2010

Reactive localization in underwater wireless sensor networks

Mohamed Watfa  
*University of Wollongong in Dubai*

Tala Nsouli  
*American University of Beirut*

Maya Al Ayyash  
*American University of Beirut*

Omar Ayyash  
*American University of Beirut*

Publication Details

Reactive Localization in Underwater Wireless Sensor Networks

Mohamed K. Watfa*
Computer Science Department
University of Wollongong in Dubai,
Dubai, UAE
mohamedwatfa@uowdubai.ac.ae

Tala Nsouli, Maya Al-Ayache, Omar Ayyash
Computer Science/Computer Engineering Department
American University of Beirut
Beirut, Lebanon
Mma,tkn02@aub.edu.lb

Abstract—In this paper, we present a novel technique for localizing an event of interest in an underwater environment. The network consists of randomly deployed identical sensor nodes. Instead of proactively localizing every single node in the network as all proposed techniques set out to do, we approach localization from a reactive angle. We reduce the localization problem to the problem of finding 4-Node Coverage, in which we form a subset of nodes such that every node in the original set is covered by four nodes belonging to this special subset — which we call the anchor nodes for simplicity. This subset of anchor nodes behaves like a backbone to the localization process. We show that in terms of energy consumption, this localization technique far surpasses others in terms of energy efficiency.

Keywords—localization; reactive localization; underwater sensor networks; energy-efficiency; k-node coverage

I. INTRODUCTION AND MOTIVATION

During the past few years, a significant interest in monitoring aquatic environments has emerged. Such a process was driven by major incentives such as scientific exploration, commercial exploitation, and coastline protection. These functions were made feasible by applying underwater communications among underwater devices. Underwater wireless sensor networks comprise a number of sensor nodes and vehicles installed at different levels of the ocean (surface, bottom, and mid-ocean) to perform various functionalities, most importantly monitoring the ocean environment. These sensor networks share some properties with ground sensor networks, most notably the large number of nodes and the limited energy constraint, which poses a challenge in deploying and managing large scale wireless sensor networks. With wireless sensor networks, whether underwater or terrestrial, localization will inevitably be discussed. The importance of localization takes shape in the fact that much of the data obtained through these sensor networks must be location-aware. All the localization schemes designed for underwater sensor networks (detailed in the Related Work section) handle the localization problem proactively. That is localization of all the nodes in the network is performed as a kind of initialization phase, meaning before the network is put to its actual use. However, if we study the motivation behind localization, we find that it is not necessary to know the location of every node in our network, since our aim is to localize an event of interest, rather than the node itself. Keeping that aim in mind, we notice that the energy expenditure incurred by a proactive localization algorithm is an unnecessary cost, and can be reduced by rendering the localization event-driven and therefore devise an energy-efficient reactive localization scheme.

Since radio frequency waves do not propagate well underwater, UWSNs resort to acoustic waves for communication. Acoustic waves are five times slower than RF waves, magnifying the propagation delay in UWSNs. Moreover, the speed of sound underwater is variable, being a complex function of temperature, pressure and salinity. Also the three-dimensional vast underwater environment poses a great challenge. The underwater environment is also governed by currents and wildlife which poses the problem of node mobility as well as the problem of interfering noise – both man-made and ambient. Moreover, underwater sensor networks are challenged by two types of path loss. The first is attenuation, which is provoked by absorption of the acoustic waves, their conversion into heat, scattering, reverberation, refraction and dispersion. The second is geometric spreading, which is best described as the spreading of sound energy due to expansion of wave fronts. The challenges that most affect localization in UWSNs are mainly the three-dimensional environment, which imposes a third dimension (unknown) to be determined by the algorithm. This calls for extra resources to make localization possible. Another challenge specific to localization is that the high delay in UWSNs is paired with a delay variance that makes the computation of RTT inaccurate (and hence its use not so effective). In our scheme, we reactively localize a node that detects an event, using a previously selected subset of anchor nodes with known positions.

The remaining sections of this paper are organized as follows. Section II provides an extensive overview of related work as compared to the technique we propose. In Section III, we elaborate on the details of our technique’s function and architecture. Section IV and V provides an evaluation of the Reactive Localization Scheme both theoretically and through simulations. Section VI concludes this paper with a summary of the work done and an outline of future work.

II. RELATED WORK

Localization of sensor nodes in terrestrial environments has been widely explored in the past. The schemes proposed can be classified under two approaches, direct approaches and indirect approaches. Direct approaches, such as GPS-based localization, involve absolute localization which does not particularly apply in underwater environments since such approaches are neither practical nor scalable nor adapt well with node mobility [1]. Indirect approaches are known as relative localization, since nodes position themselves with respect to their neighboring nodes. Commonly indirect approaches entail a small subset of nodes knowing their locations (via GPS), sending location information to neighboring sensor nodes, thus allowing them to calculate their relative locations. The localization process within the indirect approach can be classified into range-based localization and range-free localization. Range-based protocols provide more accurate location estimates as they use absolute point-to-point estimates; however they need additional complex hardware capacity thus increasing the cost. Range-free schemes are more cost-effective but provide less accurate location estimates. Range-based schemes are potentially good choices for underwater sensor networks.
Terrestrial localization has been widely investigated, but due to the several challenges posed by the underwater environment, common algorithms cannot be directly applied underwater. And thus in the recent years, authors have proposed localization schemes for small-scale underwater static networks such as [2, 3]. Some of these schemes use surface buoys and one hop communications between sensor nodes, such as GIB (GPS Intelligent Buoys) [3], and PARADIGM [2]. GIB is an “underwater GPS” system that relies on a centralized server to compute location information for nodes. PARADIGM involves autonomous underwater vehicles (AUV) computing their locations on-board. Another scheme that uses AUV is presented in [8]. Erol et al. of “AUV-Aided Localization for Underwater Sensor Networks” [8] present a localization scheme for underwater sensor networks based on the use of an AUV (Automated Underwater Vehicle) that probes the underwater sensor field and assists nodes in calculating their coordinates. The proposed scheme assumes no initial infrastructure or synchronization between the nodes. Calculations and estimations gathered while the AUV is in motion result in significantly erroneous measures.

Hahn et al. in [5] put forward a centralized scheme that involves a sensor interrogating multiple surface buoys. It entails a ping-pong style that measures the round-trip delay for estimating ranges. Opposite to that, our reactive scheme is keen on balancing an efficient communication overhead. In [4], a silent positioning scheme is proposed where sensor nodes learn their locations by passively listening to beacon messages being delivered between neighbors. However in this scheme and contrary to our proposed one, it is not certain that we have four anchor nodes covering the node to be localized. Contrary to our range-based scheme, [6] present ALS, which is area-based and range-free. It relies on the deployment of special anchor nodes that are capable of adjusting their power levels to divide a two-dimensional region into sub-regions. Our localization system is characterized by a finer position granularity than ALS. That is to say, the positions of the sensor nodes obtained are within a coordinate system rather than positions within a sub-region. Additionally, Zhou et al. of “Localization for Large-Scale Underwater Sensor Networks” [7] propose a localization scheme that approaches the problem in a range-based hierarchical manner. The process is divided into two sub-processes: anchor node localization and ordinary node localization. They tackle this by integrating a three-dimensional Euclidean distance estimation method and a recursive location estimation method. Even though Euclidean estimation reveals to perform best in anisotropic topologies, it is hindered by its large computation and communication overheads. Anchor node localization is achieved through relying on surface buoys equipped with GPS sensors. The anchor nodes localize themselves based on the “underwater GPS” scheme, GIB (GPS Intelligent Buoys) [3]. This scheme is hindered by disregard to energy constraints and high communication overhead since it adapts continuous message flooding. It also entails higher deployment cost since it relies on a relatively big number of anchor nodes. A new approach to the underwater localization problem is posed by Z. Zhou in SLMP [9] where mobility is taken into consideration. The mobility predictions are prone to failure due to the random and sudden nature of many underwater movements (tides, animal interference, ships, etc...). Yet another localization scheme for sparse 3D environments [10] transforms the three dimensional problem into a two dimensional one using projection techniques.

All the localization schemes designed for underwater sensor networks handle the localization problem proactively performed as a kind of initialization phase before the network is put to its actual use. However, if we study the motivation behind localization, we find that it is not necessary to know the location of every node in our network, since our aim is to localize an event of interest, rather than the node itself. Keeping that aim in mind, we notice that the energy expenditure incurred by a proactive localization algorithm is an unnecessary cost, and can be reduced by rendering the localization event-driven. That motivated us to devise an energy-efficient reactive localization scheme.

III. REACTIVE LOCALIZATION

In this section, we propose a scalable localization scheme for three dimensional underwater sensor networks.

A. The Architecture

The architecture in which the Reactive Localization algorithm will apply is one equipped with two types of nodes. The sensor nodes and the surface buoys. Sensor nodes are randomly deployed over the desired area such that we assume that nodes will randomly sink to different depths depending on their densities. The nodes are therefore randomly deployed in the three dimensional environment. After selecting a subset of nodes, we refer to them as anchor nodes. The surface buoys are equipped with GPS. The sink is located on the surface in a well-equipped station where information will be gathered and computation will be possible. We detail three consecutive phases to solve the localization problem in an underwater 3D network:

1) Finding the anchor nodes
   a. Find a subset of nodes that provide 4-coverage
   b. Locate the anchor nodes

2) Reactive localization of sensor nodes
   a. A sensor node detects an event
   b. The sensor node localizes itself using the anchor nodes

3) Delivery of information
   a. Assuming a routing algorithm, the node transmits its location and information about the sensed event back to the sink

B. Finding the Anchor Nodes

The first step is to find a subset of anchor nodes such that every sensor node is in the range of 4 anchor nodes. Every sensor node in the network must be covered by 4 non-coplanar anchor nodes.

\[ D(a, b) \leq \min(C(a), C(b)) \]

**Theorem 3.1:** In a k-1 dimensional environment, for a node to be localized, it must be covered by at least k nodes (k > 1).

**Proof:** As shown in Figure 1, three anchor nodes will only narrow down the choice of the location to two points. Having a fourth anchor node that is not coplanar with the first three, will make it possible to pinpoint the exact location of the sensor node in question.

\[ a \in \mathbb{R}^n \]

**Figure 1:** Importance of 4 anchors to localize a node in a 3D environment

Some points that we need to take into consideration are:

- There should exist at least D+1 anchors to uniquely localize a network in a D-dimensional space.
- To guarantee k Node-Coverage, each point should be within the sensing range of k or more sensor nodes.
- A 3D environment implies that we need 4 non-coplanar points
We elaborate on these points of Node-Coverage in order to rationalize the 4-Coverage Algorithm. We develop the idea of localizing a node in three dimensional spaces to solving for three unknowns (x, y, z). Mathematically, to be able to assign values to these three unknowns, we need four equations. The coverage algorithm guarantees 4-node coverage, which means that every sensor node in S should be within the communication range of 4 or more anchor nodes. The 4 anchor nodes, which are aware of their locations, will provide the sensor node attempting to localize itself with the needed four equations to solve for the three unknowns that will ultimately define its absolute location in the underwater medium. We will later provide a mathematical proof on how our proposed scheme effectively deals with the possibility that the four anchor nodes might be coplanar.

Algorithm 1: K-Node Coverage

1: Send Hello Messages (ID, Energy)
2: Construct set of neighbors \( N_i \)
3: Broadcast set of neighbors \( N_i \)
4: Node waits for all neighbors to respond with sets
5: if node \( i \) receives 1 message with \( ||N_i|| \leq k \), then it is critical
6: if node \( i \) receives all messages with \( ||N_i|| > k \), then it can be turned off, sends REQUEST_TO_SLEEP message (after a time proportional to energy)
7: Nodes hearing the requests sends GO_TO_SLEEP to requester with lowest energy first
8: After receiving GO_TO_SLEEP from all neighbors, we send SLEEP and turns off
9: Step 7 for other requesters

Our reactive localization scheme begins with an initialization process that determines a subset of nodes, called the anchor nodes, such that every sensor node (ordinary node) is covered by four anchor nodes. That is achieved by the K-Node Coverage Algorithm, in the case when \( k \) is set to be equal to 4. After randomly deploying the sensor nodes in the underwater environment, every node broadcasts a hello message with its ID number and energy level to its neighbors. Every node, upon receiving the hello messages from all of its neighbors, constructs a table of its neighbors, and then broadcasts a hello message to its neighbors. A node waits for time \( \tau \) till it receives the neighbor sets from all of its neighbors. At that point every node is aware of its neighbors and the neighbor set of each of it neighbors. If one of the sets received by a node is of a size equal to 4, then the receiving node is a critical node and cannot be turned off. If all of the sets received by the node are of size greater than 4, then the node may be turned off, and so it waits for a period of time inversely proportional to its energy level, and then broadcasts a REQUEST_TO_SLEEP message. By waiting for a period of time inversely proportional to its energy level, we are giving nodes with the lowest energy level the priority of going into the sleep state. Nodes hearing the REQUEST_TO_SLEEP send a GO_TO_SLEEP message to the requester with the lowest energy level first. If a node receives a GO_TO_SLEEP from all its neighbors, it will broadcast a SLEEP message and goes into a sleep state. After the completion of that phase for all requesters, the nodes that remain awake are the chosen subset we shall refer to as anchor nodes, and the nodes in the sleep state are the sensor nodes. The communication complexity of K-Node Coverage is \( O(nm) \) where \( n \) is the total number of nodes deployed, and \( m \) is the maximum number of neighbors a node has.

After finding the subset (anchor nodes), we tackle the problem of localizing the chosen nodes. To localize the anchor nodes, we resort – as previously mentioned – to regarding anchor nodes as nodes that are capable of communicating with surface buoys and localizing themselves. We assume this property for all deployed nodes since the subset of anchor nodes is determined after deployment and thus no nodes are “special”. Using existing underwater GPS systems, such as GIB [3], the anchor nodes with their ability to communicate with several surface buoys can localize themselves. Obviously, due to the complexity and energy consumption of GIB, it cannot be used on all the deployed nodes leading to our proposed work.

C. Reactive Localization of Sensor Nodes

After the anchor nodes are selected and localized, we outline the function of sensor nodes upon detecting an event. First a sensor node detects an event. The sensor node broadcasts a message to its one-hop neighbors, four of which will be acting as anchor nodes based on the 4-Coverage Algorithm. The message broadcasted will be referred to as a Localization Request Message. Once the anchor nodes receive the messages, they reply with their location information. The node hence localizes itself, using this information, by quadrilateration. We describe quadrilateration by briefly defining multilateration. Multilateration is a range-based localization scheme, in which the sensor node measures distances to anchors by time of flight (TOF). Mathematically, we need \( n+1 \) (4) linearly independent equations to solve a system in \( n \) (3) dimensions. These four messages, sent from four different anchor nodes, will make four sets of coordinates available to the node, which it uses to solve the equations:

\[
(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2 = d_i^2
\]

It follows from this definition of multilateration that quadrilateration is the localization process in which nodes measure distances from 4 reference points.

We will have two modes for sensor nodes: Localized and Non-Localized. Initially all sensor nodes (non-anchor nodes) have a Non-Localized state. These two states are governed by a timer. Once localized, a node will have a Localized status for a preset interval of time. When the time expires, the node discards its location and its status is once again Non-Localized. This process ensures that if a node that detects an event continues to detect it for a consecutive period of time, it will not have to localize itself several times.

To elongate the lives of the sensor nodes and conserve energy, we make it such that the sensor nodes have sleep/wakeup cycles. While asleep, the sensor nodes cannot communicate with each other but continue to sense the environment and try to detect events. Once an event is detected, the sensor node wakes up. Periodically, the sensor nodes wake up in case other sensor nodes are trying to contact them for self-healing. These sleep/wakeup cycles efficiently maintain energy levels and make it possible for the sensor nodes to function normally at the same time. Anchor nodes are always awake and listening for some sensor node that may attempt to contact them for localization information.

D. Delivery of Information

The idea behind this algorithm is localizing a node that detects an event and thus obtaining a rough estimation of the event's location. It is understood in this scheme that several nodes may detect the same event. In this case, all of these nodes will send localization requests. The information from all of the nodes is sent back to the sink, where the messages are interpreted and a more accurate localization for the event is obtained. This part of the process can be seen as a range-free localization of the event.
IV. THEORETICAL ANALYSIS

In this section, we provide in depth theoretical analysis and proofs of some of the stated theorems and assumptions.

A. K-Node Coverage Localizing Algorithm

Theorem 5.1: The probability that the 4 anchor nodes covering the sensor node all lie on the same plane, \( P_{\text{coplanar}} \), is 0.

**Proof:** Since anchor nodes are not selected beforehand, we have no special control on their deployment and thus locations. This poses problems, one of which is the probability of four anchor nodes involved in localizing a fifth node lying on the same plane. If the four nodes lie on the same plane, we cannot properly localize a fifth node using them. This case must be handled; we will do so by proving that this event’s probability is zero.

Consider 3 points \( A(x_A, y_A, z_A), B(x_B, y_B, z_B), C(x_C, y_C, z_C) \) of known positions and a 4th point \( D(x,y,z) \). The problem is proving \( D \in \Delta ABC \). Although \( D(x,y,z) \) might be correlated to the positions of \( A, B, C \), we make no assumptions about this correlation. However, we can safely say that \( x_D, y_D, z_D \) are logically independent and thus probabilistically independent. Moreover, due to the many factors affecting current, drift, velocities, etc… we can assume that the nodes’ distribution is sufficiently random (i.e. continuous and thus free of direct deltas and probabilistic peculiarities). To simplify the analysis, we will assume that the distribution of \( x, y, z \) are normal distributions. Let \( \eta(\mu, \sigma) \) be the normal distribution with mean \( \mu \) and variance \( \sigma \). \( x \sim \eta(\mu_x, \sigma_x), y \sim \eta(\mu_y, \sigma_y) \) and \( z \sim \eta(\mu_z, \sigma_z) \). For \( A, B, C, \) to be coplanar, \( \theta_1 = \theta_2 = \theta_3 = \theta_4 \). Let \( \theta = \theta_1 = \theta_2 = \theta_3 = \theta_4 \) be the unit vector in the direction of \( \theta \) and \( \eta = AB \times AC \) (normal vector). This means that:

\[
\begin{align*}
\frac{x - x_D}{\eta} &= \frac{y - y_D}{\eta} = \frac{z - z_D}{\eta} \\
\text{which implies that}
\end{align*}
\]

Since \( x_D, y_D, z_D \) are Gaussian random variables, then \( x_D x_A, y_D y_A, z_D z_A \) are also Gaussian. Moreover \( x_D, y_D, z_D \) are Gaussian (only manipulation with constants and elements are jointly Gaussian). So \( \begin{bmatrix} x_D \\ y_D \\ z_D \end{bmatrix} \) is a Gaussian vector and thus \( \begin{bmatrix} x_D' \\ y_D' \\ z_D' \end{bmatrix} \) has a Rayleigh distribution. The above vectors are unit vectors and thus for them to be equal they must have the same angles \( \theta \); However, it is proven that in such vectors, \( \theta \) and \( \emptyset \) have uniform distributions. So, the problem reduces to the probability of

\[
\begin{align*}
\theta &\sim \theta^* \\
\emptyset &\sim \emptyset^* \\
\emptyset_n &\sim \emptyset_n
\end{align*}
\]

The probability of which is identically 0 (since continuous uniform distance). So, we can conclude that the probability that the 4 anchor nodes covering the sensor node all lie on the same plane, \( P_{\text{coplanar}} \), is 0.

Definition 5.1: A node is critical when one of its neighbors needs it to be k-covered.

Lemma 5.1: Critical nodes are never turned off.

**Proof:** Consider a node \( s_i \). The node \( s_i \) waits for time \( \tau \) to receive the set of neighbors \( N_i \) from every neighbor \( s_j \). Considering bidirectional links, \( s_i \) will be an element of every \( N_j \) received by \( s_i, s_i \in N_j \) if \( s_i \) received \( N_j \) from \( s_j \). We consider three cases:

- If for some \( j, ||N_j|| < k: s_j \) has less than \( k \) neighbors including \( s_i \), and in that case \( s_i \) cannot be localized, since it is not covered by \( k \) nodes. In that case \( s_i \) is considered critical, and is kept turned on.

- If for some \( j, ||N_j|| = k: s_j \) has exactly \( k \) neighbors including \( s_i \), and in that case \( s_i \) is critical since to be localized, \( s_i \) needs to be covered by \( k \) nodes. Since \( s_j \) has exactly \( k \) neighbors, then each one of its neighbors is critical for it to be localized. In that case \( s_i \) is considered critical, and it does not send a REQUEST_TO_SLEEP message, and hence remains in the awake phase.

- If for some \( j, ||N_j|| > k: s_j \) has more than \( k \) neighbors including \( s_i \), and in that case any of \( s_j \)’s neighbor can be turned off since we only need \( k \) nodes covering it. In that case \( s_i \) is not considered critical, and it broadcasts a REQUEST_TO_SLEEP message, which is received by all its neighbors. The node does not sleep yet, until it receives a GO_TO_SLEEP message from all of its neighbors.

Since the algorithm guarantees that only non-critical nodes send a REQUEST_TO_SLEEP message, then critical nodes are guaranteed to remain awake.

V. ANALYSIS AND EVALUATION

In our simulation experiments, 500 sensor nodes are randomly distributed in a 100m x 100m x 100m region. We define node density as the expected number of nodes in a node’s neighborhood; hence node density is equivalent to node degree. We control the node density by changing the communication range of each node while keeping the area the same. We study the differences as compared to other underwater localization schemes (mainly [7], [8], and [12]).

A. Localization Coverage

Localization coverage is defined as the ratio of localizable nodes to the total number of nodes. Clearly, as node density increases, localization coverage increases. Once the nodes are dense enough so that the subset of anchor nodes can be sufficiently completed, then localization coverage will be at 100% and errors will be small. Since a complete set implies that the condition of every sensor node being covered by four anchor nodes is achieved and hence whenever a sensor node needs to be localized, it can be localized. In other words, every node is hence localizable. The percentage of coverage increases linearly as node density increases. It also increases as the subset grows to incorporate more anchor nodes. This implies that we may be able to overcome the low localization coverage in sparse networks by making our subset larger. In comparison to the hybrid scheme and the recursive scheme [12] (Figure 2), our algorithm is slightly lower that the hybrid scheme in terms of localization coverage with lower density; however, it quickly catches up to achieve the same results with more accuracy. We notice that the difference is not very big at the beginning because we choose our anchor nodes to optimally cover the nodes in the area, but the hybrid algorithm achieves slightly better coverage due to their use of recursion.

B. Localization Error

In general, the localization error is higher when the nodes are sparse since the subset of nodes we choose may be lacking in the sense that a node may not have four other nodes that cover it. At higher density, the error should resemble the error faced by other schemes. At a certain density that will provide what we have come to refer to as a “complete subset”, the error will have reached a minimum beyond which it will no longer decrease no matter how the density increases. Compared to the
hybrid scheme (Figure 3), we notice that our algorithm begins with a slightly higher percentage of error at lower density; however, this quickly changes. And while error continues to decrease as we increase the node density in our algorithm, their error percentages are almost constant all throughout since the recursion in their algorithm leads to a propagation of error through the system. As for the AUV-Aided Scheme [10], we notice that their errors fluctuate and are hence unreliable since the error is dependent on a chosen interval for the AUV to transmit signals. For a higher node density, our algorithm far surpasses both in terms of accuracy.

![Figure 2: Localization coverage of different approaches.](image)

Figure 2: Localization coverage of different approaches.

![Figure 3: Localization error of different approaches.](image)

Figure 3: Localization error of different approaches.

C. Communication Overhead

We study the communication overhead relative to node density as compared to the hybrid scheme (Figure 4). On average, our communication cost is less than their communication cost. Although we start out with higher communication cost, our algorithm compensates as mentioned before by decreasing the communication cost after the initialization phase. Also on average the communication cost is higher on low node density since the nodes will continuously try to find a fourth reference point in order to localize themselves. Then, as the node density increases and the subset becomes more “complete”, the communication cost decreases as there will be less need for self-healing algorithms.

![Figure 4: Communication cost of different approaches.](image)

Figure 4: Communication cost of different approaches.

VI. CONCLUSIONS

In this paper, we proposed a reactive localization scheme that is both scalable and distributed. The algorithm consists of three consecutive steps and is capable of self-healing. Analysis and evaluation of our scheme show that it is superior in terms of conserving node energy and hence allowing the system to live longer. It also reduces the communication overhead imposed by other underwater localization algorithms.

REFERENCES


