

Sea Depth Measurement with Restricted Floating Sensors

Zheng Yang, Mo Li, and Yunhao Liu
Hong Kong University of Science and Technology
{yangzh, limo, liu}@cse.ust.hk

Abstract

Sea depth monitoring is a critical task to ensure the safe operation of harbors. Traditional schemes largely rely on labor-intensive work and expensive hardware. This study explores the possibility of deploying networked sensors on the surface of sea, measuring and reporting sea depth of given areas. We propose a Restricted Floating Sensors (RFS) model, in which sensor nodes are anchored to the sea bottom, floating within a restricted area. Distinguished from traditional stationary or mobile sensor networks, the RFS network consists of sensor nodes with restricted mobility. We construct the network model and elaborate the corresponding localization problem. We show that by locating such RFS sensors, the sea depth can be estimated without the help of any extra ranging devices. A prototype system with 25 Telos sensor nodes is deployed to validate this design. We also examine the efficiency and scalability of this design through large-scale simulations.

1. Introduction

We conducted a field study in H. H. Harbor, which is currently the second largest harbor for coal transportation in China. It has experienced rapid development over the past 5 years, and its coal transporting capability has increased from 1.6 million tons per year in 2002 to 6.7 million tons per year in 2006. However, this harbor currently suffers from the increasingly severe problem of silt deposition along its sea route. H. H. Harbor has a sea route that is 19 nautical miles long and 800m wide at the entrance, including an inner route and an outer route. The sea route is designed to have a water depth of 13.5m to allow for the passage of ships that weigh over 50 thousand tons. Since the sea route has been in operation, it has always been threatened by the movement of silt from the shallow sea area within 14 nautical miles outside the route entrance. In the event that the sea route is silted up, ships of large tonnages must wait for entering the harbor to prevent grounding, and ships of small tonnages need

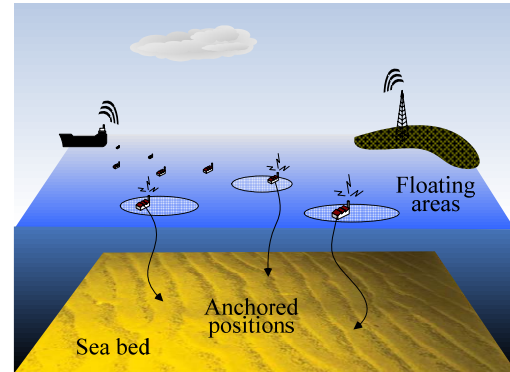


Figure 1: Restricted floating sensors on the sea

be piloted into the harbor. Monitoring the extent of siltation reliably is critical in order to ensure the safe operation of H. H. Harbor.

The uncertainty and the high instant intensity of the siltation make monitoring the extent of siltation extremely expensive and difficult. The amount of siltation in H. H. Harbor is affected by many factors, among which tide and wind blow are the most dominating. While the tides produce a periodic influence on the movement of silt, the highly variable nature of wind brings more incidental and intensive effects. For example, records show that strong winds with wind forces of 9 to 10 on the Beaufort scale hit H. H. Harbor from Oct. 10th to Oct. 13th in 2003. The storm surge brought 970,000m³ of silt to the sea route, which suddenly decreased the water depth from 9.5m to 5.7m and blocked most of the ships weighing more than 35 thousand tons. The harbor administration hired three boats equipped with active sonars to cruise the 380km² shallow sea area around the harbor for several days. Monitoring sea depth costs this harbor more than 18 million US dollars per year.

In this work, we explore the possibility of deploying networked sensors on the sea surface for sea depth measurement. Different from deployments on ground, sensor nodes, in this scenario, will generally not be stationary at their original deployed places, but float by many factors, e.g. ocean current, wind blow, and tide

etc. Therefore, we anchor the sensor nodes to the sea bottom by ropes to restrict their floating movements. Otherwise, they may float out of our interested monitoring area. We call them the Restricted Floating Sensors (RFS). Figure 1 illustrates a RFS network deployed in a sea area. As shown, different sea depths result in different sizes of floating areas.

To map sea depth, both geographic positions and water depths are necessary. We find that such measurements can be acquired by figuring out sensors' floating areas. Section 5 details the process how we obtain water depth from the localization results of floating areas without any extra ranging devices. Thus our problem changes to determining the floating area of each sensor. Since the sensor nodes in the RFS network can float around, the traditional localization approaches for stationary sensors cannot work. On the other hand, simply treating the RFS network as a mobile sensor network and blindly applying those localization approaches for mobile WSNs does not capture the special nature of the RFS network. In the RFS network, sensor nodes float within restricted areas, providing us possibilities to capture their mobility models. By understanding RFS mobility behaviors, we can achieve higher accuracy with reduced overhead.

In this paper, we give an elaborate analysis on the localization problem in the RFS network. We build network models and establish the localization objective as locating the floating area of each sensor node. We equip a small portion of the network nodes with external locating devices such as GPS receivers, called seed nodes; while others are non-seed nodes. All sensor nodes estimate their distances from each other. Our approach applies different computation schemes for efficiently localizing the floating areas of the seed and non-seed nodes based on range distance estimations. By locating the sensor nodes, we can accordingly infer the sea depth at the anchor positions in a practical way, which prevents relying on other expensive specialized devices like sonar pingers.

We validate our design by launching a prototype system with 25 *Telos* sensor nodes off the seashore in HKUST campus. The results show that our prototype achieves less than 0.5m sea depth estimation error averagely. We conduct a large scale simulation to further test the system performance and scalability under various network settings. With precise distance measurements assumed, we can obtain the sea depth estimation with an average relative error within 20%.

The rest of the paper is organized as follows. In Section 2, we formally define the RFS network model and formulate the localization problem for RFS. We describe our localization approaches for seeds and non-seeds in Section 3 and Section 4. In Section 5, we

discuss measuring sea depth based on the floating area localization. We present the experiments and the results in Section 6. We summarize related work in Section 7 and conclude this work in Section 8.

2. The Network Model

In this section, we first give a definition of the more general Restricted Mobile Sensor (RMS) network.

DEFINITION 2.1 – RMS network. A sensor is called a restricted mobile sensor, if it is capable of movement but its movement is restricted within a local area of the application field. A network composed of restricted mobile sensors is called a RMS network. The RFS is a typical RMS network. Once anchored at a point, the sensor node floats on the sea surface but within a restricted area.

DEFINITION 2.2 – Floating area. In a RMS network, the movement of a sensor is limited in a restricted area. The restricted area may have different shapes due to different constraints of RMS networks. In the RFS network, each sensor node floats on the sea surface within a disk area centered at its anchor. This disk area is called the floating area of the sensor. We use $o(c, r)$ to denote a disk floating area, where c and r represent the centre and radius of the disk area, respectively. In practice, c is the anchored position of each sensor and r is determined by the length of the rope and the sea depth at the anchored position.

DEFINITION 2.3 – Floating model. In the RFS network, each sensor floats within its floating area. The movement is affected by many factors, e.g. ocean current, wind blow, tide etc. The factors above can hardly be modeled and mostly affect with randomness. In this case, the current position of a sensor is considered independent of its previous positions under non-negligible intervals between consecutive sampling times. Each sensor is assumed to appear in the floating area under uniform distribution and the probability distribution of the sensor position is given by:

$$f(x, y) = \begin{cases} \frac{1}{\pi r^2}, & (x, y) \in o(c, r) \\ 0, & \text{otherwise} \end{cases}$$

DEFINITION 2.4 – RFS network model. The targeted RFS network $N(S, O)$ consists of a set of sensors S and the corresponding set of floating areas O . Each sensor s_i moves within its floating area o_i under the floating model. The floating areas of different sensors are assumed non-overlapped, i.e. $\forall s_i, s_j \in S$ within floating area $o_i(c_i, r_i)$ and $o_j(c_j, r_j)$, $dist(c_i, c_j) > r_i + r_j$. This assumption prevents the possibility that two sensors get too close and their ropes become twisted with each other. This assumption is realistic in practice as

the sensor communication range is usually multiple times the radius of sensor floating area.

DEFINITION 2.5 – Neighborhood of RFS. In traditional sensor networks, the neighbors of a sensor s are defined as the set of sensors that have direct communications with s . While the neighborhood is relatively stable in static sensor networks, it is highly dynamic in mobile sensor networks. As a restricted mobile sensor network, RFS network shares similarity with traditional mobile sensor networks in that each sensor node has dynamic connections with its neighboring nodes. However, the locality of sensor movement in the RFS network constrains this dynamic effect. Therefore, we are able to introduce a more proper definition of neighborhood for RFS. Sensor s_i and s_j are defined to be neighbors *iff.* they can communicate with each other in their entire floating areas. Each node has direct communication with its neighbor nodes at any time. Under this definition, we obtain a stable neighborhood in RFS networks.

DEFINITION 2.6 – Floating area localization. In RFS networks, sensor nodes move within their floating areas under the probabilistic floating model. The localization issue in RFS networks is to obtain the floating areas instead of the instantaneous locations. The floating area localization in the RFS network indicates the process of locating the floating area $o(c, r)$ of each sensor, including the central anchor position c and the radius r . In the following, localization means floating area localization if not elsewhere specified.

DEFINITION 2.7 – Error. Let $o(c, r)$ be the floating area of sensor s and $\hat{o}(\hat{c}, \hat{r})$ be the estimated floating area. The localization error includes two parts: (1) error on the estimated anchor position $e_c = \text{dist}(\hat{c}, c)$; (2) error on the estimated radius $e_r = |\hat{r} - r|$. The relative error of floating area $o(c, r)$ is defined as a 2D vector $E(\hat{o}, o) = (e_c/r, e_r/r)$. The average localization error of a RFS network $N(S, O)$ is defined as

$$E(N) = \frac{1}{|O|} \sum_{o \in O} E(\hat{o}, o).$$

We design the Floating Area Localization Algorithm (FALA) to localize sensor nodes in a RFS network. In the localization process, all sensor nodes are able to measure the distances between themselves through RSS measurements. Other superior techniques like TOA, TDOA and AOA can be applied for higher ranging accuracy.

As a statistic based algorithm, FALA yields the localization result after a series of data sampling. During each sampling process, seeds collect their locations and non-seeds process the distance measurements. FALA applies different schemes for locating seeds and

non-seeds. Although seeds are able to know their instantaneous locations, further computation based on the location information is needed to determine their floating areas. For non-seeds, FALA derives their floating areas from distance information through a sequential process.

FALA includes four steps: sampling, seed floating area computing, non-seed floating area computing, and continuous data collection and accuracy improvement.

3. FALA for Seeds

As equipped with localization devices, seeds are aware of their instant positions. We carry out a series of samplings on seed positions. After a period of time, each seed node records a set of positions it resides in at different time. We estimate the floating area of seed nodes from the position sets.

Obviously, all sampled positions of a seed are certainly in its floating area under the floating model. In other words, its floating area should be a disk area at least containing all sampled positions. Thus, we can transform the localization problem to figuring out a disk area which covers a set of positions.

Apparently, there are many feasible disk areas, among which the smallest one should be considered the maximum likelihood estimation because it provides the highest probability of the occurrence of a set of positions. Thus, the smallest one is then considered the estimation of the floating area of the seed. As the sampled positions accumulate, the floating area is asymptotically approached. The problem is formulated as follows.

Given a set P of n points in the plane, find the smallest enclosing disk for P , that is, the smallest disk that contains all the points of P .

For simplicity, we assume that no three points are collinear and no four points are cocircular. In computational geometry, this problem is often called the Minimum Enclosing Disk (MED) problem.

It is not difficult to find a brute force solution to the problem which takes $O(n^4)$ running time. However, such an algorithm introduces intensive computational cost which is likely not suitable for the resource restricted sensor nodes.

A randomized algorithm[9, 17] for MED problem has been proposed in computational geometry domain, which takes $O(n)$ expected running time. It is observed that when a point is outside the MED of all other points, it must lie in the boundary of MED of all points. The following theorem [9] illustrates this observation.

Theorem: Let P be a set of points in the plane. Let R be a possibly empty set of points with $R \cap P = \emptyset$. Let $D(P, R)$ denote the minimum enclosing disk of P that contains R on its boundary. Then we have,

- (a) If a point $p \in D(P \setminus \{p\}, R)$, then $D(P, R) = D(P \setminus \{p\}, R)$;
- (b) Otherwise $D(P, R) = D(P \setminus \{p\}, R \cup \{p\})$.

Based on this theorem, the randomized algorithm *RMED* computes the MED of a given set P of positions. At the very beginning, we have no idea about which point lies on the boundary of MED, so the seed runs *RMED*(P, null) as a start.

Algorithm *RMED*(P, R)

- 1: if $P = \emptyset$ or $|R| = 3$,
 - 2: then $D :=$ the disc defined by R .
 - 3: else choose a random $p \in P$,
 - 4: $D := \text{RMED}(P \setminus \{p\}, R)$;
 - 5: If $p \notin D$,
 - 6: then $D := \text{RMED}(P \setminus \{p\}, R \cup \{p\})$.
 - 7: return D .
-

The *RMED* algorithm can be change to work in an incremental manner [9]. That is to say, being informed the current position from its positioning device at each sampling time, a seed updates its existing minimum enclosing disk and obtains a refined approximation of the floating area. In this online version of algorithm, seeds can start localization as early as possible without waiting for all n sampling positions collected. This feature well suits the data acquisition pattern of seed sampling process. Moreover, the updating process takes only $O(1)$ expected running time.

According to our algorithm, the approximated floating area \hat{o} is always smaller than the real one o . The error between o and \hat{o} keeps decreasing during the updating processes in which \hat{o} expands towards o . To minimize the estimation error, the seed needs to collect more sample data. However, a large sample capacity usually implies a long period of sampling. Therefore, we need to properly choose a sample capacity aiming for an acceptable accuracy.

We conduct a simulation to analyze the error of our estimation at different sample capacities of $n = 2, 3, 5, 10$ and 20 . The simulation results are shown in Figure 2. We find that when the number of samplings is 10, 80% of cases have less than 20% relative error of radius estimation and when the number of samplings is 20, 90% of cases have less then 10% error. In most applications, a number ranging from 10 to 20 induces an acceptable sample capacity for seeds to compute their floating areas.

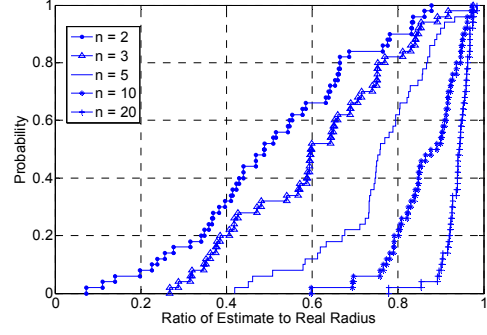


Figure 2: Cumulative distribution function of the estimated to real radius ratio

4. FALA for Non-seeds

When seeds have localized their floating areas, we need to utilize them as referees to locate non-seeds. Trilateration from referees is a widely used method to localize static nodes in stationary sensor networks. However, due to the dynamic property, directly using trilateration for RFS leads to poor accuracy. In this section, we propose a new scheme for locating non-seeds based on statistical measurements.

4.1. The Framework of Non-seed FALA

Before looking inside the non-seed FALA, we first define two concepts about computed and computable sensor nodes.

DEFINITION 4.1 – Computed and computable sensor nodes. We call a sensor node computed if its floating area is already known. If a non-computed sensor node has k ($k \geq 3$) computed neighbors, it is a computable sensor node.

The non-seed FALA is an iterative process, gradually transforming computable sensors to computed sensors. Figure 3 plots a deployment of four sensors: a non-seed s , with the floating area o unknown, and its three neighbors $\{s_i \mid 1 \leq i \leq 3\}$. Assume all s_i are computed nodes, that is, their floating area $o_i(c_i, r_i)$ are known. Let d_i denote the distance between s and s_i . Our goal is to estimate the floating area o of s . Clearly, with one time measurement there exists uncertainty for floating area computation. As shown in Figure 3, another disk area o' different from o is also possible to be a candidate of the floating area of s , because the current position of s which satisfies all distance constraints also resides in o' . We cannot distinguish the real area from o and o' at this stage. That means, it is impossible to calculate the floating area of s under a single time observation of d_i .

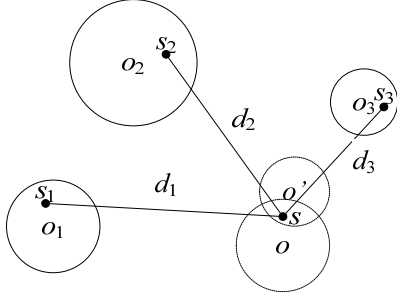


Figure 3: Non-seed localization

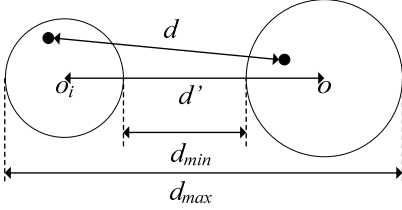


Figure 4: Anchor distance estimation

The distance measurement d_i varies all the time due to the movement of s and s_i . We observe that multiple samplings can alleviate the uncertainty for localization. If we treat o and o' as two sets of points in plane, when s moves to some position in $o - o'$, the distance sampling information negates the possibility of o' being the floating area. Furthermore, we know that d_i only depends on $o(c, r)$ and $o_i(c_i, r_i)$, irrespective of any other floating areas o_j ($j \neq i$). Hence, the sample distribution is determined by r_i , r , and the distance between two anchored positions $d'_i = \text{dist}(c, c_i)$, indicating that, to some extent, the samples statistics can imply r_i , r , and d'_i . Therefore, the relationship between the sample statistics and the parameters r_i , r , and d'_i is of great importance, based on which non-seeds can localize their floating areas.

Without loss of generality, we only consider s and a calculated neighbor s_i , as shown in Figure 4. For simplicity, we use d and d' instead of d_i and d'_i to elaborate the non-seed FALA. We know d varies all the time while d' is a static value. Let D denote the random variable of d and let d_i denote the observed value of D at sampling time t_i .

Three steps are included in the floating area computation of non-seed s , described as below.

1. A non-seed s samples the distance measurements d between s and its neighbors.
2. Based on sample statistics, s calculates d' and r .
3. If s is computable, it calculates the anchor position c by trilateration based on d' .

In step 1, the non-seed s carries out a sampling process. In step 2, s estimate the hidden parameters based on distance samples. We consider two methods

for exploring the relationship between sample statistics and the hidden parameters, based on the geometrical relationship and regression analysis respectively.

In step 3, although sensor nodes are mobile, their anchored positions are static. Thus, it is possible to solve a typical point localization problem for locating anchored positions. On the premise that the distances from an unknown anchor position to three known anchored positions are obtained, trilateration can be conducted to calculate the unknown anchored position. With c and r , this step completes the floating area computation of s and s becomes a computed sensor node.

4.2. Geometrical Relationship

A simple method for estimating d' and r is to explore the geometrical relationship between the two floating areas of s and s_i . We define $d_{max} = \max(D)$ and $d_{min} = \min(D)$ as the minimum and maximum values of D . As shown in Figure 4, d_{max} and d_{min} are obtained in two extreme situations. According to the geometrical relationship, we have:

$$d' = \frac{d_{max} + d_{min}}{2}$$

$$r = \frac{d_{max} - d_{min}}{2} - r_i$$

Such a method is simple to implement and takes little computation cost. In practice, it is reasonable to regard $\max(d_i)$ and $\min(d_i)$ as the estimation of d_{max} and d_{min} respectively. However, the extreme cases may not occur in sampling, under which we will get bad estimations of d_{max} and d_{min} . In addition, non-negligible ranging errors of existing approaches also heavily degrade the effectiveness of the method. On the contrary, the statistical method, based on sampling distributions, less suffers from this.

4.3. Regression Analysis

As we have observed, the distribution of D , to some extent, reflects the hidden parameters r and d' . This fact allows us to design a method to estimate r and d' based on sample statistics.

In our analysis, since r_i is a known parameter, we introduce two coefficients θ_1 and θ_2 , such that $r = \theta_1 \times r_i$ and $d' = \theta_2 \times r_i$. The sample statistics include the mean $\hat{\mu}$ and the standard deviation $\hat{\sigma}$ of samples, defined by

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n d_i$$

$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \hat{\mu})^2}.$$

A simulation study is conducted to explore the relationship between hidden parameters (θ_1 and θ_2) and sample statistics ($\hat{\mu}$ and $\hat{\sigma}$). Figure 5 gives us an important intuition about the relationship, that is, there exists linear relation between the parameters and the sample statistics.

We now synthetically take account of the impact of both θ_1 and θ_2 by using multiple regression analysis. Let β be a 2×3 coefficient matrix, our general form of two-variable linear regression equation is as follows:

$$\begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \end{bmatrix} = \beta \times \begin{bmatrix} \theta_1 \\ \theta_2 \\ 1 \end{bmatrix}.$$

Using least squares technique, we have

$$\beta = \begin{bmatrix} 0.0951 & 0.9820 & 0.0409 \\ 0.4507 & 0.0035 & 0.1956 \end{bmatrix}.$$

In summary, a node s first collects sample distances between itself and its neighbor s_i , and then calculates the statistics $\hat{\mu}$ and $\hat{\sigma}$. According to the regression model, s determines θ_1 and θ_2 ; and finally completes the estimation of r and d .

Errors of our regression model may come from two sources: (1) the residuals in regression analysis and (2) the inaccuracy of $\hat{\mu}$ and $\hat{\sigma}$. The residual figure, Figure 6, illustrates that the error of our linear regression model is relatively small if we consider the usual values of θ_1 and θ_2 . The inaccuracy of $\hat{\mu}$ and $\hat{\sigma}$ is usually due to a small sample capacity.

Taking error analysis of $\hat{\mu}$ as an example, the size of sample can be determined by the accuracy constraints. The normality test of sample data, Figure 7, suggests the sampling distribution is almost normal. Thus, the statistic

$$\frac{\hat{\mu} - \mu}{\hat{\sigma} / \sqrt{n}}$$

possesses a t -distribution. Then we get

$$P\left(\hat{\mu} - t_{\alpha/2, n-1} \frac{\hat{\sigma}}{\sqrt{n}} < \mu < \hat{\mu} + t_{\alpha/2, n-1} \frac{\hat{\sigma}}{\sqrt{n}}\right) = 1 - \alpha,$$

and the interval estimation of μ

$$\left(\hat{\mu} - t_{\alpha/2, n-1} \frac{\hat{\sigma}}{\sqrt{n}}, \hat{\mu} + t_{\alpha/2, n-1} \frac{\hat{\sigma}}{\sqrt{n}}\right)$$

is a $(1 - \alpha) \times 100\%$ confidence interval for the mean μ . The length of the interval is

$$l = 2t_{\alpha/2, n-1} \frac{\hat{\sigma}}{\sqrt{n}}.$$

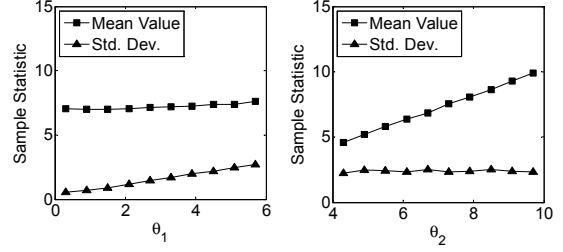


Figure 5: Hidden parameters vs. sample statistics

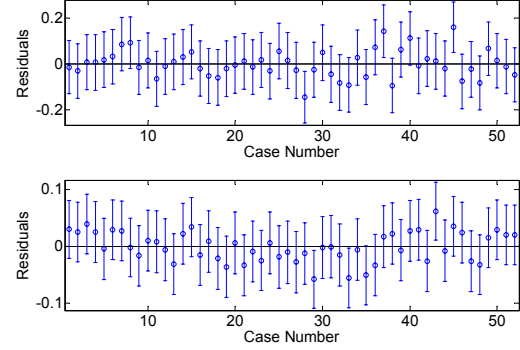


Figure 6: Residuals of regression analysis

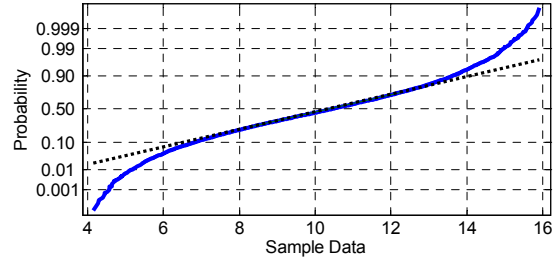


Figure 7: Normal probability plot of sample data

According to the t -distribution, $n = 30$ deserves a 90% confidence interval with an acceptable accuracy $l = 0.6204 \hat{\sigma}$.

5. Sea Depth Measurement by FALA

The ultimate goal of this work is to estimate the depth of the sea. By utilizing FALA, we can efficiently localize the sensor nodes in the network. When we use a rope with length L to anchor the sensor node on the sea of depth h ($L > h$), the sensor node floats within the disk area of radius $r = \sqrt{L^2 - h^2}$, as shown in Figure 8. After localization, we obtain the floating area of a node, achieving its center c as well as its radius r . We can then easily calculate the sea depth at position c . This calculation involves neither extra measurements nor hardware costs.

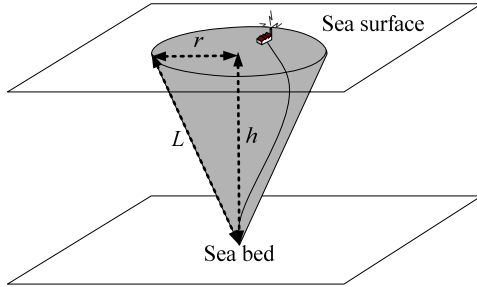


Figure 8: Geometrical structure of sea depth measurement



Figure 9: Prototype system deployment

When the sea depth of a monitoring region is deep, longer ropes are necessary. In this situation, the gravity of ropes cannot be ignored. When sensor nodes are on the boundaries of their floating areas, ropes cannot be straight but form a curve with steep upper part and mild lower part. Such a curve can be seen as a part of catenary. We can calculate the sea depth according to localization results and the equation of catenary [3].

6. Performance Evaluation

We first examine the effectiveness of our design by deploying a prototype system off the seashore. A large scale simulation is further conducted to test the system scalability under varied network parameters.

We evaluate FALA using three metrics: $E(c) = e_c/r$, $E(r) = e_r/r$, defined in Section 2, and $E(h) = |\hat{h} - h|/h$ to evaluate the error of sea depth measurement. In some previous literatures, the location error is represented relative to the hop size (the maximum communication range of a node) [6, 7]. However, for FALA

evaluation, if we use the communication range as a benchmark to measure location error, a 1m error contributes the same impact to a small floating area as to a large floating area, i.e. a 2m radius area and a 10m radius area. To diminish this unfairness, we adopt the relative error against the radii of floating areas in the evaluation. Since the communication range of each sensor node is usually 5~15 times larger than the radius of its floating areas in our experiment, the estimate errors are usually several times than they are against the sensor communication range.

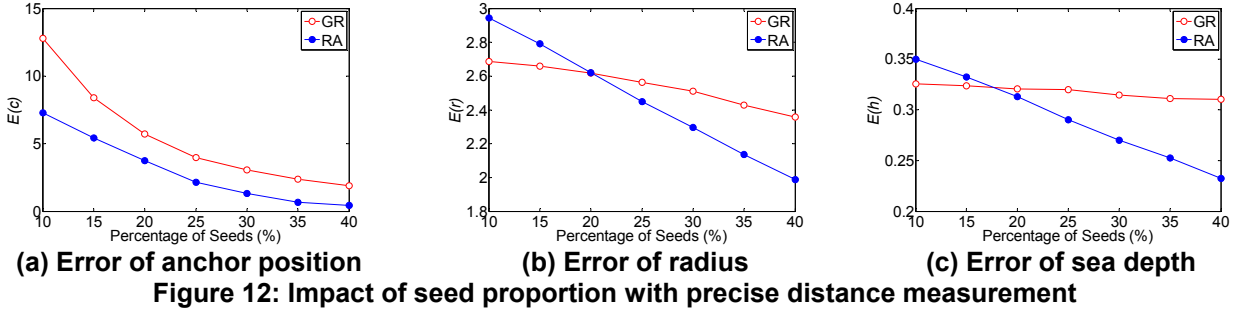
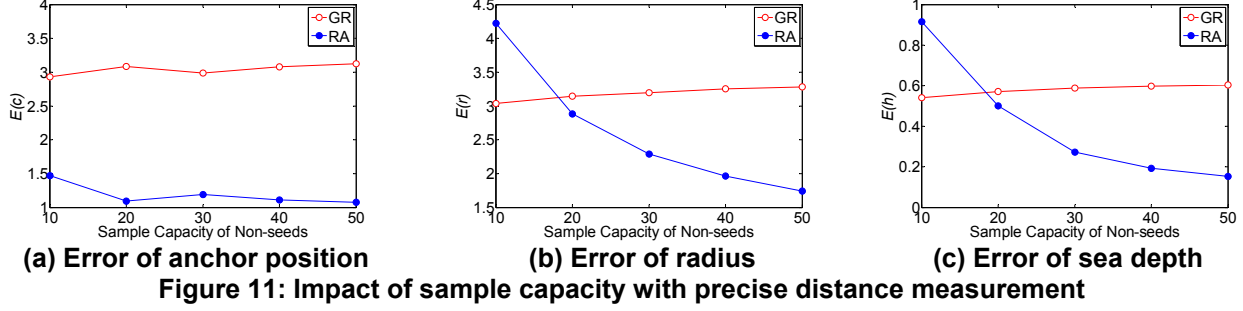
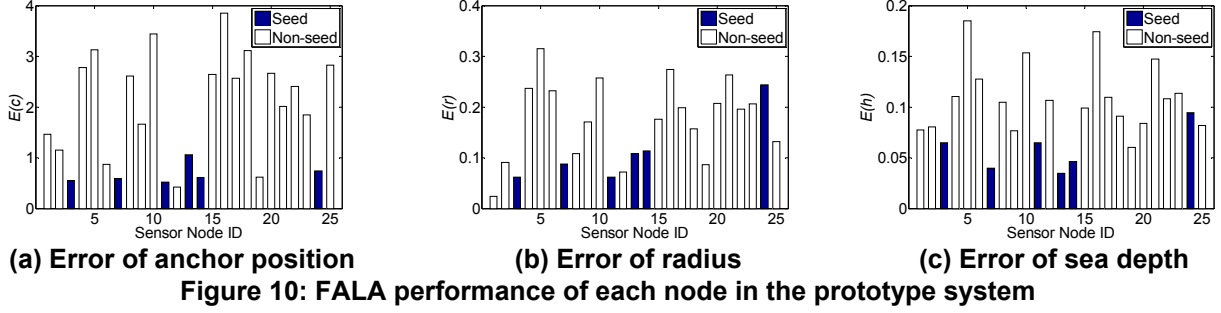
6.1 Prototype Experiment

To better understand the systematic behaviors of FALA, we deploy a prototype with 25 nodes off the seashore on university campus. The hardware layer of the prototype is constructed on the *Telos* motes with Atmel128 processor and CC2420 transceiver. We fit each node with a lightweight supporting shelf, which floats on the sea surface and raises the sensor node 150cm high above the sea surface. 25 such assembled floating nodes are anchored on a $100\text{m} \times 100\text{m}$ sea area where the water depth is around 4~7m. Figure 9 exhibits our deployment.

We utilize RSSI values from the transceivers to estimate the distances between nodes. The transmitting power of sensor nodes is set to 1mW and transmitting range could reach as far as 40m with more than -95dbm receiving signal strength. We construct a distance estimator according to the most widely used signal propagation model: the log-normal shadowing model [15]. Due to the coarse and non-monotone correspondence between the RSSI and distance in the real measurements, the relative error of the distance estimation can be up to 150%, which heavily limits the accuracy of FALA. We believe more precise distance estimating techniques such as TDOA or TOA based approaches will help to achieve better accuracy.

Figure 10 plots the FALA performance in our prototype system. The error of anchored position, as shown in Figure 10(a), is around 0.5~1 for seeds and 0.5~4 for non-seeds. For radius estimation, the error is around 0.05~0.3, illustrated in Figure 10(b). In Figure 10(c), we can see the relative error of sea depth is around 0.03~0.2. From Figure 10, seeds basically outperform non-seeds in all three metrics.

In practice, two factors limit our prototype from more accurate results: (1) the seawater near the seashore moves a little regularly rather than completely affected by randomness, which makes errors on our floating model assumptions; (2) the large errors in our RSSI based ranging technique contributes much to the estimation error of FALA.



6.2 Large-scale Simulation

We generate networks of 900 nodes randomly distributed in a square sea region. In our simulation, the sea region is designed to be a $600\text{m} \times 600\text{m}$ square and has a water depth around 10m. When a RFS network is deployed in this sea region, the radii of floating areas are 2~6m, which are determined by sea depth and rope length. A typical communication range of the sensor nodes is 30m and the average degree of network topology is 8. In all our measurements, we integrate the results from 100 network instances.

In our simulation, we varied two parameters, the proportion of seeds and the sample capacity, to examine FALA under different network settings. We test the performances of geometrical relation (GR) method and regression analysis (RA) method proposed in Section 4.

Precise distance measurements

We first assume precise distance measurement to explore the ideally achievable accuracy of FALA. Figure 11(a) plots the average error of anchored position. The

error of RA is below 1.5, which is lower than GR as the sample capacity varies in a wide range. When the size of sample is larger than 20, the extra gain from RA becomes trivial. Therefore, 20 can be a good choice of sample capacity considering the tradeoff between the accuracy and overhead. As shown in Figure 11(b), the average radius error of RA consistently decreases as the sample size increases; while GR is slightly getting worse. RA outperforms GR when sample capacity is larger than 20. The average error of sea depth, investigated in Figure 11(c), follows the similar pattern as the radius estimation.

We also examine the impact of the seed density on FALA, highlighted in Figure 12. All performance metrics get better when the seed density increases. There is notable gap between GR and RA in Figure 12(a). The error of RA is less than 2 when 25% seeds exist. In figure 12(b) we examine the radius estimation. We observe that both RA and GR yield smaller errors when inserting more seeds and RA is better than GR when seed proportion is larger than 20%. Figure 12(c) shows the error on sea depth measurements. Again, it follows the similar pattern as radius error does and RA yields the error from 0.35 to 0.2 when seed proportion varies from 10% to 40%.

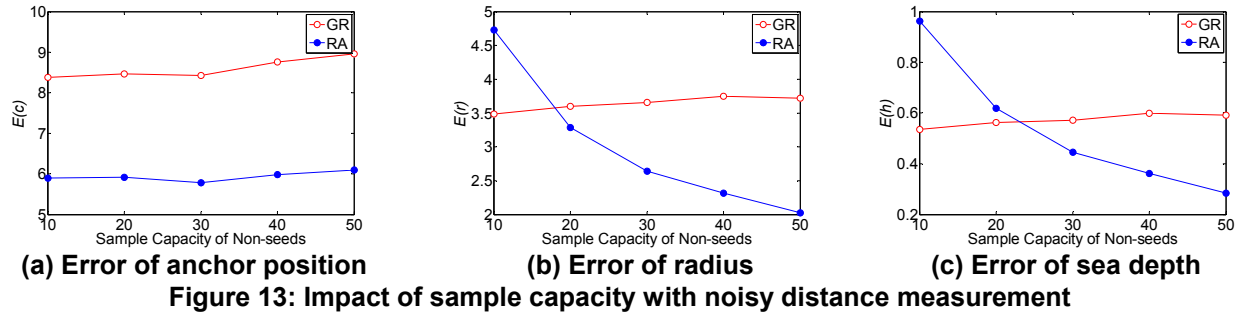


Figure 13: Impact of sample capacity with noisy distance measurement

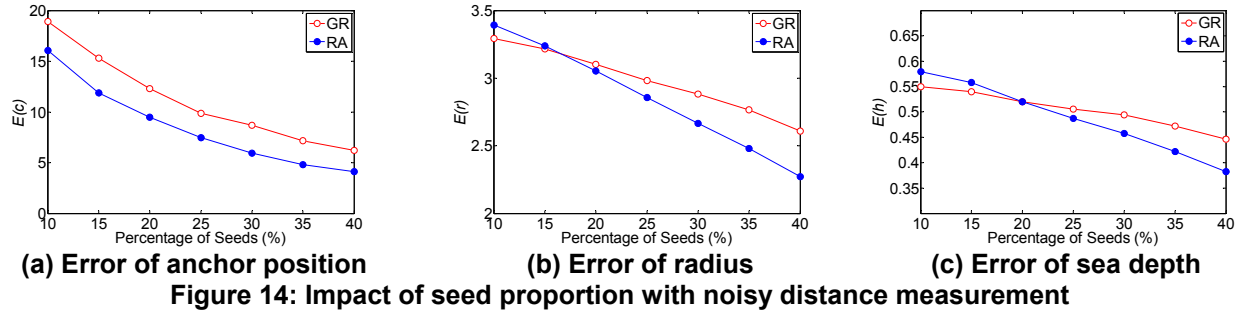


Figure 14: Impact of seed proportion with noisy distance measurement

Noisy distance measurements

We further evaluate FALA under noisy distance measurements. In our simulation, we introduce a zero-mean Gaussian noise with standard deviation of 50% of the real values into distance measurements.

Again, RA outperforms GR, as shown in Figure 13 and 14. Compared with Figure 11 and 12, the error of all performance metrics is larger than the corresponding errors with precise distance measurement. Especially, for anchored position estimation, the error can be 3 times larger. Clearly, a noisy distance measurement heavily degrades the performance of FALA. In such situations, 30 sample data and 30% seeds are necessary for precise localization.

7. Related Work

Recent advances in WSNs attract the attention of a lot of researchers [1, 13, 16, 18] with many efforts made for locating sensors [6, 8, 10]. According to the targeted environments, previous localization approaches can be classified into two types: for static sensor networks and for mobile sensor networks.

The static localization problem has been extensively studied for WSNs. The proposed localization approaches typically use a small number of seed nodes that are aware of their location. Moreover, ranging measurements [6, 10, 11, 14, 16] (in range-based approaches) or neighborhood information [4, 8, 12] (in range-free approaches) are utilized to locate non-seed nodes. All these approaches assume the invariability of sensor locations. Once a sensor node knows its location, it can be used as a beacon to locate other sensor nodes. Such a strategy fails in our RFS context due to the movement of sensors.

Some of the static localization approaches [12, 14] can be extended to conform to the mobile environment. Most of them, however, cannot yield results in real time and thus suffer from estimation latency and inaccuracy brought on by sensor movements. Bergamo and Mazzini's [2] is one of the first works related to the localization problem in mobile sensor networks. Two fixed seeds are assumed transmitting across the entire network and other nodes can measure the received signal strength accurately. L. Hu and D. Evans propose a statistic based localization approach for mobile sensor networks in [7] based on the MCL method [5], which originates from a mobile localization problem in robotics. Mobility creates obstacles to accurate localization, resulting in large errors and heavy communication cost. In addition, dense seed deployment is required in that proposed approach.

None of above schemes considers a restricted movement model for sensor nodes. Directly using those localization approaches does not capture the special movement behaviors of RFS networks. Hence, they suffer from either inaccurate localization results or unnecessary estimation overhead.

8. Conclusions and Future Work

We discuss a novel sea depth measurement application using wireless sensor networks. We define the localization problem in RFS networks and introduce the concept of floating area localization, so as to determine the floating areas of sensor nodes. A statistical approach, FALA, is designed, based on which the sea depth can be acquired without expensive sonar systems. A prototype with 25 *Telos* nodes is deployed on a sea surface, and intensive large-scale simulations are con-

ducted to examine the efficiency and scalability of the proposed approach.

This work is still at its early stage. The future work leads into following directions.

(1) One assumption in our floating model is that the sensors float within their anchored areas under randomness, which in our prototype test is shown to be inadequate. The seawater near the seashore moves a little regularly rather than completely affected by randomness. The wave may also introduce errors of estimations. Thus a well model of the behaviors of the sea will help diminish their negative impact or even make use of their regularity to achieve more accuracy.

(2) The system scalability is also an important issue we need pay special attention to. Since the RSSI based distance measurement bears a large error, there is a trend of error propagation on our estimations when the network size significantly increases, especially under a small percentage of seeds. Whether or not we are able to design a sound collaborating mechanism at the layer of network topology, so that we can suppress the localization errors throughout the network, is a significant but challenging issue.

(3) Sea depth estimation is of great interests and importance for many sea monitoring applications. Our FALA approach yields the estimations of sea depth by utilizing the result of the floating area localizations, reducing the cost. This approach, however, also has its own limitations, e.g. the anchor of each sensor can actually get buried by the silt, which leads to inaccurate estimations as the time passes by. Is there any other light-weight approach for measuring the sea depth? Due to the intensive needs on the sea depth measurement and the difficulty of employing infrastructures at sea, we believe WSN is one of the best candidates for this application.

(4) The Restricted Floating Sensors describe a general model for sensor deployment which might be suitable for many sensing applications carried out on the sea. Under different contexts of the sensing applications, we might concern different factors of the network besides the locations, such as sensor coverage, network connectivity, data samplings, etc. Due to the nature of restricted mobility, the RFS network introduces the intermediate dynamics between the static network and mobile network. By developing mechanisms over the dynamics but taking advantage of the restriction on the mobility, can we achieve higher efficiency? We believe it is non-trivial and highly related to the concerned factors and the application context.

We are currently continuing this project for answering part of above questions and deploying a real working system together with the research group from Ocean University of China.

Acknowledgement

This work is supported in part by the Hong Kong RGC grant HKUST6169/07E, the National Basic Research Program of China (973 Program) under grant No. 2006CB303000, the National High Technology Research and Development Program of China (863 Program) under grant No.2007AA01Z180, and NSFC Key Project grant No. 60533110.

References

- [1] X. Bai, S. Kumar, Z. Yun, D. Xuan, and T.-H. Lai, "Deploying Wireless Sensors to Achieve Both Coverage and Connectivity," in Proceedings of ACM MobiHoc, 2006.
- [2] P. Bergamo and G. Mazzini, "Localization in Sensor Networks with Fading and Mobility," in Proceedings of IEEE PIMRC, 2002.
- [3] W. H. Beyer, *CRC Standard Mathematical Tables*, 28 ed: CRC Press, 1987.
- [4] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less Low Cost Outdoor Localization for Very Small Devices," *IEEE Personal Communications Magazine*, 2000.
- [5] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo Localization for Mobile Robots," in Proceedings of IEEE ICRA, 1999.
- [6] D. Goldenberg, P. Bihler, M. Cao, J. Fang, B. Anderson, A. S. Morse, and Y. R. Yang, "Localization in Sparse Networks using Sweeps," in Proceedings of ACM MobiCom, 2006.
- [7] L. Hu and D. Evans, "Localization for Mobile Sensor Networks," in Proceedings of ACM MobiCom, 2004.
- [8] M. Li and Y. Liu, "Rendered Path: Range-Free Localization in Anisotropic Sensor Networks with Holes," in Proceedings of ACM MobiCom, 2007.
- [9] Mark de Berg, M. v. Kreveld, M. Overmars, and O. Schwarzkopf, *Computational Geometry: Algorithms and Applications*, 2nd ed: Springer-Verlag, 2000.
- [10] L. M. Ni, Y. Liu, Y. C. Lau, and A. Patil, "LANDMARC: Indoor Location Sensing Using Active RFID," in Proceedings of IEEE PerCom, 2003.
- [11] D. Niculescu and B. Nath, "Ad Hoc Positioning System (APS) using AoA," in Proceedings of IEEE INFOCOM, 2003.
- [12] D. Niculescu and B. Nath, "DV Based positioning in Ad hoc Networks," *Journal of Telecommunication Systems*, 2003.
- [13] S. Patten, B. Krishnamachari, and R. Govindan, "The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks," in Proceedings of ACM/IEEE IPSN, 2004.
- [14] A. Savvides, C. Han, and M. B. Strivastava, "Dynamic Fine-grained Localization in a Ad-hoc Networks of Sensors," in Proceedings of ACM MobiCom, 2001.
- [15] S. Y. Seidel and T. S. Rappaport, "914 MHz Path Loss Prediction Models for Indoor Wireless Communications in Multi-floored Buildings," *IEEE Transactions on Antennas and Propagation*, vol. 40, pp. 209-217, 1992.
- [16] Q. Wang, R. Zheng, A. Tirumala, X. Liu, and L. Sha, "Lightning: A Fast and Lightweight Acoustic Localization Protocol using Low-End Wireless Micro-Sensors," in Proceedings of IEEE RTSS, 2004.
- [17] E. Welzl, *New Results and New Trends in Computer Science*: Springer-Verlag, 1991.
- [18] G. Xing, C. Lu, Y. Zhang, Q. Huang, and R. Pless, "Minimum Power Configuration in Wireless Sensor Networks," in Proceedings of ACM MobiHoc, 2005.