deployed number of normal nodes.

Categories and Subject Descriptors
C.2 [Computer Systems Organization]: Computer Communication Networks

General Terms
Theory

Keywords
Distributed, Sensor Networks, Localization

1. INTRODUCTION
In recent years, with rapid advances in Micro Electro Mechanical Systems (MEMS) technology, Wireless sensor networks (WSNs) have received extensive interest lately. It is getting popular due to its low cost and small size and its applications in military and civilian surveillance. However, wireless sensor networks have a few inherent limitations, e.g., limited hardware, limited transmission range, and large scale network system and the traditional protocol or mechanism cannot use in WSNs. Hence, several issues need to research in WSNs to construct an efficient and robust network. For example, sensor nodes have limited computation capability and limited power supply and therefore low complexity algorithms and power saving schemes, respectively should be designed.

In wireless sensor networks, location of nodes plays an important role in most applications. When sensors are deployed over a network, they only have connectivity information with neighbors but do not know their location information. In some situations, the problem can have easy to solve if location information of the node is available, i.e. when nodes have location information, routing path can be observed easily, and coverage hole can be easily detected. Knowing relative location of sensors allows the location-based addressing and routing protocols, which can improve network robustness and energy-efficiency effectively. Recent research results show that nodes with location information lead to increase performance of applications and reduce power consumption. In addition, more accurate location information leads the more accurate of result that application needs. In summary, localization is an essential part of WSNs.

Normally, sensors are intended to be low-cost disposable devices, and currently developed solutions such as global position system (GPS) [1] are inadequate for the hardware and power-limited sensors. Traditional localization techniques are not well suited for these requirements. Including a global positioning system (GPS) receiver on each device is cost and energy prohibitive for many applications, not sufficiently robust to jamming for military applications, and limited to outdoor applications. Local positioning systems (LPS) [2] rely on high-capability base stations being deployed in each coverage area, an expensive burden for most low-configuration wireless sensor networks. Hence, automatic localization of the sensors in wireless networks is a key enabling technology. The overwhelming reason is that a sensor’s location must be known for its data to be meaningful. As an additional motivation, sensor location information can be extremely useful for scalable, and geographic routing algorithms.

In this paper, a novel localization scheme is proposed to cal-
calculate the relative location of the nodes distributively with help of anchor and beacon nodes. The rest of the paper is organized as follows. An overview of the related work is presented in Section 2 and our proposed localization scheme is described in Section 3. The performance analysis of our algorithm is made in Section 4 of the paper. Concluding remarks are made in Section 5.

2. RELATED WORK

Localization in wireless sensor networks is different from traditional wireless communication technology. There has been an increasing interest in the localization technique for WSNs in recent years and many localization algorithms have been proposed [3][4][5]. Constraint on limited hardware supports and power supply, sensor nodes can only find its approximate location information. In order to find node's location efficiently and simply, various localization algorithms have been proposed. The localization algorithms can further be divided into Range-based and Range free localization schemes. The Range-based localization scheme uses measurements of distance or angle to estimate node's location. According to signal propagation and receive time, two kinds of technology are mentioned to obtain the distance. They are: TOA (Time of arrival) [6], TODA (Time of difference of arrival) [7]. TOA method is used to obtain the range between sender and receiver nodes by signal arrival time. TODA technique is based on the difference in time between two different signals arrival time and is widely proposed as a necessary measurement method in localization solution for WSNs. The algorithms proposed in [8] and [9] are self-organized methods to establish the relative coordinate system on every known nodes through the TODA.

Angle of arrival (AOA) technique [10] is another range-based localization algorithm. In this algorithm, normal nodes have ability to detect the angle to neighbor nodes by directional antenna or smart antenna. By angle information, we also can calculate the node's position. However, these three methods require additional equipments and hardware supports, which may incur additional cost and energy consumption. Hence, these protocols seem less suitable for the low-power WSNs. In Globe Position System (GPS), few beacon nodes obtain their absolute location information by GPS and other unknown nodes estimate their location information by receiving the beacon packets from the beacon nodes. In [11], authors propose a localization scheme called approximate point in triangular test (APIT) algorithm. In APIT, each beacon node first broadcasts the beacon packet to neighbor nodes, which is later flooded into the whole network. Then each unknown node determines if it is within a particular triangle formed by a set of beacon nodes. Finally, unknown node estimates its location by the center of gravity of the overlapped area. Although location information of the unknown nodes can be obtained by this algorithm, still there are some problems exist in it. First, the accuracy relies on heavily percentage of beacon nodes, and communication cost is high as each node needs to listen many times to different beacon packets. Besides, the complexity of computations is high when the unknown node estimates the overlapped area.

A range-free localization scheme called DV (distance vector) hop is proposed in [12][13]. It uses topological information and number of hops to alternative the real distance. In the beginning the beacon node floods the packet with hop count and node ID to the rest of the network. Unknown nodes compute the average hop size of their nearest beacon node, translate the number of hops into real distance and estimate their position. However, some drawbacks exist in DV-hop algorithm, since localization accuracy depends on the node density. Besides, irregular deployment will cause the inaccuracy of average hop size and communication cost are still high. In order to improve the accuracy of location information, a distributed location estimation scheme (DLS) has been proposed in [14]. In this algorithm, each beacon node exchanges the node ID and location information to all nodes of the network. The unknown node calculates its own estimated rectangle (ER) and regards the center of ER is his location. In [5], the authors propose a distributed range-free algorithm, called Concentric Anchor Beacon (CAB) localization algorithm. In CAB, each beacon node emits several beacon packets with different power levels and each node maintains a table that includes the ID, location, transmit power level and constraint region of the beacon node. Each normal node determines the particular ring or circle it belongs to within range of different anchors. From the intersection points of different rings, the average of those intersection points is estimated as the location of a node.

Although CAB uses few beacon nodes for localization, but it still has some drawbacks. Firstly, it is not a good method for beacon nodes to transmit packets with different power level. Moreover, averaging the intersection point is not accurate result. If some nodes have the same intersection points, then the algorithm will give same location information to those nodes. From the discussion of those two kinds of localization schemes, it is clear that each of them have unique properties. In range-based scheme, it can provide more accurate location estimation, but need additional equipments. In range free scheme, low cost location system can be built, but estimated location is not accurate enough than range based scheme. In our work, we propose a range free scheme, which is cost effective. In order to get accuracy of the localization, we propose analytical methods to correct the errors and to get more realistic position of a node.

3. DISTRIBUTED LOCALIZATION (DIL) ALGORITHM

Let us consider a rectangular outdoor monitoring region to find location of the nodes. In our localization algorithms, nodes are classified as Normal, Beacon and Anchor nodes to find location of the normal nodes. Normal nodes and beacon nodes are deployed randomly on the monitoring region and normal nodes have no location information. However, beacon nodes have location information with higher capacity of computation and more energy resource. Anchor nodes have larger communication range and are deployed manually. In our protocol, it is assumed that anchor nodes provide angle information to each normal nodes of the network and percentage of anchor nodes is less than the beacon nodes. As shown in Figure 1, the whole network is divided into several clusters, and only one anchor node is deployed in each cluster. In our localization process, it is assumed that there must be at least one beacon node around the normal nodes in order to ensure that normal nodes get enough information to calculate their position. Besides, at most three beacon
nodes in each cluster are used to find position of a normal node. The deployment strategy of those three types (anchor, beacon and normal) of nodes are described in subsection

3.1 Node Deployment Strategy

In our localization system, normal nodes should receive enough information from the beacon and anchor nodes to calculate their location information correctly. In order to ensure every normal node get enough information, we need to design the node deployment strategy to optimize our localization algorithm. As per our assumption, normal nodes get angle information from the anchor nodes. Hence, first the anchor nodes are deployed manually to make sure that the entire monitoring region is fully covered. It is assumed that the size of the monitoring region is mbyn, and communication range of the anchor node is Rc. The deployment of the anchor nodes is made based on the equation (1) and (2).

\[
x = \sqrt{2}Rc + (2 \times \sqrt{2}Rc) \times a
\]

and

\[
y = \sqrt{2}Rc + (2 \times \sqrt{2}Rc) \times b
\]

where, \(a\) ranges from 0 to \([\frac{m}{2Rc}]\), and \(b\) ranges from 0 to \([\frac{n}{2Rc}]\). After deployment of the anchor nodes, a small percentage of beacon nodes that is more than the number of the anchor nodes are deployed on the monitoring regions, randomly. Then, large percentage of normal nodes that are more than the number of the beacon nodes are deployed randomly.

3.2 Localization Algorithm

In this section, we describe our localization algorithm. Prior to this, we introduce the distance measurement mechanism of the normal nodes from the received signal strength indicator (RSSI) value of the beacon nodes. We propose algorithm to compute coordinate of each node based on the angle information from the anchor nodes and distance information from the beacon nodes, as described below.

3.2.1 Distance Measurement

The received signal strength indicator (RSSI) is one type of distance estimation technology to obtain the distance between transmitter and receiver [15]-[16]. This measurement technology is based on a standard feature found in most wireless devices and is attractive as they do not need any additional hardware support. When the transmitter sends packet to receiver, receiver obtains the RSSI value as the inverse square of the distance. Most sensor network research assumes that the propagation of signal is an over idealization, e.g. free space model. In fact, the fading and shadowing effects must be considered because of the noise and obstacle. Experimental results [17] show that many well-designed protocols in WSNs fail in a realistic wireless environment. Typically, the mean RSS decays between transmitter and receiver (T-R) can be predicted by some radio propagation model. The log normal shadowing model is a most commonly used propagation model that considers the shadowing effect, whether in outdoor or indoor environment. This model indicates that the average received signal strength decreases logarithmically with distance. In general, the average path loss for an arbitrary T-R separation can be expressed as given in equation (3).

\[
P_r(d) = P_t(d_0) - 10n\log\left(\frac{d}{d_0}\right) + x_
\]

where \(n\) is the path loss exponent, which depends on the specific propagation environment. \(d\) is the distance between T-R, and \(P_r(d)\) represents the received signal strength (RSS). \(P_t(d_0)\) represents the transmission power at reference distance \((d_0)\). The term \(X_n\) is a random variable which accounts for the random variation of the path loss, and is supposed to be Gaussian distribution with zero mean random variable (in dB) with standard deviation \(\sigma\) (also in dB). Based on equation (3), we can obtain the distance \(d\) from equation (4).

\[
d = d_0 \times 10^{\frac{P_t(d_0) - P_r(d) - X_n}{10n}}
\]

3.2.2 Coordinate Computation

Upon measuring the distance between the beacon and normal node based on the RSSI value and as described in subsection 3.2.1, we propose algorithm to estimate the coordinate of a normal node. According to our assumptions, at least one beacon node sends information to the normal node for localization and an anchor node sends angle information to the normal node. Upon receiving beacon packet from one beacon node and angle information from the anchor node, a normal node waits for a predefined timeout \(T_n\) to receive RSSI value from other beacon nodes of its cluster.

Let, \((x, y)\) be the coordinate of the normal node, \((x_1, y_1)\) be the location of beacon node \(B_1\), and \((x_n, y_n)\) be the location of anchor node. Distance between the beacon and normal node is \(d_1\), which is estimated as described in section 3.2.1. Let, the angle between the anchor node’s x-axis and the line joining the normal and anchor node be \(\theta\). Based on these information, we can obtain two equations. We consider the linear equation that passes through the anchor and normal node as shown in Figure 2 and given in equation (5).

\[
y = x \tan \theta + k
\]

where \(k\) is a constant, which is obtained by substituting
location of the normal node.

\[ k = y_n - x_n \tan \theta \]  \hspace{1cm} (6)

Considering the boundary of the sensing range of a beacon node as equation of a circle, we get equation (7).

\[ (x - x_1)^2 + (y - y_1)^2 = d_1^2 \]  \hspace{1cm} (7)

Substituting equation (5) into (7) and upon simplification we obtain equation (8).

\[ (1 + \tan^2 \theta)x^2 - (2x_1 + 2y_1\tan \theta - 2ktan \theta)x + R = 0 \]  \hspace{1cm} (8)

where \( R \) is

\[ R = x_1^2 + k^2 - 2ky_1 + y_1^2 - d_1^2 \]  \hspace{1cm} (9)

Hence, location of the normal node \((x, y)\) can be estimated from equation (8) as given in equation (9), which is obvious.

\[ x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \]  \hspace{1cm} (10)

where \( a \) and \( b \) are coefficient of \( x^2 \) and \( x \), respectively and \( R \) represents a constant term. Then, we substitute equation (10) in equation (5) to get the \( y \) coordinate. However, it could be possible that a normal node may receive beacon packets from two or three beacon nodes. It is to be noted that each cluster can have at most three beacon nodes as per our assumptions. As described previously, first a normal node listens to the network and checks the arrival of the beacon packets. Normal node continues to wait for \( T_n \) units and maintains a coordinate table as shown in Table 1 to record the beacon packet’s information. Each normal node maintains the coordinate table with four fields. They are the ID of the beacon node, location of the beacon node and RSSI value of the beacon node from which beacon packet is received. Besides, the last field records possible estimated location (P-Loc) information of the normal node. Once the waiting time expires, it starts computing its location from all of the received data.

For example, suppose a normal node receives beacon packet from two different beacon nodes \( B_1 \) and \( B_2 \). As shown in Figure 3(a), from the sensing range of \( B_1 \), the line joining the normal and anchor node can have two possible coordinates \( P_1 \) and \( P_1' \). Similarly, from the sensing range of \( B_2 \), another two possible coordinates \( P_2 \) and \( P_2' \) can be obtained. Then, the normal node compares the distance between each combination of points i.e., \( P_1 \) with \( P_2 \) or \( P_2' \) with \( P_1' \) or any other pairs. Finally it chooses the point having the minimum distance or very negligible distance. As shown in Figure 3(a), obviously points \( P_1' \) and \( P_2' \) are selected as the most possible location of the nodes. As shown in Figure 3(b), if more than two beacon packets are received from three different beacon nodes, normal node continues to update the coordinate table and use the same procedure to compute the possible coordinates \( P_3 \) and \( P_3' \) and determines the correct coordinate. Since, there may be slight differences between the final coordinates, the error estimation and correction methods as described in subsection 3.3. can be used to find the most accurate location of the normal node.

The detail procedure of executing the localization algorithm

![Figure 2: Location computation of normal node with help of one beacon node.](image)

![Table 1: Coordinate table](image)

<table>
<thead>
<tr>
<th>BN-ID</th>
<th>BN-loc.</th>
<th>RSSI</th>
<th>P-loc</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_1</td>
<td>(X_{B_1}, Y_{B_1})</td>
<td>RSSI_{B_1}</td>
<td>P_3</td>
</tr>
<tr>
<td>B_2</td>
<td>(X_{B_2}, Y_{B_2})</td>
<td>RSSI_{B_2}</td>
<td>P_2</td>
</tr>
<tr>
<td>B_3</td>
<td>(X_{B_3}, Y_{B_3})</td>
<td>RSSI_{B_3}</td>
<td>P_3'</td>
</tr>
<tr>
<td>B_4</td>
<td>(X_{B_4}, Y_{B_4})</td>
<td>RSSI_{B_4}</td>
<td>P_4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

![Table 2: Distributed Localization (DIL) Algorithm](image)

**Initial:**
- Initialize: Waiting time \( T_n \) for each normal node;
- Initialize: All fields of coordinate table = \{ φ \};

**Do**

- Start: Node deployment strategy;

**For each Anchor nodes:**
  - Check: Neighbors of normal nodes;
  - Measure: Angle information for each neighbor of normal nodes;
  - Transmit: Angle information to each normal nodes;

**For each Beacon nodes:**
  - Broadcast the beacon packet;

**For each Normal nodes:**
  - Setup: Waiting time \( T_n \);
  - While \( T_n \) is not expired
    - do
      - Listen the network;
      - If Any beacon packet is arriving
        - Translate: RSSI into Distance;
        - Computation: Update the coordinate table;
      - End If
      - Calculate: Final result from all entries of the table;
    - Output: Normal node’s location;
  - End

End
of a normal node taking one or maximum three beacon nodes, is given in Table 2.

### 3.3 Localization Error Estimation

It is to be noted that we propose the distributed localization algorithm taking three different types of nodes. We use location information of at least one or at most three beacon nodes to calculate the location of normal nodes. The anchor nodes do not provide location information neither to beacon nor to normal nodes. In our algorithm, they can provide angle information only to the normal nodes. Since, we consider at most three beacon nodes to calculate location of the normal nodes, it could be possible that a normal node may calculate three different locations from the RSSI values received from three different beacon nodes. Hence, we propose here a probabilistic method for improving the location accuracy of the normal node. If only one beacon node is used to calculate the location of the normal node, it is obvious that only one pair of coordinate is estimated as the location of the normal node. However, presence of more beacon nodes can enhance the accuracy of the localization, of course with increased cost. Hence, we propose our system with two or three beacon nodes for error analysis as follows.

Consider three beacon nodes A, B and C are located at different location but within communication range of a normal node. Let, $S_A$, $S_B$ and $S_C$ be the received signal strength (RSS) by a normal node from those beacon nodes A, B and C, respectively. Considering beacon node A, the case of error determination with respect to nodes B and C is

$$P_{A\rightarrow B} = P(f_A(S_A) < f_B(S_A))$$

and

$$P_{A\rightarrow C} = P(f_A(S_A) < f_C(S_A))$$

Denote $\mu_A$ and $\mu_B$ as the expected values of $P_A$ and $P_B$, respectively. If $\mu_A$ larger than $\mu_B$, the error identification will be from $S(f_A = f_B)$ to $\infty$, as shown in Figure 4.

The probability of error determination from A to B when $\mu_A$ is smaller than $\mu_B$ is

$$P_{A\rightarrow error} = 1 - \int_{S(f_A = f_B)} f_A(S) \cdot dS$$

If we consider three beacon nodes A, B and C at the same time, the probability of error determination can be made as a combination among any two nodes out of those three beacon nodes, which can be similar to the above case. If $\mu_A$ is the largest (or smallest), i.e. if we assume that $\mu_A > \mu_B > \mu_C$, then error determination at A could be estimated as given in equation (14).

$$P_{A\rightarrow error} = 1 - \int_{S(f_A = f_B) \cup S(f_A = f_C)} f_A(S) \cdot dS$$

In another case, if $\mu_A > \mu_B$ but $\mu_C > \mu_A$ (and vice versa), then correct determination of localization could be in between $S(f_A = f_B)$ and $S(f_A = f_C)$, as shown in Figure 5.

### 4. PERFORMANCE ANALYSIS

In this section, we analyze performance of our distributed localization algorithms through simulation. The detail description of the simulation setups and results are given as follows.
4.1 Simulation Setups

We have simulated our algorithm using ns 2.29. An outdoor environment with size of the monitoring region $200 \times 200 \ m^2$ is setup in our simulation. The number of deployed nodes over the said monitoring region varies from 250 to 400 nodes including normal, beacon and anchor nodes. The ratio of nodes with known position to nodes with unknown position also varies between 80% to 90%. Communication range of all normal nodes are fixed to 20 meters. The value of path loss exponent is set as 2. Each beacon node transmits beacon packet in an interval of 2 ms. The initial energy resource of each sensor node is considered as 5 joules, which is decreased by 0.3 joules in each transmission. IEEE 802.15.4 medium access mechanism and AODV routing protocol are considered in our simulation.

4.2 Simulation Results

In our simulation, we find out the average estimated error for different situations, which is defined as the difference between the estimated coordinate and real coordinate. As shown in Figure 6, the average estimated localization error for different number of beacon nodes with fixed number of total nodes ($N$) are analyzed. In this simulation, the number of anchor nodes are also fixed, though different lines are obtained for different number of $N$. From this figure, it is observed that the estimated localization error decreases if number of beacon nodes increases. Besides, if more number of nodes $N$ are deployed to the monitoring region, the estimated error also decreases. It is to be noted that the average estimated error is more than 9m when the number of beacon nodes is 35. Hence, in order to get more localization accuracy of the network, deployment of more beacon nodes is essential.

In equation (3), we use the path loss shadowing model to be our propagation model. In this model, $X_\sigma$ is a random variable with standard deviation $\sigma$, which affects the RSSI value and thereby causes error in the estimated localization. As shown in Figure 7, we simulated the percentage of deployed normal nodes with different standard deviation ($\sigma$) to study the average estimated localization error. The number of anchor and beacon nodes in this experiment is fixed. It is noticed that the average estimated localization error is more for large value of the standard deviation. It is reasonable, as the large value of standard deviation means the degree of probability distribution is large, and therefore the average error is increased. Besides, the estimated localization error increases, if percentage of deployed normal nodes is increased.

Figure 8 indicates how the average estimated localization error is affected for different percentage of deployed normal nodes. This experiment is carried out for different communication range of the beacon nodes with fixed number of anchor nodes equals to 9. From this figure, it is found that the average estimated error is reduced, if communication range of the beacon nodes is increased. This situation happens, since most of the normal nodes can receive enough beacon packets to calculate their location and thereby reducing the localization error. However, if percentage of normal node increases, average estimated localization error also increases, which is compatible with the results given in Figure 7.

In Figure 9, we show the average localization time for different number of beacon nodes. It is observed that the average localization time decreases if number of beacon nodes increases. This is expected, as more beacon nodes means more packets are transmitted to calculate the location of the normal nodes. However, if the number of beacon nodes increases, the average localization time also increases, which is due to the increased energy consumption of the beacon nodes.
In Figure 9, the localization time for different number of network size $N$ with different number of beacon nodes is shown. Here, the localization time increases with increase in number of beacon nodes. It is due to when number of beacon nodes increases, a normal node waits for $T_n$ units and therefore the communication time is also increased. From Figure 9, it is interesting to note that the variation in localization time is very less although the network size changes. The analysis of average residual power for different communication range with different percentage of normal nodes is presented in Figure 10. In this experiment, first we measure the residual power of the beacon, anchor and normal nodes and then average their power. As shown in Figure 10, we observe that the average energy consumption is increased, if number of normal nodes is increased. Besides, the residual power decreases if communication range of the normal nodes increases. This is because of more power consumption due to higher communication range.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a novel distributed localization algorithm using at most three beacon nodes. The advantage of our algorithm is that it can work even if only one beacon node provides location information to a normal node. Besides, we propose the error correction methods using probability distribution that gives a solid theoretical basis to verify the localization is correct or not. Our simulation results also satisfy the conditions of our algorithms. We feel that our algorithm can calculate the location of nodes with most simplest ways and with less time complexity, which is quite suitable for the memory and energy constraint sensors. In our future work, we will continue more simulation and analysis and compare with the standard localization algorithms.

6. REFERENCES


